TABLE OF CONTENTS

PART II: THE GEOCHEMICAL TOOLBOX

hapter 1: Introduction	1
Geochemistry	1
This Book	2
The Philosophy of Science	4
Building Scientific Understanding	4
The Scientist as Skeptic	5
Elements, Atoms, and Chemical Bonds	6
The Periodic Table	6
Electrons and Orbits	7
Some Chemical Properties of the Elements	9
Chemical Bonding	
A Brief Look at the Earth	
Structure of the Earth	
Plate Tectonics and the Hydrologic Cycle	
Earth Materials	
References and Suggestions for Further Reading	
napter 2: Energy, Entropy and Fundamental Thermodynamic Concepts	
The Thermodynamic Perspective	20
Thermodynamic Systems and Equilibrium	20 2∩
Fundamental Thermodynamic Variables	∠0 ??
Faultions of State	∠∠ ⊃1
Ideal Gas I aw	∠± 24
Equations of State for Pool Cases	
Equation of State for Other Substances	20
Tomporature Absolute Zero and The Zeroth Law Of Thermodynamics	
Enorgy and The First Law of Thermodynamics	
Energy and The Trist Law of Thermodynamics	
Work	······2/ 28
Path independence. Exact differentials State functions, and the First Law	
The Second Law and Entropy	
Statement	
Statistical Machanics: A Microscopic Paranactiva of Entropy	
Integrating Easters and Exact Differentials	
Enthology	
Entitialpy	
Constant Volume Heat Canacity	
Constant Process Heat Capacity	
Constant Pressure field Capacity	
Energy Associated with volume Occupied and the Kelationship between C_v and C_p	
Real Capacity of Solids: A Problem in Quantum Physics	
Kelationship of Entropy to Other State Variables	
Additive Nature of Silicate Heat Capacities	
Ine Inira Law and Absolute Entropy	
Statement of the Third Law	
Absolute Entropy	
Calculating Enthalpy and Entropy Changes	
Enthalpy Changes Due to Changes in Temperature and Pressure	
Changes in Enthalpy Due to Reactions and Changes of State	
Entropies of Reaction	
Free Energy	
Helmholtz Free Energy	
Gibbs Free Energy	54
Criteria for Equilibrium and Spontaneity	54
Temperature and Pressure Dependence of the Gibbs Free Energy	55
The Maxwell Relations	

References and Suggestions for Further Reading	
Problems	
Chapter 3: Solutions and Thermodynamics of Multicomponent Systems	61
Introduction	61
Phase Equilibria	61
Some Definitions	61
The Gibbs Phase Rule	64
The Clapeyron Equation	65
Solutions	65
Raoult's Law	67
Henry's Law	68
The Chemical Potential	68
Partial Molar Quantities	68
Definition of Chemical Potential and Relationship to Gibbs Free Energy	69
Properties of the Chemical Potential	69
The Gibbs-Duhem Relation	70
Derivation of the Phase Rule	71
Ideal Solutions	71
Chemical Potential in Ideal Solutions	71
Volume, Enthalpy, Entropy, and Free Energy Changes in Ideal Solutions	72
Real solutions	74
Chemical Potential in Real Solutions	74
Fugacities	75
Activities and Activity Coefficients	76
Excess Functions	
Electrolyte Solutions	80
The Nature of Water and Water-Electrolyte Interaction	80
Some Definitions and Conventions	
Activities in Electrolytes	
Solid Solutions and Their Activities	
Mixing on Site Model	
Local Charge Balance Model	
Equilibrium Constants	
Derivation and Definition	
The Law of Mass Action	
$K_{\rm D}$ Values, Apparent Equilibrium Constants and the Solubility Product	
Henry's Law and Gas Solubilities	
Temperature Dependence of Equilibrium Constant	
Pressure Dependence of Equilibrium Constant	
Practical Approach to Electrolyte Equilibrium	
Choosing Components and Species	
Mass Balance	
Electrical Neutrality	
Equilibrium Constant Expressions	
Oxidation and Reduction	
Redox in Aqueous Solutions	
Redox in Magmatic Systems	10/
References and Suggestions for Further Reduing	
Problems	109
Chapter 4: Applications of Thermodynamics to the Earth	
Introduction	
Activities in Non-Ideal Solid Solutions	113
Mathematical Models of Real Solutions	113
Exsolution Phenomena	116
Thermodynamics and Phase Diagrams	119
The Thermodynamics of Melting	
Thermodynamics of Phase Diagrams for Binary Systems	

	Geothermometry and Geobarometry	125
	Theoretical considerations	125
	Practical Thermobarometers	126
	Thermodynamic Models of Magmas	
	Structure of Silicate Melts	135
	Magma Solution Models	
	Reprise: Thermodynamics of Electrolyte Solutions	141
	Mean Ionic Quantities	141
	Activities in Wish Jonis Strongth Colutions	111
	Activities in Fight Joint Strength Solutions	144
	References and Suggestions for Further Reading	149
	Problems	152
Ch	apter 5: Kinetics: The Pace of Things	155
	Introduction	
	Reaction Kinetics	156
	Elementary and Overall Reactions	156
	Reaction Machanisme	156
	Poaction Pates	150
	Reaction Remeley Reactions	160
	Rates of Complex Reactions.	1
	Steady-State and Equilibrium	165
	Relationships between Kinetics and Thermodynamics	167
	Principle of Detailed Balancing	167
	Enthalpy and Activation Energy	
	Aspects of Transition State Theory	168
	Diffusion	173
	Diffusion Flux and Fick's Laws	173
	The Diffusion Coefficient and Diffusion in Multicomponent Systems	176
	Diffusion in Solids and the Temperature Dependence of the Diffusion Coefficient	181
	Diffusion in Liquids	
	Surfaces, Interfaces, and Interface Processes	
	The Surface Free Energy	184
	The Kelvin Effect	185
	Nucleation and Crystal Crowth	186
	Advention	100
	Ausorphon.	102
	Catalysis	192 105
	Circula Orbitation and Leaching.	195
	Simple Oxides	195
	Silicates	196
	Diagenesis	200
	Compositional Gradients in Accumulating Sediment	200
	Reduction of Sulfate in Accumulating Sediment	
	References and Suggestions for Further Reading	204
	Problems	206
Ch	anter 6. Aquatic Chemistry	210
	Inter dustion	010
	Introduction	
	Acid-Base Reactions	
	Proton Accounting, Charge Balance, and Conservation Equations	211
	The Carbonate System	214
	Conservative and Non-Conservative Ions	217
	Total Alkalinity and Carbonate Alkalinity	218
	Buffer Intensity	221
	Complexation	
	Stability Constants	
	Water-Related Complexes	
	Other Complexes	
	Complexation in Fresh Waters	228
	Dissolution and Precipitation Reactions	220
	Dissolution of Carbonatos by Ground and Surface Waters	200 22∩
	Dissolution of Carbonates by Ground and Sufface Waters	

Solubility of Mg	231
Solubility of SiO_2	
Solubility of Al(OH) ₃ and Other Hydroxides	
Dissolution of Silicates and Related Minerals	
Clavs and Their Properties	
Clav Mineralogy	
Ion-Exchange Properties of Clays	
Mineral Surfaces and Their Interaction With Solutions	
Adsorption	244
Development of Surface Charge and the Electric Double Laver	
References and Suggestions for Further Reading	255
Problems	255
Chapter 7: Trace Elements in Igneous Processes	
Introduction	258
What is a Trace Element?	
Bohavior of the Elements	260
Coldschmidt's Classification	
The Coochemical Periodic Table	
Distribution of Trace Elements Botwoon Co. ovisting Phases	
The Partition Coofficient	
Factors Coverning the Value of Partition Coefficients	
Tamparature and Pressure Dependence of the Partition Coefficient	
Ionia Size and Charge	
Ionic Size and Charge	
Compositional Dependency	
Mineral-Liquid Partition Coefficients for Maric and Ultramaric Systems	
Crystal-Field Effects	
Crystal Field Influences on Transition Metal Partitioning	
Face Element Distribution During Partial Melting	
Equilibrium or Batch Melting	
rracuonal Melung	
Zone Kerining	
Multiphase Solids	
Continuous Melting	
Frace Element Distribution during Crystallization	
Equilibrium Crystallization	
Fractional Crystallization	
In Situ Crystallization	
Crystallization in Open System Magma Chambers	
Summary of Trace Element variations during Melting and Crystallization	
References and Suggestions for Further Reading	
Problems	
Chapter 8: Radiogenic Isotope Geochemistry	
Introduction	318
Physics of the Nuclous and the Structure of Nucloi	310
Nuclear Structure and Energetics	210
The Decay of Evolted and Unstable Nuclei	
Regise of Padiogenia Isotone Coochemistry	
Coochropology	
Decay Systems and Their Applications	
Rh_Sr	200
Sm-Nd	
ош-тиц I 11-Hf	
Lu-111 ₽₀ ∩₀	
La-Ce II Th Dh	
U-111-1 U	

U and Th Decay Series Isotopes	
Isotopes of He and Other Rare Gases	
Cosmogenic and Fossil Isotopes	
³⁶ Cl in hydrology	356
¹⁰ Be in Subduction Zone Studies	357
Cosmic-Ray Exposure Ages of Mateorites	
Foceil Nuclides	358
References and Suggestions for Further Reading	358
Problems	
Chapter 9: Stable Isotope Geochemistry	
Introduction	
Scope of Stable Isotope Geochemistry	
Notation	
Theoretical Considerations	
Equilibrium Isotope Fractionations	
Kinetic Isotope Fractionations	
Isotope Geothermometry	
Isotope Fractionation in the Hydrologic System	
Isotope Fractionation in Biological Systems	
Carbon Isotope Fractionation During Photosynthesis	
Nitrogen Isotope Fractionation in Biological Processes	
Oxygen and Hydrogen Isotope Fractionation by Plants	
Biological Fractionation of Sulfur Isotopes	
Isotopes and Diet: You Are What You Eat	
Paleoclimatology	
The Marine Quaternary δ^{18} O Record and Milankovitch Cycles	
The Record in Glacial Ice	
Soils and Paleosols	
Hydrothermal Systems and Ore Deposits	
Water-Rock Ratios	
Sulfur Isotopes and Ore Deposits	
Stable Isotopes in the Mantle and Magmatic Systems	
Stable Isotopic Composition of the Mantle	
Stable Isotopes in Crystallizing Magmas	
Combined Fractional Crystallization and Assimilation	
Isotopes of Boron and Lithium	
References and Suggestions for Further Reading	
Problems	

CHAPTER 1: INTRODUCTION

CHAPTER 1: INTRODUCTION

1.1 GEOCHEMISTRY

he term "geochemistry" was first used by the Swiss chemist Schönbein in 1838. You might guess, merely from the etymology of the word, that the field of *geochemistry* is somehow a marriage of the fields of *geology* and *chemistry*. That would be a good guess. But just how are chemistry and geology combined within geochemistry; what is the relationship between them? Perhaps the best explanation would be to state that *in geochemistry*, we use the tools of chemistry to solve geological problems; that is, we use chemistry to understand the Earth and how it works. The Earth is part of a family of heavenly bodies, our Solar System, that formed simultaneously and are closely related. Hence, the realm of geochemistry extends beyond the Earth to encompass the entire Solar System. The goals of geochemistry are thus no different from those of other fields of earth science; just the approach differs. On the other hand, while geochemists have much in common with other chemists, their goals differ in fundamental ways. For example, our goals do not include elucidating the nature of chemical bonding or synthesizing new compounds, although these may often be of interest and use in geochemistry. Though geochemistry is a subdiscipline of earth science, it is a very broad topic. So broad in fact that no one can really master it all; geochemists invariably specialize in one or a few aspects, such as atmospheric chemistry, geochemical thermodynamics, isotope geochemistry, marine chemistry, trace element geochemistry, soil chemistry, etc.

Geochemistry has flourished in the quantitative approach that has dominated earth science in the second half of the twentieth century. This quantitative approach has produced greater advances in the understanding of our planet in the last 50 years than in all of prior human history. The contributions of geochemistry to this advance have been simply enormous. Much of what we know about how the Earth and the Solar System formed has come from research on the chemistry of meteorites. Through geochemistry, we can quantify the geologic time scale. Through geochemistry, we can determine the depths and temperatures of magma chambers. Through geochemistry, mantle plumes were recognized. Through geochemistry, we know that sediments can be subducted into the mantle. Through geochemistry, we know the temperatures and pressures at which the various metamorphic rock types form and we can use this information, for example, to determine the throw on ancient faults. Through geochemistry, we know how much and how fast mountain belts have risen. Through geochemistry, we are learning how fast they are eroding. Through geochemistry, we are learning how and when the Earth's crust formed. Through geochemistry, we are learning when the Earth's atmosphere formed and how it has evolved. Through geochemistry, we are learning how the mantle convects. Through geochemistry, we are learning how cold the ice ages were and what caused them. The evidence of the earliest life, 3.8 gigayears (billion, or 10⁹ years, which we will henceforth abbreviate as Ga), is not fossilized remains, but chemical traces of life. Similarly, the tenuous evidence that life existed on Mars about the same time is also largely chemical. Not surprisingly, instruments for chemical analysis have been key part of probes sent to other heavenly bodies, including Venus, Mars, Jupiter. Geochemistry lies at the heart of environmental science and environmental concerns. Problems such as acid rain, the ozone hole, the greenhouse effect and global warming, water and soil pollution are geochemical problems. Addressing these problems requires a knowledge of geochemistry. Similarly, most of our non-renewable resources, such as metal ores and petroleum, form through geochemical processes. Locating new sources of these resources increasing requires geochemical approaches. In summary, every aspect of earth science has been advanced through geochemistry.

Though we will rarely discuss it in this book, geochemistry, like much of science, is very much driven by technology. Technology has given modern geochemists tools that allow them to study the Earth in ways that pioneers of the field could not have dreamed possible. The electron microprobe allows us to analyze mineral grains on the scale of microns in minutes; the electron microscope allows us to view the same minerals on almost the atomic scale. Techniques such as X-ray diffraction, nuclear magnetic resonance, and Raman and infrared spectroscopy allow us to examine atomic ordering and

Geochemistry

CHAPTER 1: INTRODUCTION

bonding in natural materials. Mass spectrometers allow us to determine the age of rocks and the temperature of ancient seas. Ion probes allow us to do these things on micron scale samples. Analytical techniques such as X-ray fluorescence and inductively coupled plasma spectrometry allow us to perform in minutes analyses that would days using "classical" techniques. All this is done with greater precision and accuracy than was possible just a few decades ago. Mega-computers with gigahertz of power and gigabytes of memory allow us to perform in seconds thermodynamic calculations that would have taken years or lifetimes half a century ago; the tera-computers just around the corner will offer us even more power. New instruments and analytical techniques now being developed promise even greater sensitivity, speed, accuracy, and precision. Together, these advances will bring us ever closer to our goal understanding the Earth and its cosmic environment.

1.2 This Book

The intent of this book is to introduce you to geochemistry and to further your understanding of the Earth through it. To do this, we must first acquire the tools of the trade. Every trade has a set of tools. Carpenters have their saws and T-squares; plumbers have their torches and wrenches. Psychologists have their blot tests, physicians their stethoscopes, accountants their balance sheets, geologists have their hammers, compasses, and maps. Geochemists too have a set of tools. These include not only a variety physical tools such as analytical instruments, but interpretative tools that allow them to make sense of the data these instruments produce. The first part of this book, entitled *The Geochemical Toolbox*, is intended to familiarize you with the tools of geochemistry. These include the tools of thermodynamics, kinetics, aquatic chemistry, trace element geochemistry, and isotope geochemistry. Once we have a firm grip on these tools, we can use them to dissect the Earth in the second part of the book, entitled *Understanding the Earth*. We begin at the beginning, with the formation of the Earth and the Solar System. We then work our way upward through the Earth, from the mantle and core, through the crust and hydrosphere, and finally into the atmosphere.

In filling our geochemical toolbox, we start with the tools of physical chemistry: thermodynamics and kinetics. Thermodynamics is perhaps the most fundamental tool of geochemistry; most other tools are built around this one. For this reason, Chapters 2, 3, and 4 are devoted to thermodynamics. Thermodynamics allows us to predict the outcome of chemical reactions under a given set of conditions. In geochemistry we can, for example, predict the sequence of minerals that will crystallize from a magma under given conditions of temperature and pressure. The mineral assemblage of the resulting igneous rock, however, will not be stable at some other temperature and pressure. Thermodynamics allows us to predict the new suite of minerals that replace the original igneous ones. Thus thermodynamics provides enormous predictive power for the petrologist. Since geologists and geochemists are more often concerned with understanding the past than with predicting the future, this might seem to be a pointless academic exercise. However, we can also use thermodynamics in the reverse sense: given a suite of minerals in a rock, we can use thermodynamics to determine the temperature and pressure conditions under which the rock formed. We can also use it to determine the composition of water or magma from which minerals crystallized. This sort of information has been invaluable in reaching our understanding of how the Earth has come to its present condition. We can use this information to determine the amount of uplift experienced by a mountain range, the temperature at which an ore deposit formed, or the composition of ancient seas.

Thermodynamics has an important limitation: it is useful only in equilibrium situations. The rate at which chemical systems achieve equilibrium increases exponentially with temperature. Thermodynamics will be most useful at temperatures relevant to the interior of the Earth, say 500° C and above, because equilibrium will be closely approached in most cases. At low temperatures, that is, temperatures relevant to the surface of the Earth, many geochemical systems will not be in equilibrium and are governed by partly or largely by kinetics, the subject of Chapter 5. Kinetics deals with the rates and mechanisms of reactions. In this chapter, we will also touch upon such topics as diffusion and mineral surfaces. We will see that kinetics is intimately related to thermodynamics.

In Chapter 6, we see how tools of physical chemistry are adapted for use in dealing with natural solutions, the subject of aquatic chemistry. Much of the Earth's surface is covered by water, and water

Geochemistry

CHAPTER 1: INTRODUCTION

usually is present in pores and fractures to considerable depths even on the continents. This water is not pure, but is instead a solution formed by interaction with minerals. In Chapter 6, we acquire tools that allow us to deal with the interactions among dissolved species, including both among themselves and with the solids with which they come in contact. These interactions include phenomena such as dissolution and precipitation, complexation, adsorption and ion exchange. Clays are often the products of water-rock interaction and they have some very interesting chemical properties, so we will have a particularly close look at this group of minerals. The tools of aquatic chemistry are essential to understanding processes such as weathering and precipitation of sedimentary minerals, as well as dealing with environmental problems.

In Chapter 7, we move on to trace element geochemistry. In this chapter we will see that trace elements have provided remarkable insights into the origin and behavior of magmas. Without question, their value to geochemists far outweighs their abundance. There are several reasons for this. Their concentrations vary much more than do those of the more abundant elements, and their behavior tends often to be simpler and easier to treat than that of major elements (a property we will come to know as Henry's Law). Geochemists have developed special tools for dealing with trace elements; the objective of Chapter 7 is to become familiar with them.

Chapters 8 and 9 are devoted to isotope geochemistry. In Chapter 8, we learn that radioactive decay adds the important element of time; radioactivity is nature's clock. By learning to read this clock, we now know the age of the Earth and the continents, and we have gained some perspective on the rate and manner of evolution of the Earth. We can also use the products of radioactive decay, "radiogenic elements", as tracers. By following these tracers much as we would dye in fish tank, we can follow the evolution of a magma, the convection pattern of the mantle, and the circulation of the oceans. The isotopes of another set of elements vary not because of radioactive decay, but because of subtle differences in their chemical behavior. These "stable isotopes" are the subject of Chapter 9. The subtle differences in isotopic abundances of elements such as H, C, N, O, and S have, among other things, revealed the causes of the ice ages, provided insights into the composition of the ancient atmosphere, and reveal the diets of ancient peoples. Stable isotope geochemistry is the last of our geochemical tools.

With our toolbox full, we examine the Earth from the geochemical perspective in the second part of the book. We begin in Chapter 10 by looking at "the big picture": the cosmos and the Solar System. We learn how the chemical elements were formed, and how they, in turn, formed our Solar System and the Earth. We will find the tools of thermodynamic and isotope geochemistry particularly valuable in this Chapter. We will focus particularly closely on meteorites, because the chemistry of these objects provides the best record of the early history of the Solar System. Meteorites also provide essential information about the composition of the Earth as a whole, which will in turn be valuable to us in the following chapter.

In Chapter 11, we begin our inside-out geochemical tour of the Earth. First, we consider the composition of the Earth as a whole, then see how the Earth has differentiated into two major reservoirs: the mantle and core. We pay particular attention to the mantle. Though remote, the mantle is hardly irrelevant. It is important for several reasons. First, it constitutes 1/2 of the mass of the Earth. Second, the reservoirs we are most familiar with, the crust, the hydrosphere, and the atmosphere, have all formed from the mantle. Third, most geologic processes are ultimately a result of processes occurring within the mantle, processes such as convection and melting. In Chapter 12, we return to more familiar territory: the Earth's crust. We will find that geochemistry has provided much of our knowledge of how the crust has formed and how it has differentiated. We will find the tools of isotope and trace element geochemistry particularly useful in our examination of the solid Earth.

The next three chapters focus on processes at the surface of the Earth. Here water is the dominant substance, and the tools of thermodynamics, kinetics, and aquatic chemistry will be of great use. In Chapter 13, we will take a close look at the interaction between water and the Earth's surface, and processes such as weathering and soil formation. We will see how these processes control the chemistry of streams, rivers, and lakes. Life is also an important force in shaping the face of our planet. The chemistry of living organisms is part of biochemistry and not geochemistry, so we will treat intracellular processes only very briefly. However, organisms produce a vast array of chemicals that find their way into

Geochemistry

CHAPTER 1: INTRODUCTION

the physical environment. In Chapter 14, we will examine the role these organic chemicals play in aquatic chemistry. We will also see how these chemicals are transformed into substances of great geological and societal interest: oil, gas, and coal. Most of the water at the surface of the Earth is in the oceans, so we devote Chapter 15 to marine chemistry. We will find that oceans are a fascinating example of an "open" geochemical system, with material constantly flowing both into and out. We will see that in the face of this constant change, geological and biological processes together produce a solution with very uniform concentrations of the major species, but highly variable concentrations of the minor ones.

Before we begin our study of geochemistry, we will devote the remainder of the chapter to reviewing some "fundamentals". First, we briefly examine the philosophy and approach that is common to all science. Then we review the most fundamental aspects of chemistry: how matter is organized into atoms and how these atoms interact to form compounds. Finally, we review a few fundamental aspects of the Earth.

1.3 The Philosophy of Science

This book will concentrate on communicating to you the body of knowledge we call geochemistry. Geochemistry is just part of a much larger field of human endeavor known as science. Science is certainly among humanity's greatest successes; without it, our current civilization would not be possible. Among other things, it would simply not be possible to feed, cloth, and shelter as many people as live today. This phenomenal success is due in large part to the philosophy of science.

Science consists of two parts: the knowledge it encompasses and the approach or philosophy that achieves that knowledge. The goal of all science is to understand the world around us. The arts and humanities also seek understanding. Science differs from those fields as much by its approach and philosophy as by its body of knowledge.

A common approach and philosophy unite the great diversity of fields that we collectively call science. When one compares the methods and tools of a high energy physicist with those of a behavioral biologist, for example, it might at first seem that they have little in common. Among other things, their vocabularies are sufficiently different that each would have difficulty communicating his or her research to the other. In spite of this, they share at least two things. The first is a criterion of "understanding". Both the physicist and the behavioral biologist attempt to explain their observations by the application of a set of rules, which, by comparison to the range of phenomena considered, are both few and simple. Both would agree that a phenomenon is understood if and only if the outcome of an experiment related to that phenomenon can be predicted beforehand by applying those rules to measured variables*. The physicist and biologist also share a common method of seeking understanding, often called the "scientific method".

1.3.1 Building Scientific Understanding

Science deals in only two quantities: *observations* and *theories*. The most basic of these is the *observation*. Measurements, data, analyses, experiments, etc. are all observations in the present sense. An observation might be as simple a measurement of the dip and strike of a rock formation or as complex as the electromagnetic spectrum of a star. Of course, it is possible to measure both the dip of rock strata and a stellar spectrum incorrectly. Before an observation becomes part of the body of scientific knowledge, we would like some reassurance that it is right. How can we tell whether observations are right or not? The most important way to verify an observation is to *replicate* it *independently*. In the strictest sense, 'independent' means by a separate observer, team of observers, or laboratory, and preferably by a different technique or instrument. It is not practicable to replicate every observation in this manner, but critical observations, those which appear to be inconsistent with existing theories or which test the predictions of newly established ones should be, and generally are, replicated. But even replication does not guarantee that an observation is correct.

^{*} Both would admit that chance, or randomness, can affect the outcome of any experiment (though the affect might be slight). By definition, the effect of this randomness cannot be predicted. Where the effects of randomness are large, one performs a large collection, or ensemble, of experiments and then considers the average result.

Geochemistry

CHAPTER 1: INTRODUCTION

Observations form the basis of *theories*. Theories are also called models, hypotheses, etc. *Scientific understanding is achieved by constructing and modifying theories to explain observations*. Theories are merely the products of the imagination of scientists, so we also need a method of sorting out 'correct' theories from 'incorrect' ones. Good theories not only explain existing observations, but make predictions about the outcome of still unperformed experiments or observations. *Theories are tested by performance of these experiments and comparison of the results with the predictions of the theory*. If the predictions are correct, the theory is accepted and the phenomenon considered to be understood, at least until a new and different test is performed. If the predictions are incorrect, the theory is discarded or modified. When trying to explain a newly discovered phenomenon, scientists often reject many new theories before finding a satisfactory one. But long-standing theories that successfully explain a range of phenomena can usually be modified without rejecting them entirely when they prove inconsistent with new observations.

Occasionally, new observations are so inconsistent with a well-established theory that it must be discarded entirely and a new one developed to replace it. Scientific 'revolutions' occur when major theories are discarded in this manner. Rapid progress in understanding generally accompanies these revolutions. Such was the case in physics in the early 20th century when the quantum and relativity theories replaced Newtonian theories. The development of Plate Tectonics in the 1960's and 1970's is an excellent example of a scientific revolution in which old theories were replaced by a single new one. A range of observations including the direction of motion along transform faults, the magnetic anomaly pattern on the sea floor, and the distribution of earthquakes and volcanoes were either not predicted by, or were inconsistent with, classical theories of the Earth. Plate tectonics explained all these and made a number of predictions, such as the age of the seafloor, that could be tested. Thus scientific understanding progresses through an endless cycle of observation, theory construction and modification, and prediction. In this cycle, theories can achieve "acceptance", but can never be proven correct, because we can never be sure that it will not fail some new, future test.

Quite often, it is possible to explain observations in more than one way. That being the case, we need a rule that tells us which theory to accept. When this occurs, the principle is that the theory that explains the greatest range of phenomena in the *simplest* manner is always preferred. For example, the motion of the Sun across the sky is quite simple and may be explained equally well by imaging that the Sun orbits the Earth as visa versa. However, the motions of the planets in the sky are quite complex and require a very complex theory if we assume they orbit the Earth. If we theorize that the Earth and the other planets all orbit the Sun, the motions of the planets become simple elliptical orbits and can be explained by Newton's three laws of motion. The geocentric theory was long ago replaced by the he-liocentric theory for precisely this reason. This principle of simplicity, or elegance, also applies to mathematics. Computer programmers call it the KISS (Keep It Simple, Stupid!) Principle. In science, we can sum it up by saying: *don't make nature any more complex than it already is*.

1.3.2 The Scientist as Skeptic

Though we often refer to "scientific facts", there are no facts in science. A fact, by definition, cannot be wrong. Both observations and theories can be, and sometimes are, wrong. Of course, some observations (e.g., the Sun rises each morning in the East) and theories (the Earth revolves around the Sun) are so oft repeated and so well established that they are not seriously questioned. But remember that the theory that the Sun revolves around the Earth was itself once so well established that it was not seriously questioned.

One of the ways science differs from other fields of endeavor is that *in science nothing is sacred*. It is best to bear in mind the possibility, however remote, that any observation or theory can be wrong. Conversely, we must also accept the possibility that even the wildest observations and theories might be correct: in quantum physics, for example, there is a great range of well-replicated observations that can only be labeled as bizarre (see, for example, Gribbin, 1984). 'Intuition' plays a greater role in science that most scientists might be willing to admit, even though scientific intuition is often very useful. Nevertheless, our intuition is based largely on our everyday experience, which is very limited compared to the range of phenomena that science attempts to understand. As a result, our intuition often deceives us. Sometimes we must put it aside entirely. That a clock will run slower if it moves faster, or

Geochemistry

CHAPTER 1: INTRODUCTION

that an electron can behave both as a wave and a particle, or that continents move great distances are all very counter-intuitive observations, but all are (apparently) correct. Thus *skepticism is one of the keys to good science*. In science, *never totally believe anything, but never totally disbelieve anything either*.

1.4 Elements, Atoms, and Chemical Bonds: Some Chemical Fundamentals

1.4.1 THE PERIODIC TABLE

We'll begin our very brief review of chemical fundamentals with the Periodic Table. In Dmitri Mendeleyev's[‡] day, chemistry and geochemistry were not as distinct as they are today. Chemists were still very much occupied with discovering new elements, and they generally sought them in natural materials. For a variety of reasons, therefore, the Mendeleyev's periodic table provides a good point of departure for us.

Mendeleyev's periodic table of the elements was the sort of discovery that produces revolutions in science. Chemistry had evolved tremendously through the first half of the nineteenth century. Between the publication of Lavoisier's *The Elements of Chemistry*, often considered the first modern text in chemistry, in 1789 and Mendeleyev's 1869 paper, the number of known elements had increased from 23 to 67. The concepts of the atom and the molecule were well established, and role of electromagnetic forces in chemical interactions was at least partly understood. Nevertheless, the structure of atoms, and how this structure governed chemical properties of the atom were to be twentieth century discoveries (though there were some interesting prescient theories). Mendeleyev's great contribution was to show that properties of the elements are a periodic function of atomic weights. Like all good scientific theories, this one made predictions: Mendeleyev was not only able to predict the discovery of then unknown elements, such as B, Sc, Ga, and Ge, but also their characteristics and probable mode of discovery. The periodic table led the way not only to the discovery of the remaining elements, but also to understanding the fundamental controls on chemical behavior.

Figure 1.1 shows the periodic table as we know it today. Like most theories, Mendeleyev's has gone through some revision since it was first proposed. Most importantly, we now organize the periodic table based on *Atomic Number* rather than atomic weight. The atomic number of an element is its most important property, and is determined by the number of protons in the nucleus (thus the terms *atomic numberand proton number* are synonymous). The number of protons in turn determines both the number of electrons and how these electrons are organized.

The mass of an atom is a function of both the proton number and the *neutron number*, i.e., the number of neutrons in the nucleus^{*}. Generally, several possible numbers of neutrons can combine with a given number of protons to form a stable nucleus (we will discuss nuclear stability in greater detail in Chapter 8). This gives rise to different isotopes of the same element, i.e., atoms that have the same atomic number but different masses. For example, helium has 2 stable isotopes: ³He and ⁴He. Both ³He and ⁴He⁺ have 2 protons (and a matching number of electrons), but ⁴He has 2 neutrons while ³He has only 1.

[‡] Dmitri Ivanovich Mendeleyev was born in Tobolsk, Russia in 1834. He became professor of chemistry at St Petersburg in 1866. His periodic table was the sort of discovery that noble prizes are awarded for, but it came before the prize was established. He was honored, however, by having element number 101, medelevium, named for him. Mendeleyev died in 1906.

^{*} The neutron mass and proton mass are almost identical.

⁺ By convention, the mass number, which is the sum of protons and neutrons in the nucleus, of an isotope is written as a preceeding superscript. However, for historical reasons, one says "helium–4". Note also that the atomic number or proton number can be readily deduced from the chemical symbol (atomic number of He is 2). The neutron number can be found by subtracting the proton number from the mass number. Thus the symbol "⁴He' gives a complete description of the nucleus of this atom.

CHAPTER 1: INTRODUCTION



Figure 1.1. The Periodic Table showing symbols and atomic numbers of naturally occurring elements. Many older periodic tables number the groups as IA-VIIIA and IB-VIIB. This version shows the current IUPAC Convention.

The *atomic weight* of an element depends on both the masses of its various isotopes and on the relative abundances of these isotopes. This bedeviled nineteenth century chemists. William Prout (1785-1850), an English chemist and physiologist, had noted in 1815 that the densities of a number of gases were integer multiples of the density of hydrogen (e.g., 14 for nitrogen, 16 for oxygen). This law appeared to extend to many elemental solids (e.g., 12 for C, 28 for Si) as well, and it seemed reasonable that this might be a universal law. But there were puzzling exceptions. Cl, for example, has an atomic weight of 35.45 times that of hydrogen. The mystery wasn't resolved until Thompson demonstrated the existence of 2 isotopes of Ne in 1918. The explanation is that while elements such as H, N, O, C, and Si consist almost entirely of a single isotope, and thus have atomic weights very close to the mass number of that isotope, natural Cl consists of about 75% ³⁵Cl and 25% ³⁷Cl[‡].

1.4.2 Electrons and Orbits

We stated above that the atomic number of an element is its most important property. This is true because the number of electrons is determined by atomic number, as it is the electronic structure of an atom that largely dictates it chemical properties. The organization of the elements in the periodic table reflects this electronic structure.

The electronic structure of atoms, and indeed the entire organization of the periodic table, is determined by quantum mechanics and the quantization of energy, angular momentum, magnetic moment, and spin of electrons. Four quantum numbers, called the principal, azimuthal, magnetic, and spin quantum numbers and conventionally labeled n, l, m, and m_s , control the properties of electrons associated with atoms. The first of these, n, which may take values 1, 2, 3, ..., determines most of the

[‡] The actual mass of a atom depends on the number of electrons and the nuclear binding energy as well as the number of protons and neutrons. However, the mass of the electron is over 1000 times less than the mass of the proton and neutron, which are about equal, and the effect of nuclear binding energy on mass was too small for 19th century chemists to detect.

Geochemistry

CHAPTER 1: INTRODUCTION

electron's energy as well as mean distance from the nucleus. The second, *l*, which has values 0, 1, 2, ... *n*-1, determines the total angular momentum and the shape of the orbit. The third, *m*, which may have values *-l*, ...0...*l*, determines the *z* component of angular momentum and therefore the orientation of the orbit. The fourth, m_s , may have values of $-\frac{1}{2}$ or $+\frac{1}{2}$ and determines the electron's spin. The first three quantum numbers result in the electrons surrounding the nucleus being organized into *shell*, *subshells*, and *orbitals**. The *Pauli Exclusion Principle* requires that no two electrons in an atom may have identical values of all four quantum numbers. Because each orbital corresponds to a unique set of the first 3 quantum numbers, only 2 electrons may occupy a given orbital, and then only if they have different spins. In Chapter 8 we will see that the properties of the nucleus are also dictated by quantum mechanics, and that the nucleus may also be thought of as having a shell structure.

Each shell corresponds to a different value of the principal quantum number. The periodic nature of chemical properties reflects the filling of successive shells as additional electrons (and protons) are added. Each shell corresponds to a 'period', or row, in the periodic table. The first shell (the K shell) has one subshell, the *1s*, consisting of a single orbital (with quantum numbers n=1, l=0, m=0. The *1s* orbital accepts up to 2 electrons. Thus period 1 has 2 elements: H and He. If another electron and proton are added, the electron is added to the first orbital, *2s*, of the next shell (the L shell). Such a configuration has the chemical properties of Li, the first element of period 2. The second shell has 2 subshells, *2s* (corresponding to l=0) and *2p* (corresponding to l=1). The *p* subshell has 3 orbitals (which correspond to values for *m* of -1, +1, and 0), p_x , p_y , and p_z , so the L shell can accept up to 8 electrons. Thus period 2 has 8 elements.

There are some complexities in the filling of orbitals beyond the L shell. In the M shell, which corresponds to period 3, there is the possibility of putting electrons in the *d* subshell (l=2), but this is energetically less favorable than electrons going in the subsequent shell. Thus the 3*d* subshell is vacant in period 3 element in their ground states, and in the first 2 elements of period 4. Only when the 4*s* orbital

is filled do electrons begin to fill the 3*d* orbitals. The 5 3*d* orbitals are filled as one passes up the first transition series metals, Sc through Zn. This results in some interesting chemical properties, which we will consider in Chapter 7. Similarly, the second and third transition series metals correspond to filling of the 4*d* and 5*d* orbitals. The Lanthanide and Actinide rare earth elements correspond to the filling of the 4f and 5f shells are filled (again resulting in some interesting properties, which we will consider subsequently). The predicted sequence in which orbitals are filled and their energy levels are shown in Figure 1.2. Figure 1.3 shows the electronic configuration of the elements.



Figure 1.2. The predicted sequence of orbital energies for electrons in atoms. *S* levels can hold 2 electrons, *p*, *d*, and *f* can hold 6, 10, and 14 respectively.

^{*} It is often convenient to think of the electrons orbiting the nucleus much as the planets orbit the Sun. This analogy has its limitations. The electron's position can not be precisely specified as can a planet's. In quantum mechanics, the Schrödinger Wave function, ψ (or more precisely, ψ^2) determines the probability of the electron being located in a given region about the atom. As an example of failure of the classical physical description of the atom, consider an electron in the *1s* orbital. Both quantum number specifying angular momentum, *l* and *m*, are equal to 0, and hence the electron has 0 angular momentum, and hence can not be in an orbit in the classical sense.

CHAPTER 1: INTRODUCTION

H 15 ¹																ſ	He 15 ²
Li 25 ¹	Be 25 ²]										$B_{2s^2 2p^1}$	$C_{2s^{2}2p^{2}}$	$N_{2s^2 2p^3}$	$0_{25^2 2p^4}$	F 25 ² 2p ⁴	$Ne_{2s^22p^6}$
Na 35 ¹	Mg 35 ²	1									Ì	Al 35 ² 3p ¹	Si 35 ² 3p ²	P 35 ² 3p ³	S 35 ² 3p ⁴	Cl 35 ² 3p ⁴	Ar 35 ² 3p ⁶
K 4s ¹	Ca 45 ²	Sc 4s ² 3d ¹	Ti 45 ² 3d ²	V 45 ² 3d ³	Cr 4s ² 3d ⁴	Mn 4s ² 3d ⁵	Fe 45 ² 3d ⁶	Co 4 <i>s</i> ² 3 <i>d</i> ⁷	Ni 45 ² 3d ⁸	Cu 4s ² 3d ⁹	Zn 45 ² 3d ¹⁰	$\operatorname{Ga}_{4s^24p^1}$	Ge 45 ² 4p ²	$\operatorname{As}_{4s^24p^3}$	Se 45 ² 4p ⁴	Br 45 ² 4p ⁴	K 45 ² 4p ⁶
Rb 55 ¹	Sr 55 ²	Y 5 <i>s</i> ² 4 <i>d</i> ¹	Zr 5 <i>s</i> ² 4 <i>d</i> ²	Nb 55 ² 4d ³	Mo 5 <i>s</i> ² 4 <i>d</i> ⁴	Tc 55 ² 4d ⁵	Ru 55 ² 4d ⁶	Rh 55 ² 4d ⁷	Pd 55 ² 4d ⁸	Ag 55 ² 4d ⁹	Cd 5 <i>s</i> ² 4 <i>d</i> ¹⁰	In 55 ² 5p ¹	Sn 55 ² 5p ²	Sb 55 ² 5p ³	Te 55 ² 5p ⁴	 55 ² 5p ⁴	Хе ^{55²5р⁶}
Cs 65 ¹	Ba 65 ²	La 65 ² 5d ¹	Hf 65 ² 5d ²	Ta 65 ² 5d ³	W 6s ² 5d ⁴	Re 65 ² 5d ⁵	Os 6s ² 5d ⁶	Ir 65 ² 5d ⁷	Pt 65 ² 5d ⁹	Au 65 ² 5d ¹⁰	Hg 65 ² 5d ¹⁰	TI 65 ² 6p ¹	Pb 65 ² 6p ²	Bi 65 ² 6p ³	Ро 65 ² 6р ⁴	At 65 ² 6p ⁴	Rd 65 ² 6p ⁶
Fr 751	Ra 75 ²	Ac 75 ² 6d ¹	۱.	<u>.</u>								•	•	•		<u> </u>	
		`\	La	Ce	$\Pr_{6s^2 f^3}$	$\frac{Nd}{6s^2 df^2}$	Pm	Sm	5 Eu	Gd	Tb	Dy	Ho	$Er_{6s^2 4f^4}$	Tm_{65^2}	Yb	Lu
		١	AC 75 ² 6d	Th 75 ² 6d	Pa Pa 75°6d '5f	U ² 75 ² 6d ¹ 5f	Np 75°6d ¹ 5f	Pu 4 75 ² 5f ⁶	0.5 49	0.5 54 49	0.5 49	0.5 49		0.5 4)	0.5 49		<i>55 Ju</i> 4j

Figure 1.3. The Periodic Table of naturally occurring elements showing the electronic configuration of the elements. Only the last orbitals filled are shown, thus each element has electrons in the orbitals of all previous group 18 elements (noble gases) in addition to those shown. Superscripts indicate the number of electrons in each subshell.

1.4.3 Some Chemical Properties of the Elements

It is only the most loosely bound electrons, those in the outermost shells, that participate in chemical bonding, so elements sharing a similar outermost electronic configuration tend to behave similarly. Elements within the same column, or *group*, share outer electronic configurations and hence behave in a similar manner. Thus the elements of group 1, the *alkalis*, all have 1 electron in the outermost *s* orbital, and behave in a similar manner. The group 18 elements, the *noble*, *or rare*, *gases*, all have a filled *p* subshell, and behave similarly.

Let's now consider several concepts that are useful in describing the behavior of atoms and elements: *ionization potential, electron affinity,* and *electronegativity*. The *First Ionization Potential* of an atom is the energy required to remove (i.e., move an infinite distance away) the least tightly bound electron. This is energy absorbed by the electron in reactions such as:

$$Na \rightarrow Na^+ + e^-$$
 1.1

The *Second Ionization Potential* is the energy required to remove a second electron, etc. The first ionization potential of the elements is illustrated in Figure 1.4.

The *electron affinity* is the energy given up in reactions such as:

$$F + e^- \rightarrow F^-$$
 1.2

Electronegativity is another parameter that is often used to characterize the behavior of the elements. It is a relative, unitless quantity determined from the differences in bond energy between an A-B molecule and the mean energies of A-A and B-B molecules. *Electronegativity quantifies the tendency of an element to attract a shared electron when bonded to another element.* For example, F has a higher electronegativity than H (the values are 3.8 and 2.5 respectively), thus the bonding electron in hydrogen fluoride, HF, is more likely to be found in the vicinity of F than of H. It is also useful in characterizing the nature of chemical bonds between elements, as we shall see in a subsequent section. Electronegativities of the elements are shown in Figure 1.5.

CHAPTER 1: INTRODUCTION



Figure 1.4. First ionization potential of the elements.

In general, first ionization potential, electron affinity, and electronegativities, increase from left to right across the periodic table, and to a less degree from bottom to top. This reflects the shielding of outer electrons, particularly those in *s* orbitals, by inner electrons, particularly those in *p* orbitals, from the charge of the nucleus. Thus the outer 3s electron of neutral sodium is effectively shielded from the nucleus and is quite easily removed. On the other hand, the 2p orbitals of oxygen are not very effectively shielded, and it readily accepts 2 additional electrons. With the addition of these 2 electrons, the 2p orbital is filled and the 3s orbital effectively shielded, so there is no tendency to add a third electron. With the outer *p* (and *s*) orbitals filled, a particularly stable configuration is reached. Thus Ne has little tendency to either add or give up an electron.

The number of electrons that an element will either give up or accept is known as its *valence*. For elements in the 'wings' of the periodic table (i.e., all except the transition metals), valence is easily determined simply by counting how far the element is horizontally displaced from Group 18 in the periodic table. For Group 18, this is 0, so these elements, the noble gases, have 0 valence. For group 1 it is 1, so these elements have valence of +1; for group 17 it is -1, so these elements have valence of -1, etc. Valence of the transition metals is not so simply determined, and these elements can have more than 1 valence state. Most, however, have valence of 2 or 3, though some, such as U, can have valences as high as 6.

A final characteristic that is important in controlling chemical properties is *ionic radius*. This is deduced from bond length when the atom is bonded to one or more other atoms. Positively charged atoms, or *cation*, have smaller ionic radii than do negatively charged atoms, or *anions*. Also, ionic radius decreases as charge increases. This decrease is due both to loss of electrons and to shrinking of the orbits of the remaining electrons. The latter occurs because the charge of the nucleus is shared by fewer CHAPTER 1: INTRODUCTION

H 2.1]															Γ	He
Li 1.0	Be 1.5]										B 2.0	C 2.5	N 3.0	0 3.5	F 4.0	Ne
Na 0.9	Mg 1.2											Al 1.5	Si 1.8	P 2.1	S 2.5	Cl 3.0	Ar
K 0.8	Ca 1.0	Sc 1.3	Ti 1.5	V 1.6	Cr 1.6	Mn 1.5	Fe 1.8	Co 1.9	Ni 1.9	Cu 1.9	Zn 1.6	Ga 1.6	Ge 1.8	As 2.0	Se 2.5	Br 2.8	К
Rb 0.8	Sr 1.0	Y 1.2	Zr 1.4	Nb 1.6	Mo 1.8	Tc 1.9	Ru 2.2	Rh 2.2	Pd 2.2	Ag 1.9	Cd 1.7	In 1.7	Sn 1.8	Sb 1.9	Te 2.1	 2.5	Xe
Cs 0.7	Ba <i>0.9</i>	La 1.0	∖Hf <i>₹.3</i>	Та 1.5	W 1.7	Re 1.9	Os 2.2	lr 2.2	Pt 2.2	Au 2.4	Hg 1.9	TI 1.8	Pb 1.9	Bi 1.9	Po 2.2	At 2.2	Rd
Fr <i>0.7</i>	Ra 0.9	Ac 1.1													!	<u>L</u>	
		`\	La 1.0	Ce 1.0	Pr 1.0	Nd 1.0	Pm 1.0	Sm 1.0	Eu 1.1	Gd 1.1	Tb 1.1	Dy 1.1	Ho 1.1	Er 1.1	Tm 1.2	Yb 1.2	Lu 1.2
			Ac	Th	Pa	l u	Np	Ри								1	

1.1	1.3	1.4	ĭ.4	1.4	1.4			

Figure 1.5. Electronegativities of the elements.

electrons and hence has a greater attractive force on each. In addition, ionic radius increases downward in each group in the periodic table, both because of addition of electrons to outer shells and because these outer electrons are increasingly shielded from the nuclear charge by the inner ones. Ionic radius is important in determining important geochemical properties such substitution in solids, solubility, and diffusion rates. Large ions need to be surrounded, or *coordinated*, by a greater number of oppositely charged ions than do smaller ones. The ionic radii of the elements are illustrated in Figure 1.6.

We can now summarize a few of the more important chemical properties of the various groups in the periodic table. Group 18, does not participate in chemical bonding in nature, hence the term 'noble gases'. Group 1 elements, the alkalis, readily accept an electron (they are *electropositive*) and hence are highly reactive. They tend to form *ionic bonds* rather than *covalent* ones and hence tend not to form bonds that are as strong as other elements. They tend to be quite soluble in aqueous solutions. Because they have only a +1 charge, their ionic radii tend to be larger than other cations. Group 2 elements, the alkaline earths, have these same characteristics, but somewhat moderated. Group 17 elements, the halogens, are highly electronegative and readily accept an electron, are highly reactive, form ionic bonds, and are quite soluble. Their ionic radii tend to be larger than more highly charged anions. Elements of Groups 13-16 tend to form bonds that are predominantly covalent. As a result, they tend to be less reactive and less soluble (except where they form soluble radicals, such as SO_4^{-}) than Group 1, 2, and 17 elements. Finally, the transition metals are a varied lot. Many form strong bonds (generally with O in nature) and are fairly insoluble, particularly the highly charged ones. Some, the 'noble metals' (Ru, Rh, Pd, Os, Ir, Pt) in particular, tend to be very unreactive and insoluble. The rare earths are of interest because all have 2 electrons in the 6s outer orbital, and vary only by the number of electrons in the 4f shell. Their bonding behavior is thus quite similar. They vary systematically in ionic radius, which makes them of great geochemical interest.

1.4.4 CHEMICAL BONDING

1.4.4.1 Covalent, Ionic, and Metal Bond

Except for the noble gases, atoms rarely exist independently: they are generally bound to other atoms in molecules, crystals, or ionic radicals. Atoms bind to one another through transfer or sharing of electrons, or through electrostatic forces arising from uneven distribution of charge in atoms and molecules. A bond that results from the total transfer of electrons from one atom to another is known as an *Ionic Bond*. A good example is the bond between Na and Cl in a halite crystal. In this case, the Na atom

CHAPTER 1: INTRODUCTION



Ionic Radii



(the electropositive element) gives up an electron, becoming positively charged, to the Cl atom (the electronegative element), which becomes negatively charged. Electrostatic forces between the Na⁺ and the Cl⁻ ions hold the ions in place in the crystal. When electrons are shared between atoms, such as in the H₂O molecule or the SiO₄⁴⁻ radical, the bond in known as covalent. In a *Covalent Bond*, the outer electrons of the atoms involved are in hybrid orbits that encompass both atoms.

The ideal covalent and ionic bonds represent the extremes of a spectrum: most bonds are neither wholly covalent nor wholly ionic. In these intermediate cases, the bonding electrons will spend most, but not all, of their time associated with one atom or another. Electronegativity is useful in describing the degree of ionicity of a bond. Though there is no clear distinction between a covalent and ionic bond, a bond is considered ionic when the difference in the electronegativity of the two atoms involved is greater than 2. In Figure 1.5, we see that metals (generally those elements on the left hand side of the periodic table) tend to have low electronegativities while the non-metals (those elements on the right) have high electronegativities. Thus bonds between metals and non-metals (e.g., NaCl) will be ionic while those between non-metals (e.g., CO_2) will be covalent, as will bonds between two like atoms (e.g., O_2).

Another type of bond occurs in pure metal and metal alloy solids. In the metallic bond, valence electrons are not associated with any single atom or pair of atoms; rather they are mobile and may be found at any site in the crystal lattice. Since metals rarely occur naturally (iron meteorites and the Earth's core are notable exceptions), this type of bond is less important in geochemistry than other bonds.

Ionically bonded compounds tend to be good conductors of heat and electricity and have high melting temperatures. They also tend to be hard, brittle and highly soluble in water. Covalently bonded compounds tend to be good conductors of heat, but not of electricity. They are typically harder and less brittle than ionic solids but less soluble. In molecular solids, such as ice, atoms within the molecule are covalently bonded. The molecules themselves, which occupy the lattice points of the crystal, are bonded to each other through van der Waals and/or hydrogen bonds. Such solids are comparatively weak and soft and generally have low melting points.

Molecules in which electrons are unequally shared have an asymmetric distribution of charge and are termed *polar*. A good example is the hydrogen chloride molecule. The difference in electronegativity between hydrogen and chlorine is 0.9, so we can predict that bonding electron will be shared, but associated more with the Cl atom than with the H atoms in HCl. Thus the H atom will have a partial positive charge, and the Cl atom a partial negative charge. Such a molecule is said to be a *dipole*. The *dipole momen*, which is the product of one of the charges (the two charges are equal and opposite) times the distance between the charges, is a measure of the asymmetric distribution of charge. Dipole moment is usually expressed in debye units (1 debye = 3.3356×10^{-34} coloumb-meters).

Geochemistry

CHAPTER 1: INTRODUCTION

1.4.4.2 Van der Waals Interactions and Hydrogen Bonds

Covalent and ionic bonds account for the majority of bonds between atoms in molecules and crystals. There are two other interactions that play a lesser role in interactions between atoms and molecules, van der Waals interactions and hydrogen bonds. These are much weaker but nevertheless play an important role in chemical interactions, particularly where water and organic substances are involved.

Van der Waals interactions arise from asymmetric distribution of charge in molecules and crystals. There are three sources for van der Waals interactions: *dipole-dipole attraction, induction,* and *London dispersion* forces. As we noted above, many molecules, including water, have permanent dipole moments. When two polar molecules encounter each other, they will behave much as two bar magnets: they will tend to orient themselves so that the positive part of one molecule is closest to the negative part of another (Figure 1.7a). This results in a net attractive force between the two molecules. When the distance between molecules is large compared to the distance between charges within molecules, the energy of attraction can be shown to be:

$$U_{D-D} = -\frac{2}{3} \frac{\mu^4}{r^6} \frac{1}{kT}$$
 1.3

where U_{D-D} is the interaction energy, μ is the dipole moment, T is temperature (absolute, or thermodynamic temperature, which we will introduce in the next chapter), k is a constant (Boltzmann's constant, which we shall also meet in the next chapter), and r is distance. We don't want to get lost in equations at this point; however, we can infer several important things about dipole-dipole interactions just from a quick glance at it. First, the interaction energy depends inversely on the sixth power of distance. Many important forces, such as electromagnetic and gravitational forces, depend on the inverse square of distance. Thus we may infer that dipole-dipole forces become weaker with distance very rapidly. Indeed, they are likely to be negligible unless the molecules are very close. Second, the interaction energy depends on the forth power of the dipole moment, so that small differences in dipole moment will result in large differences in interaction energy. For example, the dipole moment of water (1.84 debyes) is less than twice that of HCl (1.03 debyes), yet the dipole interaction energy between two water molecules (716 J/mol) is nearly 10 times as great as that between two HCl molecules (72.24 J/mol) at the same temperature and distance (298 K and 5Å). Finally, we see that dipole interaction energy will decrease with temperature.

Dipole molecules may also polarize electrons in a neighboring molecule and distort their orbits in such a way that their interaction with the dipole of the first molecule is attractive. This is known as the

induction effect (Figure 1.7b). The induction energy also depends on the inverse sixth power of intermolecular distance, but only on the square of the dipole moment of the molecules involved. In addition, another parameter, the polarizability of a molecule, is also needed to describe this effect. In general, the attraction arising from induction is less important than from dipole-dipole interaction. However, because it depends only on the square of dipole moment, the induction attraction can be larger than the dipoledipole attraction for some weakly dipolar molecules.

Finally, van der Waals forces can also occur as a consequence of fluctuations of charge distribution on molecules that occur on time scales of 10⁻¹⁶ seconds. These are known as London dispersion forces. They



b. The Induction Effect

Figure 1.7. Van der Waals interactions arise because of the polar nature of some molecules. Illustrated here are (a) dipole-dipole interactions, which occur when two dipolar molecules orient themselves so oppositely charged sides are closest, and (b) the induction effect, which arise when the electron orbits of one molecule are perturbed by the electromagnetic field of another molecule.

Geochemistry

CHAPTER 1: INTRODUCTION

arise when the instantaneous dipole of one molecule induces a dipole in a neighboring molecule. As was the case in induction, the molecules will orient themselves so that the net forces between them are attractive.

The total energy of all three types of van der Waals interactions between water molecules is about 380 J/mol assuming an intermolecular distance of 5 Å and a temperature of 298 K (25° C). Though some interaction energies can be much stronger (e.g., CCl_4 , 2.8 kJ/mol) or weaker (1 J/mol for He), an energy of a few hundred Joules per mole is typical of many substances. By comparison, the hydrogen-oxygen bond energy for each H—O bond in the water molecule is 46.5 kJ/mol. Thus van der Waals interactions are quite weak compared with typical intramolecular bond energies.

The *Hydrogen Bond* is similar to van der Waals interactions in that it arises from nonsymmetric distribution of charge in molecules. However, it differs from van der Waals interactions in a number of ways. First, it occurs exclusively between hydrogen and strongly electronegative atoms, namely oxygen, nitrogen and fluorine. Second, it can be several orders of magnitude stronger than van der Waals interactions, though still weak by comparison to covalent and ionic bonds. The exact nature of hydrogen bonds is not completely understood. They arise principally from electrostatic interactions. In the water molecule, binding between oxygen and hydrogen results in hybridization of *s* and *p* orbitals to yield two bonding orbitals between the O and two H atoms and 2 non-binding sp^3 orbitals on the oxygen. The latter are prominent on the opposite side of the O from the hydrogens. The hydrogen in one water molecule, carrying a net positive charge, is attracted by the non-binding sp^3 electrons of the oxygen of another water molecule, forming a hydrogen bond with it (Figure 1.8).

Hydrogen bonds typically have energies in the range of 20-40 kJ/mol. These are much higher than expected for electrostatic interactions alone and indeed approach values similar to intra-molecular bond energies. Thus there is the suspicion that the some degree of covalency is also involved in the hydrogen bond. That is to say that the non-binding electrons of oxygen are to some degree shared by with the hydrogen in another molecule. Hydrogen bonds are perhaps most important in water, where they account for some of the extremely usual properties of this compound, such as its high heat of vaporization, but they can also be important in organic molecules and are present in HF and ammonia as well.

1.5 A Brief Look at the Earth



1.5.1 Structure of the Earth

The Earth consists of 3 principal layers: the core, the mantle, and the crust (Figure 1.9). The core, roughly 3400 km thick and extending about half way to the surface, consists of Fe-Ni alloy and can be subdivided into an inner and outer core. The outer core is liquid while the inner core is solid. The mantle is about 3000 km thick and accounts for about 2/3 the mass of the Earth; the core accounting for the other 1/3. The crust is quite thin by comparison, nowhere thicker than 100 km and

Figure 1.8. Hydrogen bonding between water molecules. Hydogen positions shown as red; sp^3 hybrid orbitals in oxygen shown as dark gray. The δ + and δ - indicate partial positive and negative charges respectively.

Geochemistry

CHAPTER 1: INTRODUCTION

usually much thinner. Its mass is only about 0.5% of the mass of the Earth. There are 2 fundamental kinds of crust: oceanic and continental. Ocean crust is thin (about 6 km) and is nowhere older than about 200 million years. The continental crust is thicker (about 35-40 km thick on average) and relatively permanent.

Both the crust and the mantle consist principally of silicates. The mantle is comparatively rich in iron and magnesium, so ferromagnesian silicates, such as olivine and pyroxenes, dominate. Rocks having these compositional characteristics are sometimes called *ultramafic*. The continental crust is poor in iron and magnesium, and aluminosilicates such as feldspars dominate. Rocks of this composition are sometimes referred to as *felsic*. The oceanic crust is intermediate in composition between the mantle and continental crust and has a *mafic* composition, consisting of a roughly 50-50 mix of ferromagnesian minerals and feldspar. These differences in composition lead to differences in density, which are ultimately responsible for the layering of the Earth, the density of each layer decreasing outward. The continental crust is the least of dense of these layers. The fundamental reason why continents stick out above the oceans is that continental crust is less dense than oceanic crust.

1.5.2 Plate Tectonics and the Hydrologic Cycle

Two sources of energy drive all geologic processes: solar energy and the Earth's internal heat. Solar energy drives atmospheric and oceanic circulation, and with them, the hydrologic cycle. In the hydrologic cycle, water vapor in the atmosphere precipitates on the land as rain or snow, percolates into the soil and, through the action of gravity, makes it way to the oceans. From the oceans, it is evaporated into the atmosphere again and the cycle continues. The hydrologic cycle is responsible for two very important geologic processes: weathering and erosion. Weathering, a topic we will consider in more detail in Chapter 13, causes rocks to breakdown into small particles and dissolved components. The particles and dissolved matter are carried by the flow of water (and more rarely by wind and ice) from high elevation to areas of low elevation. Thus the effect of the hydrologic cycle is to level the surface of the planet.

The Earth's internal heat is responsible for tectonic processes, which tend to deform the surface of the planet, producing topographic highs and lows. The internal heat has two parts. Some fraction of the heat, estimated to be between 25% and 75%, originated from the gravitational energy released when the Earth formed. The other fraction of internal heat is produced by the decay of radioactive elements, principally uranium, thorium, and potassium, in the Earth. The Earth's internal heat slowly decays over geologic time as it migrates to the surface and is radiated away into space. It is this migration of heat out of the Earth that drives tectonic processes. Heat causes both the outer core and the mantle to convect, as hot regions rise and cold regions sink. Convection within the outer core gives rise to the Earth's magnetic field, and may have other, as yet not understood, geologic consequences. Convection in the mantle is responsible for deformation of the Earth's crust as well as volcanism.



Figure 1.9. The Earth in cross-section. The outer rocky part of the planet, the mantle and crust, consists principally of silicates and is 2885 km thick. The core, divided into a liquid outer core and a solid inner core, consists of iron-nickel alloy and is 3486 km thick.

Geochemistry

CHAPTER 1: INTRODUCTION

The great revolution in earth science in the 1960's centered on the realization that the outer part of the Earth was divided into a number of "plates" that moved relative to one and other. Most tectonic processes, as well as most volcanism, occur at the boundaries between these plates. The outer part of the Earth, roughly the outer 100 km or so, is cool enough (<1000 ° C) that it is rigid. This rigid outer layer is known as the lithosphere and comprises both the crust and the outermost mantle (Figure1.10). The mantle below the lithosphere is hot enough (and under sufficient confining pressure) that it flows, albeit extremely slowly, when stressed. This part of the mantle is known as the asthenosphere. Temperature differences in the mantle create stresses that produce convective flow. It is this flow that drives the motion of the lithospheric plates. The motion of the plates is extremely slow, a few tens of centimeters per year at most and generally much less. Nevertheless, on geologic time scales they are sufficient to continually reshape the surface of the Earth, creating the Atlantic Ocean, for example, in the last 200 million years.

Rather than thinking of plate motion as being driven by mantle convection, it would be more correct to think of plate motion as part of mantle convection. Where plates move apart, mantle rises to fill the gap. As the mantle does so, it melts. The melt rises to the surface as magma and creates new oceanic crust at volcanos along mid-ocean ridges (Figure 1.10). Mid-ocean ridges, such as the East Pacific Rise and the Mid-Atlantic Ridge, thus mark divergent plate boundaries. As the oceanic crust moves away from the mid-ocean ridge it cools, along with the mantle immediately below it. This cooling produces a steadily thickening lithosphere. As this lithosphere cools, it contracts and its density increases. Because of this contraction, the depth of the ocean floor increases away from the mid-ocean ridge. When this lithosphere has cooled sufficiently, after 100 million years or so, it becomes more dense than the underlying asthenosphere. The lithosphere may then sink back into the mantle in a process known as *subduction.* As the lithosphere sinks, it creates deep ocean trenches, such as the Peru-Chile Trench, or the Marianas Trench. Chains of volcanos, known as island arcs, invariably occur adjacent to these deep sea trenches. The volcanism occurs as a result of dehydration of the subducting oceanic crust and lithosphere. Water released from the subducting oceanic crust rises into the overlying mantle, causing it to melt. The island arcs and deep sea trenches are collectively called *subduction zones*. Subduction zones thus mark *convergent plate boundaries*. It is primarily the sinking of old, cold lithosphere that drives the motion of plates. Thus the lithosphere does not merely ride upon convecting mantle, its motion is actually part of mantle convection.

The density of the continental crust is always lower than that of the mantle, regardless of how cold the crust becomes. As a result, it cannot be subducted into the mantle. The Indian-Eurasian plate boundary is a good example of what happens when two continental plates converge. Neither plate readily subducts and the resulting compression has produced, and continues to uplift, the Himalayan



Figure 1.10. Cross section of the Earth illustrating relationships between lithosphere and asthenosphere and plate tectonic processes. Oceanic crust and lithosphere are created as plates diverge at mid-ocean ridges and are eventually subducted back into the mantle. Continental lithosphere is thicker and lighter than oceanic lithosphere and not easily subducted.

Geochemistry

CHAPTER 1: INTRODUCTION

Mountains and the Tibetan Plateau. This area of the continental crust is not only high it is also deep. The crust beneath this region extends to depths of as much as 100 km, nearly three times the average crustal thickness. Rocks within this thickened crust will experience increased temperatures and pressures, leading to *metamorphism*, a process in which new minerals form in place of the original ones. In the deepest part of the crust, melting may occur, giving rise to granitic magmas, which will then intrude the upper crust.

The topographically high Himalayas are subject to extremely high rates of erosion, and the rivers draining the area carry enormous quantities of sediment. These are deposited mainly in the northern Indian Ocean, building the Ganges and Indus Fans outward from the continental margin. As the mountains erode, the mass of crust bearing down on the underlying asthenosphere is reduced. As a result of the decreased downward force, further uplift occurs.

The third kind of plate boundary is known as a transform boundary and occurs where plates slide past one and other. A good example of this type of plate boundary is the San Andreas Fault system of California. Here the Pacific Plate is sliding northward past the North American Plate. The passage is not an easy one, however. The two plates occasionally stick together. When they do, stresses steadily build up. Eventually, the stress exceeds the frictional forces holding the plates together, and there is a sudden jump producing an earthquake. Earthquakes are also common in subduction zones and along mid-ocean ridges. They are much rarer in the interior of plates.

Most volcanism and crustal deformation occur along plate boundaries. A small fraction of volcanos, however, is located in plate interiors and appears to be entirely unrelated to plate tectonic processes. Crustal uplift also occurs in association with these volcanos. Two good examples are Hawaii and Yellowstone. These phenomena are thought to be the result of mantle plumes. Mantle plumes are convective upwellings. In contrast to the convective upwelling occurring along mid-ocean ridges, which is typically sheet-like, mantle plumes appear to be narrow (~100 km diameter) and approximately cylindrical. Furthermore, it appears that mantle plumes rise from much deeper in the mantle, perhaps even the core-mantle boundary, than convection associated with plate motion.

1.5.3 Earth Materials

The most abundant elements in the Earth are O and Fe (both close to 32%), Mg (~15%), Si (~14%), Ni (~1.8%), Ca (1.7%), and Al (1.6%). The majority of the Earth's Fe and Ni are in the core. The remaining rocky part of the Earth, the mantle and crust, consists of ~44% O, ~23% Mg, 21% Si, 8% Fe, 2.5% Ca, and 2.4% Al. As a consequence, the outer part of the Earth consists principally of compounds known as *silicates*. Silicates are compounds based on the silica tetrahedron, consisting of a silicon atom surrounded by four oxygens (Figure 11.13a). The bonds between the oxygens and silicon are about 50% covalent–50% ionic and are quite strong. The silicon atom shares an electron with each of the four oxygens. Since oxygen has two valence electrons, each oxygen can form an additional bond. There are two possibilities: the oxygen can form a second bond (which is usually more ionic) with another metal atom or it can form a bond with a second silicon. This latter possibility leads to linking of silica tetrahedra to form rings, chains, sheets, or frameworks. Oxygens bound to two silicons are called *bridging oxygens*.

In orthosilicates, the silica tetrahedra are either completely independent or form dimers, that is, two linked tetrahedra. A good example of a mineral of this type is olivine, whose structure is illustrated in Figure 11.11b. The chemical formula for olivine is $(Mg,Fe)_2SiO_4$. The notation (Mg,Fe) indicates that either magnesium or iron may be present. *Olivine* is an example of a solid solution between the Mg-end member, forsterite (Mg_2SiO_4) , and the Fe-end member, fayalite (Fe_2SiO_4) . Such solid solutions are quite common among silicates. As the formula indicates, there are two magnesium or iron atoms for each silica tetrahedra. Since each Mg or Fe has a charge of +2, their charge balances the –4 charge of each silica tetrahedra. Olivine constitutes roughly 50% of the Earth's upper mantle, and is thus one of the most abundant minerals in the Earth. At great pressures, it disproportionates to a Fe-Mg oxide (magnesiowüstite) and $(Mg,Fe)SiO_3$ (Mg-perovskite). At the same time, silicon atoms become octahedrally coordinated (surrounded by 6 oxygens rather than 4 as in the silica tetrahedra).

In chain silicates, the silica tetrahedra are linked together to form infinite chains (Figure 1.11c), with two bridging oxygen per tetrahedra. Minerals of this group are known as *pyroxenes* and have the gen-

Geochemistry

CHAPTER 1: INTRODUCTION

eral formula $XSiO_3$ where X is some metal, usually Ca, Mg, or Fe, which is located between the chains. Two pyroxenes, orthopyroxene ((Mg,Fe)SiO₃) and clinopyroxene (Ca(Mg,Fe)Si₂O₆ are very abundant in the Earth's upper mantle as well as mafic igneous rocks. The pyroxenes wollastonite (CaSiO₃) and jadite (NaAlSi₂O₆) are found exclusively in metamorphic rocks.

In double chain silicates, an additional one half bridging oxygen per tetrahedra joins two chains together. Minerals of this group are known as *amphiboles*, which occur widely in both igneous and metamorphic rocks. Among the important minerals in this group are hornblende (Ca₂Na(Mg,Fe)₄ Al₃Si₈O₂₂(OH)₂), tremolite-actinolite (Ca₂(Mg,Fe)₅Si₈ O₂₂(OH)₂), and glaucophane (Ca₂(MgFe)₃Al₅Si₈O₂₂(OH)₂). These minerals all contain OH as an essential component (Cl or F sometimes substitutes for OH). They are thus examples of hydrous silicates.

Sharing of a third oxygen

links the tetrahedra into sheets, forming the sheet silicates (Figure 1.11d). This group includes *micas* such as biotite

 $(K(Mg,Fe)_3AlSi_3O_{10}(OH)_2)$ and muscovite $(KAl_3Si_3O_{10} (OH)_2)$, talc $(Mg_3Si_4O_{10} (OH)_2)$, and *clay minerals* such as kaolinite $(Al_2Si_2O_5 (OH)_4)$. As in amphiboles, OH is an essential component of sheet silicates. These minerals can form through weathering and are thus primary sedimentary minerals. Many of them are found in igneous and metamorphic rocks as well.

When all four oxygens are shared between tetrahedra, the result is a framework. The simplest framework silicate is quartz (Si O_2), which consists solely of linked SiO₄ tetrahe-The other important dra. group of framework silicates are the *feldspars*, of which there are three end-members: sanidine $(KAlSi_3O_8)$, albite (NaAl Si₃O₈), and anorthite (Ca $Al_2Si_2O_8$). The calcium and sodium feldspars form the plagioclase solid solution, which is stable through a large temperature range. Sodium and potassium feldspars, collectively called alkali feldspar, form more limited solid solutions. Feldspars are the most abundant minerals in the Earth's crust.



Pyroxene Structure

Figure 1.11. The silica tetrahedron and the structure of silicate minerals. a. The silica tetrahedron consists of a central silicon atom bound to 4 oxygens. b. In orthosilicates such as olivine, the tetrahedra are separate and each oxygen is also bound to other metal ions that occupy interstitial sites between the tetrahedra. c. In pyroxenes, the tetrahedra each share two oxygen and are bound together into chains. Metal ions are located between the chains. d. In sheet silicates, such as talc, mica, and clays, the tetrahedra each share 3 oxygens and are bound together into sheets.

Geochemistry

CHAPTER 1: INTRODUCTION

Silicates are the most abundant minerals in the Earth, but not the only ones. Other classes of minerals include oxides such as magnetite (Fe₃O₄) and ilmenite (FeTiO₃), carbonates such as calcite (CaCO₃), sulfates such as gypsum CaSO₄·2H₂O, hydroxides such as gibbsite (Al(OH)₃), and sulfides such as pyrite FeS₂.

References and Suggestions for Further Reading

Gribbin, J. R. 1984. In Search of Schrödinger's Cat: Quantum Physics and Reality. New York: Bantam Books.

Other Texts in Geochemistry:

Broecker, W. S. and V. M. Oversby, 1971. *Chemical Equilibria in the Earth.* New York: McGraw-Hill. Henderson, P. 1986. *Inorganic geochemistry*. Oxford: Pergamon Press.

Richardson, S. M. and H. Y. McSween. 1988. *Geochemistry: Pathways and Processes*. New York: Prentice Hall.

Brownlow, A. H. 1996. Geochemistry. New York: Prentice Hall.

Krauskopf, K. B. and D. K. Bird. 1995. Introduction to Geochemistry. New York: McGraw-Hill.

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

CHAPTER 2: ENERGY, ENTROPY AND FUNDAMENTAL THERMODYNAMIC CONCEPTS

2.1 The Thermodynamic Perspective

when that metamorphic rocks weathers. Thus thermodynamics allows us to understand (in the sense that we defined understanding in Chapter 1) a great variety of geologic processes.

Thermodynamics embodies a *macroscopic* viewpoint, i.e., it concerns itself with the properties of a system, such as temperature, volume, heat capacity, and it does not concern itself with how these properties are reflected in the internal arrangement of atoms. The *microscopic* viewpoint, which is concerned with transformations on the atomic and subatomic levels, is the realm of *statistical mechanics* and *quantum mechanics*. In our treatment, we will focus mainly on the macroscopic (thermodynamic) viewpoint, but we will occasionally consider the microscopic (statistical mechanical) viewpoint when our understanding can be enhanced by doing so.

In principle, *thermodynamics is only usefully applied to systems at equilibrium*. If an equilibrium system is perturbed, thermodynamics can predict the new equilibrium state, but cannot predict how, how fast, or indeed whether, the equilibrium state will be achieved. (The field of *irreversible thermodynamics*, which we will not treat in this book, attempts to apply thermodynamics to non-equilibrium states. However, we will see in Chapter 5 that thermodynamics, through the *principle of detailed balancing* and *transition state theory*, can help us predict reaction rates.)

Kinetics is the study of rates and mechanisms of reaction. Whereas thermodynamics is concerned with the ultimate equilibrium state and not concerned with the pathway to equilibrium, kinetics concerns itself with the pathway to equilibrium. Very often, equilibrium in the Earth is not achieved, or achieved only very slowly, which naturally limits the usefulness of thermodynamics. Kinetics helps us to understand how equilibrium is achieved and why it is occasionally not achieved. Thus these two fields are closely related, and together form the basis of much of geochemistry. We will treat kinetics in Chapter 5.

2.2 Thermodynamic Systems and Equilibrium

We now need to define a few terms. We begin with the term *system*, which we have already used. A thermodynamic system is simply that part of the universe we are considering. Everything else is referred to as the *surrounding*. A thermodynamic *system* is defined at the convenience of the observer in a manner so that thermodynamics may be applied. While we are free to choose the boundaries of a system, our choice must nevertheless be a careful one as the success or failure of thermodynamics in describing the system will depend on how we have defined its boundaries. Thermodynamics often allows us this sort of freedom of definition. This can certainly be frustrating, particularly for someone exposed to thermodynamics for the first time (and often even the second or third time). But this freedom allows us to apply thermodynamics successfully to a much broader range of problems than otherwise.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS



Figure 2.1. Systems in relationship to their surroundings. The ball represents mass exchange, the arrow represents energy exchange.

A system may be related to its environment in a number of ways. An isolated system can exchange neither energy (heat or work) nor matter with its surroundings. A truly isolated system does not exist in nature, so this is strictly a theoretical concept. An adiabatic system can exchange energy in the form of work, but not heat or matter, with its surroundings, that is to say it has thermally insulating boundaries. Though a truly adiabatic system is probably also a fiction, many geologic systems are sufficiently well insulated that they may be considered adiabatic. *Closed* systems may exchange energy, in the form of both heat and work with their surrounding but cannot exchange matter. An open system may exchange both matter and energy across it boundaries. The various possible relationships of a system to its environment are illustrated in Figure 2.1.

Depending on how they behave over time, systems are said to be either in *transient* or *timeinvariant* states. Transient states are those that change with time. Time-independent states may

be either *static* or *dynamic*. A dynamic time-independent state, or *steady-state*, is one whose thermodynamic and chemical characteristics do not change with time despite internal changes or exchanges of mass and energy with its surroundings. As we will see, the ocean is a good example of a steady-state system. Despite a constant influx of water and salts from rivers and loss of salts and water to sediments and the atmosphere, it composition does not change with time (at least on geologically short time scales). Thus a steady-state system may also be an open system. We could define a static system is one in which nothing is happening. For example, an igneous rock or a flask of seawater (or some other solution) is static in the macroscopic perspective. From the statistical mechanical viewpoint, however, there is a constant reshuffling of atoms and electrons, but with no net changes. Thus static states are generally also dynamic states when viewed on a sufficiently fine scale.

Let's now consider one of the most important concepts in physical chemistry, that of *equilibrium*. One of the characteristics of the *equilibrium state* is that it is static from a macroscopic perspective, that is, it does not change measurably with time. Thus the equilibrium state is always time-invariant. While a reaction $A \rightarrow B$ may appear to have reached static equilibrium on a macroscopic scale this reaction may still proceed on a microscopic scale but with the rate of reaction $A \rightarrow B$ being the same as that of $B \rightarrow A$. Indeed, a kinetic definition of equilibrium is that the forward and reverse rates of reaction are equal.

The equilibrium state is entirely independent of the manner or pathway in which equilibrium is achieved. Indeed, once equilibrium is achieved, no information about previous states of the system can be recovered from its thermodynamic properties. Thus a flask of CO_2 produced by combustion of graphite cannot be distinguished from CO_2 produced by combustion of diamond. In achieving a new equilibrium state, all records of past states are destroyed.

Time-invariance is a necessary but not sufficient condition for equilibrium. Many systems exist in metastable states. Diamond at the surface of the Earth is not in an equilibrium state, despite its time-invariance on geologic time scales. Carbon exists in this metastable state because of kinetic barriers that inhibit transformation to graphite, the equilibrium state of pure carbon at the Earth's surface. Over-coming these kinetic barriers generally requires energy. If diamond is heated sufficiently, it will transform to graphite, or in the presence of sufficient oxygen, to CO_2 .

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

The concept of equilibrium versus metastable or unstable (transient) states is illustrated in Figure 2.2 by a ball on a hill. The equilibrium state is when the ball is in the valley at the bottom of the hill, because its gravitational potential energy is minimized in this position. When the ball is on a slope, it is in an unstable, or transient, state and will tend to roll down the hill. However, it may also become trapped in small depressions on the side of the hill, which represent metastable states. The small hill bordering the depression represents a kinetic barrier. This kinetic barrier can only be overcome when the ball acquires enough energy to roll Figure 2.2. States of a system. up and over it. Lacking that energy, it will exist in the metastable state indefinitely.



In Figure 2.2, the ball is at equilibrium when its (gravitational) potential energy lowest (i.e., at the bottom of the hill). This is a good definition of equilibrium in this system, but as we will soon see, is not adequate in all cases. A more general statement would be to say that the equilibrium state is the one toward which a system will change in the absence of constraints. So in this case, if we plane down the bump (remove a constraint), the ball rolls to the bottom of the hill. At the end of this Chapter, we will be able to produce a thermodynamic definition of equilibrium based on the Gibbs Free Energy. We will find that, for a given pressure and temperature, the chemical equilibrium state occurs when the Gibbs Free Energy of the system is lowest.

Natural processes proceeding at a finite rate are *irreversible*: under a given set of conditions; i.e., they will only proceed in one direction. Here we encounter a problem in the application of thermodynamics: if a reaction is proceeding, then the system is out of equilibrium and thermodynamic analysis cannot be applied. This is one of the first of many paradoxes in thermodynamics. This limitation might at first seem fatal, but we get around it by imagining a comparable reversible reaction. *Reversibility and* local equilibrium are concepts that allow us to 'cheat' and apply thermodynamics to non-equilibrium situations. A "reversible" process is an idealized one where the reaction proceeds in sufficiently small steps that it is in equilibrium at any given time (thus allowing the application of thermodynamics).

Local equilibrium embodies the concept that in a closed or open system, which may not be at equilibrium on the whole, small volumes of the system may nonetheless be at equilibrium. There are many such examples. In the example of mineral crystallizing from magma, only the rim of the crystal may be in equilibrium with the melt. Information about the system may nevertheless be derived from the relationship of this rim to the surrounding magma. Local equilibrium is in a sense the spatial equivalent to the temporal concept of reversibility and allows the application of thermodynamics to real systems, which are invariably non-equilibrium at large scales. Both local equilibrium and reversibility are examples of simplifying assumptions that allow us to treat complex situations. In making such assumptions, some accuracy in the answer may be lost. Knowing when and how to simplify a problem is an important scientific skill.

2.2.1 Fundamental Thermodynamic Variables

In the next two chapters we will be using a number of variables, or properties, to describe thermodynamic systems. Some of these will be quite familiar to you, others less so. Volume, pressure, energy, heat, work, entropy, and temperature are most fundamental variables in thermodynamics. As all other thermodynamic variables are derived from them, it is worth our while to consider a few of these properties.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

Energy is the capacity to produce change. It is a fundamental property of any system, and it should be familiar from physics. By choosing a suitable reference frame, we can define an absolute energy scale. However, it is changes in energy that are generally of interest to us rather than absolute amounts. Work and heat are two of many forms of energy. Heat, or thermal energy, results from random motions of molecules or atoms in a substance and is closely related to kinetic energy. Work is done by moving a mass, M, through some distance, x = X, *against* a force *F*:

$$w = -\int_0^X F \, dx \tag{2.1}$$

where *w* is work and force is defined mass times acceleration:

$$F = -M\frac{dv}{dt}$$
 2.2

(the minus signs are there because of the convention that *work done on a system is positive, work done by a system is negative*). This is, of course, Newton's first law. In chemical thermodynamics, pressure–volume work is usually of more interest. Pressure is defined as force per unit area:

$$P = \frac{F}{A}$$
 2.3

Since volume is area times distance, we can substitute equation 2.3 and dV = Adx into 2.1 and obtain:

$$w = -\int_{x_0}^{x_1} \frac{F}{A} A dx = -\int_{V_0}^{V_1} P dV$$
 2.4

Thus work is also done as a result of a volume change in the presence of pressure.

Potential energy is energy possessed by a body by virtue of its position in a force field, such as the gravitational field of the Earth, or an electric field. Chemical energy will be of most interest to us in this book. Chemical energy is a form a potential energy stored in chemical bonds of a substance. Chemical energy arises from the electromagnetic forces acting on atoms and electrons. Internal energy, which we denote with the symbol U, is the sum of the potential energy arising from these forces as well as the kinetic energy of the atoms and molecules (i.e., thermal energy) in a substance. It is internal energy that will be of most interest to us.

We will discuss all these fundamental variables in more detail in the next few sections.

2.2.1.1 Properties of State

Properties or variables of a system that depend only on the present state of the system, and not on the manner in which that state was achieved are called *variables of state* or *state functions*. *Extensive* properties depend on total size of the system. Mass, volume, and energy are all extensive properties. Extensive properties are additive, the value for the whole being the sum of values for the parts. *Intensive* properties are independent of the size of a system, for example temperature, pressure, and viscosity. They are not additive, e.g., the temperature of a system is not the sum of the temperature of its parts. In general, an extensive property can be converted to an intensive one by dividing it by some other extensive property. For example, density is the mass per volume and is an intensive property. It is generally more convenient to work with intensive rather than extensive properties. For a single component system not undergoing reaction, specification of 3 variables (2 intensive, 1 extensive) is generally sufficient to determine the rest, and specification of any 2 intensive variables is generally sufficient to determine the remaining intensive variables.

A final definition is that of a *pure substance*. A pure substance is one that cannot be separated into fractions of different properties by the same processes as those considered. For example, in most processes, the compound H_2O can be considered a pure substance. However, if electrolysis were involved, this would not be the case.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

2.3 Equations of State

Equations of state describe the relationship that exists among the state variables of a system. We will begin by considering the ideal gas law and then very briefly consider two more complex equations of state for gases.

2.3.1 Ideal Gas Law

The simplest and most fundamental of the equations of state is the *ideal gas law*^{\dagger}. It states that pressure, volume, temperature, and the number of moles of a gas are related as:

$$\mathbf{PV} = \mathbf{NRT}$$

where P is pressure, V is volume, N is the number of moles, T is thermodynamic, or absolute temperature (which we will explain shortly), and R is the ideal gas constant* (an empirically determined constant equal to 8.314 J/mol-K, 1.987 cal/mol-K or 82.06 cc-atm/deg-mol). This equation describes the relation between two extensive (mass dependent) parameters, volume and the number of moles, and two intensive (mass independent) parameters, temperature and pressure. We earlier stated that if we defined two intensive and one extensive system parameter, we could determine the remaining parameters. We can see from equation 2.5 that this is indeed the case for an ideal gas. For example, if we know N, P, and T, we can use equation 2.5 to determine V.

The ideal gas law, and any equation of state, can be rewritten with intensive properties only. Dividing V by N we obtain the *molar volume*, V. Substituting V for V and rearranging, the ideal gas equation becomes:

$$\overline{V} = \frac{RT}{P}$$
 2.6

The ideal gas equation tells us how the volume of a given amount of gas will vary with pressure and temperature. To see how molar volume will vary with temperature alone, we can differentiate equation 2.6 with respect to temperature holding pressure constant and obtain:

$$\left(\frac{\partial V}{\partial T}\right)_{P} = \frac{\partial (NRT / P)}{\partial T}$$

$$2.7$$

$$\left(\frac{\partial V}{\partial T}\right)_{P} = \frac{NR}{P}$$

$$2.8$$

which reduces to:

It would be more useful to know to *fractional* volume change rather than the absolute volume change with temperature, because the result in that case does not depend on the size of the system. To convert to the fractional volume change, we simply divide the equation by V:

$$\frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_{P} = \frac{NR}{PV}$$
 2.9

Comparing equation 2.9 with 2.5, we see that the right hand side of the equation is simply 1/T, thus

$$\frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_{P} = \frac{1}{T}$$
 2.10

The left hand side of this equation, the fractional change in volume with change in temperature, is known as the *coefficient of thermal expansion*, α :

2.8

⁺ Frenchman Joseph Gay-Lussac (1778-1850) established this law based on earlier work of Englishman Robert Boyle and Frenchman Edme Mariotte.

^{*} We will often refer to it merely as the gas constant.



CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

$$\alpha \equiv \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_{P} \qquad 2.11$$

 Table 2.1. Van der Waals Constants for Selected Gases

For an ideal gas, the coefficient of thermal expansion is simply the inverse of temperature.

The *compressibility* of a substance is defined in a similar manner as the fractional change in volume produced by a change in pressure at constant temperature:

$$\left|\beta \equiv -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_T\right| = 2.12$$

Gas	а	b
	liter ² -atm/mole ²	liter/mole
Helium	0.034	0.0237
Argon	1.345	0.0171
Hydrogen	0.244	0.0266
Oxygen	1.360	0.0318
Nitrogen	1.390	0.0391
Carbon Dioxide	e 3.592	0.0399
Water	5.464	0.0305
Benzene	18.00	0.1154

Geologists sometimes use the *isothermal bulk modulus*, K_T , in place of compressibility. The isothermal bulk modulus is simply the inverse of compressibility: $K_T = 1/\beta$. Through a similar derivation to the one we have just done for the coefficient of thermal expansion, it can be shown that the compressibility of an ideal gas is $\beta = 1/P$.

The ideal gas law can be derived from statistical physics (first principles), assuming the molecules occupy no volume and have no electrostatic interaction. Doing so, we find, $R = N_0 k$, where k is Boltzmann's constant (1.381 x 10⁻¹⁶ erg/K) and N_0 is Avagadro's Number (the number of atoms in one mole of a substance). k is a fundamentally constant that relates the average molecular energy, *e*, of an ideal gas to its temperature (K) as e = 3kT/2.

Since the assumptions just stated are ultimately invalid, it is not surprising that the ideal gas law is only an approximation for real gases, it applies best in the limit of high temperature and low pressure. Deviations are largest near the condensation point of the gas.

The compressibility factor is a measure of deviation from ideality and is defined as

$$Z = PV/NRT 2.13$$

By definition, Z=1 for an ideal gas.

2.3.2 Equations of State for Real Gases

2.3.2.1 VAN dER WAALS EQUATION:

Factors we need to consider in constructing an equation of state for a real gas are the finite volume of molecules and the attractive and repulsive forces between molecules arising from electric charges. The Van der Waals equation is probably the simplest equation of state that takes account of these factors. The Van der Waals equation is:

$$P = \frac{RT}{\overline{V} - b} - \frac{a}{\overline{V}^2}$$
 2.14

Here again we have converted volume from an extensive to an intensive property by dividing by N.

Let's examine the way in which the Van der Waals equation attempts to take account of finite molecular volume and forces between molecules. Considering first the forces between molecules, imagine two volume elements v_1 and v_2 . The attractive forces will be proportional to the number of molecules or the concentrations, c_1 and c_2 , in each. Therefore, attractive forces are proportional to $c_1 \propto c_2 = c^2$. Since c is the number of molecules per unit volume, c=n/V, we see that attractive forces are proportional to $1/\overline{V}^2$. Thus it is the second term on the right that takes account of forces between molecules. The *a* term is a constant that depends on the nature and strength of the forces between molecules, and will therefore be different for each type of gas.

In the first term on the right, \overline{V} has been replaced by $\overline{V} - b$. *b* is the volume actually occupied by molecules, and the term $\overline{V} - b$ is the volume available for movement of the molecules. Since different

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

gases have molecules of differing size, we can expect that the value of *b* will also depend on the nature of the gas. Table 2.1 lists the values of *a* and *b* for a few common gases.

2.3.2.2 Other Equations of State for Gases

The *Redlich-Kwong Equation* (1949) expresses the attractive forces as a more complex function:

$$P = \frac{RT}{\overline{V} - b} - \frac{a}{T^{1/2}\overline{V}(\overline{V} + b)}$$
2.15

The *Virial Equation* is much easier to handle algebraically than the van der Waals equation and has some theoretical basis in statistical mechanics:

$$PV = A + BP + CP^2 + DP^3 + \dots$$
 2.16

A, B, C, are empirically determined (temperature dependent) constants.

2.3.3 Equation of State for Other Substances

The compressibility and coefficient of thermal expansion parameters allow us to construct an equation of state for any substance. Such an equation relates the fundamental properties of the substance: its temperature, pressure, and volume. The partial differential of volume with respect to temperature and pressure is such an equation:

$$dV = \left(\frac{\partial V}{\partial T}\right)_P dT + \left(\frac{\partial V}{\partial P}\right)_T dP$$
 2.17

Substituting the coefficient of thermal expansion and compressibility for $\partial V/\partial T$ and $\partial V/\partial P$ respectively we have:

$$dV = V\left(\alpha dT - \beta dP\right)$$
 2.18

Thus to write an equation of state for a substance, our task becomes to determine its compressibility and coefficient of thermal expansion. Once we know them, we can integrate equation 2.18 to obtain the equation of state. These, however, will generally be complex functions of temperature and pressure, so the task is often not easy.

2.2 TEMPERATURE, Absolute ZERO, AND THE ZEROTH LAW OF THERMODYNAMICS

How do you define temperature? We discussed temperature with respect to the ideal gas law without defining it, though we all have an intuitive sense of what temperature is. We noted above that temperature of a gas is measure of the average (kinetic) energy of its molecules. Another approach might be to use the ideal gas law to construct a *thermometer* and define a temperature scale. A convenient thermometer might be one based on the linear relationship between temperature and the volume of an ideal gas. Such a thermometer is illustrated in Figure 2.3. The equation describing the relationship between the volume of the gas in the thermometer and our temperature, τ , is:

$$V = V_0 (1 + \gamma \tau)$$
 2.19

where V_0 is the volume at some reference point where $\tau = 0$ (Figure 2.3a) and γ is a scale factor. For example, we might choose $\tau=0$ to be the freezing point of water and the scale factor such that $\tau=100$ (Figure 2.3b) occurs at the boiling point of water, as is the case in the centigrade scale. Rearranging, we have:

$$\tau = \frac{1}{\gamma} \left(\frac{V}{V_0} - 1 \right)$$
 2.20

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS



Figure 2.3. An ideal gas thermometer.

Then $\tau = 0$ at $V = V_0$. If V is less than the reference volume, then temperature will be negative on our scale. But notice that while any positive value of temperature is possible on this scale, there is a limit to the range of possible negative values. This is because V can never be negative. The minimum value of temperature on this scale will occur when V is 0. This occurs at:

$$\tau_0 = -\frac{1}{\gamma} \qquad 2.21$$

Thus implicit in the ideal gas law, which we used to make this thermometer, is the idea that there is an absolute minimum value, or an absolute zero, of temperature, which occurs when the volume of an ideal gas is 0. Notice that while the value($-1/\gamma$) of this absolute zero will depend on how we designed our thermometer, i.e., on V₀, the result, that a minimum value exists, does not. We should also point out that only an ideal gas will have a volume of 0 at the absolute 0. The molecules of real gases have a finite volume, and such a gas will have a finite volume at absolute 0.

The temperature scale used by convention in thermodynamics is the Kelvin^{*} scale. The magnitude of units, called kelvins (not degrees kelvin) and designated K (not ° K), on this scale are the same as the centigrade scale, i.e., there are exactly 100 kelvins between the freezing and boiling point of water. There is some slight uncertainty (a very much smaller uncertainty than we need to concern ourselves with) concerning the value of absolute zero (i.e., the value of γ in 2.20 and 2.21). The scale has been fixed by choosing 273.16 kelvins to be the triple point of water (0.01° C). On this scale, the absolute zero of temperature occurs at 0±0.01 kelvins. The Kelvin scale should be used wherever temperature occurs in a thermodynamic equation.

Temperature has another fundamental property, and this is embodied in the *zeroth law of thermodynamics*. It is sufficiently ob-

vious from everyday experience that we might overlook it. It concerns thermal equilibrium and may be stated in several ways: *two bodies in thermal equilibrium have the same temperature* and *any two bodies in thermal equilibrium with a third are in equilibrium with each other.*

2.3 Energy and The First Law of Thermodynamics

2.3.1 Energy

The first law may be stated various ways:

- •heat and work are equivalent
- energy is conserved in any transformation;
- the change of energy of a system is independent of the path taken.[‡]

^{*} named for Lord Kelvin. Born William Thomson in Scotland in 1824, he was appointed Professor at Glasgow University at the age of 22. Among his many contributions to physics and thermodynamics was the concept of absolute temperature. He died in 1907.

[‡]This may seem intuitively obvious to us, but it wasn't to James Joule (1818-1889), English brewer and physicist, who postulated it on the basis of experimental results. It wasn't obvious to his contemporaries either. His presentation of the idea of equivalence of heat and work to the British Association in 1843 was received with "entire incredulity" and "general silence". The Royal Society rejected his paper on the subject a year later. If you think

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

All are restatements of the law of conservation of energy:

Energy can be neither created nor destroyed.

Mathematically:

$\Delta U=Q+W$ or dU=dQ+dW

2.22

Thermodynamics is concerned only with the internal energy of a system. We don't really care whether the system as a whole is in motion, i.e., whether it has kinetic energy (we do care, however, about the internal kinetic energy, or heat). For the most part, we also don't care whether it has potential energy, i.e., what its position in potential fields is, except to the extent that this influences the state of our system (e.g., pressure in the atmosphere is a function of the altitude, and hence would be of interest to us). In addition, we are almost always concerned *only* with energy *changes*, not with the absolute energy of a system. In thermodynamics ΔU , not U, is the interesting quantity.

Energy may be transferred between a system and it surroundings in several ways: heat, work, radiation, and advection (i.e., energy associated with mass gained or lost by the system). Whenever possible, we will want to chose our system such that it is closed and we don't have to worry about the latter. In most, but not all, instances of geochemical interest, radiation is not important. Thus in geochemical thermodynamics, *heat and work are the forms of energy flow of primary interest*.

2.3.2 Work

We have seen that work is the integral of force applied through a distance. Force times distance has units of energy (mass-velocity²), thus work is form of energy. The SI (Systeme International) unit of energy is the Joule = $1 \text{ kg-m}^2/\text{s}^2$.

There are several kinds of work of interest to thermodynamics, the most important of which is that involved in chemical reactions (later, when we consider oxidation and reduction reactions, we will be concerned with electrochemical work). One of the most important forms of work in classical thermo-dynamics is 'PV' work: expansion and contraction. Expressing equation 2.4 in differential form:

$$dW = -P_{ext} dV 2.23$$

Pressure is force per unit area and therefore has units of mass-time⁻²-distance⁻¹, volume has units of distance³. The product of P and V therefore has units of energy: mass-(distance/time)².[†] The negative sign arises because, *by convention, we define energy flowing into the system as positive*. Work done by the system is thus negative, work done on the system is positive. This conforms to a 1970 I.U.P.A.C. (International Union of Pure and Applied Chemistry) recommendation.

While 'PV' work is not as important in geochemistry as in other applications of thermodynamics, it is nevertheless of significant interest. There is, of course, a great range of pressures within the Earth. Systems rising within the Earth, such a magma or a hydrothermal fluid, will thus do work on their surroundings, and systems sinking, such as sediments being buried or lithosphere being subducted, will have work done on them.

We mentioned the concept of reversible and irreversible reactions, and stated that a reversible reaction is one that occurs in sufficiently small steps that equilibrium is maintained. In an expansion or contraction reaction, equilibrium is maintained and the reaction is reversible if the external pressure is equal to the internal pressure. The work done under these conditions is said to be reversible:

about it a bit, it is not so obvious — in fact, there is no good reason why heat and work should be equivalent. This law is simply an empirical observation. The proof is a negative one: experience has found no contradiction of it. German physician Julius Mayer (1814-1878) formulated the idea of conservation of energy in 1842, but his writing attracted little attention. It was Joule's experiments with heat and work that conclusively established the principle of conservation of energy. By 1850, the idea of conservation of energy began to take hold among physicists, thanks to Joule's persistence and the support of a brilliant young physicist named William Thomson, who also had been initially skeptical.

⁺ The pascal, the SI unit of pressure, is equal to 1 kg/m-s². Thus if pressure is measured in MPa (megapascals, 1 atm \approx 1 bar = 0.1 MPa) and volume in cc (= 10⁻⁶ m⁻³), the product of pressure times volume will be in joules. This is rather convenient.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

$$dW_{rev} = -PdV 2.24$$

2.3.4 Path Independence, Exact Differentials, State Functions, and the First Law

We said earlier that state functions are those that depend only on the present state of a system. Another way of expressing this is to say that state functions are path independent. Indeed, path independence may be used as a test of whether a variable is a state function or not. This is to say that if Y is a state function, then for any process that results in a change Y1 \rightarrow Y2, the net change in Y, Δ Y, is independent of how one gets from Y1 to Y2. Furthermore, if Y is a state function, then the differential dY is said to be mathematically *exact*.

Let's explore what is meant by an exact differential. An exact differential is the familiar kind, the kind we would obtain by differentiating the function u with respect to x and y, and also the kind we can integrate. But not all differential equations are exact. Let's first consider the mathematical definition of an exact differential, then consider some thermodynamic examples of exact and inexact differentials.

Consider the first order differential expression:

$$Mdx + Ndy$$
 2.25

containing variables M and N, which may or may not be functions of x and y. Equation 2.25 is said to be an *exact differential* if there exists some function u of x and y relating them such that the expression:

$$du = Mdx + Ndy 2.26$$

is the total differential of *u*, i.e.,:

$$du = \left(\frac{\partial u}{\partial x}\right)_{y} dx + \left(\frac{\partial u}{\partial y}\right)_{x} dy \qquad 2.27$$

Let's consider what this implies. Comparing 2.26 and 2.27, we see that:

$$\frac{\partial u}{\partial x} = M$$
 and $\frac{\partial u}{\partial y} = N$ 2.28

A necessary, but not sufficient, condition for 2.25 to be an exact differential is that M and N must be functions of x and y.

A general property of partial differentials is the *reciprocity relation* or *cross-differentiation identity*, which that the order of differentiation does not matter, so that:

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x}$$
 2.29

(The reciprocity relation is an important and useful property in thermodynamics, as we shall see at the end of this chapter.) If equation 2.26 is the total differential of u, it follows that:

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$
 2.30

$$\left(\frac{dM}{dy}\right)_{x} = \left(\frac{dN}{dx}\right)_{y}$$
 2.31

which is equivalent to:

Equation 2.31 is a necessary and sufficient condition for 2.25 to be an exact differential, that is, if the cross differentials are equal, then the differential expression is exact.

Exact differentials have the property that they can be integrated and an exact value obtained. This is true because they depend only on the initial and final values of the independent variables (e.g., x and y in 2.27).

Now let's consider some thermodynamic examples. Volume is a state function and we can express it as an exact differential in terms of other state functions:

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

$$dV = \left(\frac{\partial V}{\partial T}\right)_{P} dT + \left(\frac{\partial V}{\partial P}\right)_{T} dP$$
(2.17)

Substituting the coefficient of thermal expansion and compressibility for $\partial V/\partial T$ and $\partial V/\partial P$ respectively, equation 2.30 becomes:

$$dV = \alpha V dT + \beta V dP \tag{2.18}$$

According to equation 2.31, if V is a state function, then:

$$\frac{\partial(\alpha V)}{\partial P} = -\frac{\partial(\beta V)}{\partial T}$$
 2.32

You should satisfy yourself that equation 2.32 indeed holds for ideal gases and therefore that V is a state variable.

Work is not a state function, that is, the work done does not depend only on the initial and final state of a system. We would expect then that dW is not an exact differential, and indeed, this is easily shown for an ideal gas.

For PV work, dW = -PdV. Substituting equation 2.17 for dV and rearranging, we have:

$$dW = -P\left[\left(\frac{\partial V}{\partial T}\right)_{P} dT + \left(\frac{\partial V}{\partial P}\right)_{T} dP\right]$$
 2.33

Evaluating $\partial V/\partial T$ and $\partial V/\partial P$ for the ideal gas equation and multiplying through by P, this becomes:

$$dW = -NRdT + \frac{NRT}{P}dP \qquad 2.34$$

but

$$\frac{\partial NR}{\partial P} \neq \frac{\partial (NRT/P)}{\partial T}$$
 2.35

We cannot integrate equation 2.34 and obtain a value for the work done without additional knowledge of the variation of T and P because the amount of work done does not depend only on the initial and final values of T and P; it depends on the path taken. Heat is also not a state function, not an exact differential, and also path dependent. Path dependent functions always have inexact differentials; path independent functions always have exact differentials.

On a less mathematical level, let's consider how the work and heat will vary in a transformation of a system, say from state 1 to state 2. Imagine that we burn gasoline in an open container. In this case, in the transformation from state 1 (gasoline) to state 2 (combustion products of gasoline) energy is given up by the system only as heat. Alternatively, we could burn the gasoline in an engine and recover some of the energy as work (expansion of the volume of the cylinder resulting in motion of the piston). The end states of these two transformations are the same, but the amount of heat released and work done varied depending on the path we took. Thus neither work nor heat can be state functions. Energy is a state function, is path independent, and is an exact differential. Whether we burn the gasoline in an open container or an engine, the energy released will be the same. Herein lies the significance for thermodynamics of Joule's discovery: that *the sum of heat and work is independent of the path taken even though, independently, work and heat are not*.

2.4 The Second Law and Entropy

2.4.1 STATEMENT

Imagine a well-insulated box (an isolated system) somewhere in the universe (Figure 2.4). Imagine that within the box are two gases, separated by a removable partition. If we remove the partition, what happens? You know: the two gases mix completely. The process is entirely spontaneous. We have neither added energy to nor taken energy from the system, hence the first law says nothing about this process. Nor did removing the partition "cause" the reaction. This is apparent from the observation that if we reinsert the partition, the gases do not unmix. That you knew that the gases would mix (and

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

knew as well that they would not unmix upon reinserting the partition) suggests there is something very fundamental and universal about this. We need a physical law that describes it. This is the second law.

The Second Law may be stated in a number of ways:

- It is impossible to construct a machine that is able to convey heat by a cyclical process from one reservoir at a lower temperature to another at a higher temperature unless work is done by some outside agency (i.e., air conditioning is never free). §
- *Heat cannot be entirely extracted from a body and turned into work (i.e., an engine can never run 100% efficiently).*
- Every system left to itself will, on the average, change toward a condition of maximum probability.

Introducing a new state function **S** called *entropy*, we may state the second law as:

The entropy of the universe always increases.

In colloquial terms we could say:

You can't shovel manure into the rear end of a horse and expect to get hay out its mouth.



Figure 2.4. A gas-filled box with a removable partition. When the partition is removed, the gases mix. Entropy increases during this process.

The second law codifies some of our everyday experience. The first law would not prevent us from using a horse to manufacture hay from manure. It only says we can't get more joules worth of hay out than we put in as manure. We would search in vain for any other physical law that prohibited this event. Yet our experience shows that it won't happen. Indeed, this event is so improbable that we find it comical. Similarly, we know that we can convert gasoline and oxygen to carbon dioxide and water in an internal combustion engine and use the resulting energy to drive a vehicle down the road. But adding CO_2 and water to the engine and pushing the car backwards down the street does not produce gasoline and oxygen, although such a result violates no other law of physics. *The second law states that there is a natural direction in which reactions will tend to proceed*. This direction is inevitably that of higher entropy of the system and its surroundings.

2.4.2 Statistical Mechanics: A Microscopic Perspective of Entropy

Whereas energy is a property for which we gain an intuitive feel through every day experience, the concept of entropy is usually more difficult to grasp. Perhaps the best intuitive understanding of entropy can be obtained from the microscopic viewpoint of statistical mechanics. So for that reason, we will make the first of several brief excursions into the world of atoms, molecules, and quanta.

Let's return to our box of gas and consider what happens on a microscopic scale when we remove the partition. To make things tractable, we'll consider that each gas consists of only two molecules, so there are four all together, two red and two black. For this thought experiment, we'll keep track of the individual molecules, so we label them 1red, 2red, 1black, 2 black. Before we removed the partition, the red molecules were on one side and the black ones on the other. Our molecules have some thermal energy, so they are free to move around. So by removing the partition, we are essentially saying that each molecule is equally likely to be found in either side of the box.

[§]Rudolf Clausius (1822-1888), a physicist at the Prussian military engineering academy in Berlin, formulated what we now refer to as the second law and the concept of entropy in a paper published in 1850. Similar ideas were published a year later by William Thomson (Lord Kelvin), who is responsible for the word "entropy". Clausius was a theorist who deserves much of the credit for founding what we now call "thermodynamics" (he was responsible for, among many other things, the virial equation for gases). However, a case can be made that Sadi Carnot (1796-1832) should be given the credit. Carnot was a Parisian military officer (the son of a general in the French revolutionary army) interested in the efficiency of steam engines. The question of credit hinges on whether he was referring to what we now call entropy when he used the word "*calorique*".
Geochemistry

а

1

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

Before we removed the partition, there was only one possible arrangement of the system: this is shown in Figure 2.5a. Once we remove the partition, we have 4 molecules and two subvolumes, and a total of $2^4 = 16$ possible configurations (Fig. 2.5b) of the system. The *basic postulate of statistical mechanics* is: *a system is equally likely to be found in any of the states accessible to it*. Thus we postulate that each of these configurations are equally likely. Only one of these states corresponds to the original one (all red molecules on the left). Thus the probability of the system being found in its original state is 1/16. That is not particularly improbable. However, suppose that we had altogether a mole of gas ($\approx 6 \times 10^{23}$ molecules). The probability of the system ever being found again in its original state is then $\approx 2^{-10^{24}}$, which is unlikely indeed.

Now consider a second example. Suppose that we have two copper blocks of identical mass at different temperatures and separated by a thermally insulating barrier (Figure 2.6). Imagine that our system, which is the two copper blocks, is isolated in space so that the total energy of the system remains constant. What happens if we remove the insulating barrier? Experience tells us that the two copper blocks will eventually come into thermal equilibrium, i.e., their temperatures will eventually be identical.

Now let's look at this process on a microscopic scale. We have already mentioned that temperature is related to internal energy. As we shall see, this relationship will differ depending on the nature and mass of the material of interest, but since our blocks are of identical size and composition, we can assume temperature and energy are directly related in this case. Suppose that before we remove the insulation, the left block has 1 unit of energy and the right one has 5 (we can think of these as quanta, but this is not necessary). The question is, how will energy be distributed after we remove the insulation?

2 11 (1)221 0 ا 🌔 🚺 2 1 (2) I **1**2 21¹2 1 i (1) 2 🎱 (2) **1**2¦ 12 ı 🛈 💋 1 🕗 🛽 b Figure 2.5. Possible distribution of molecules of

a red and and a black

gas in a box before (a)

and after (b) removal of

a partition separating

1

2

1

In the statistical mechanical viewpoint, we cannot determine how the energy will be distributed; we can only compute the possible ways it could be distributed. Each



Figure 2.6. Two copper blocks at different temperatures separated by an insulator. When the insulator is removed and the blocks brought in contact, the blocks come to thermal equilibrium. Entropy increases in this process. of these energy distributions is then equally likely according to the basic postulate. So let's examine how it can be distributed. Since we assume that the distribution is completely random, we proceed by randomly

Since we assume that the distribution is com- them. pletely random, we proceed by randomly assigning the first unit to either to left or right block, then the second unit to either, etc. With 6 units of energy, there are already more ways of distributing it $(2^6 = 64)$ than we have space to enumerate here. For example, there are 6 ways energy can be distributed so that the left block has 1 unit and the right one 5 units. This is illustrated in Figure 2.7. However, since we can't actually distinguish the energy units, all these ways are effectively identical. There are 15 ways, or *combinations*, we can distribute energy so that the left block has 2 units and the right 4 units. Similarly there are 15 combinations where the left block has 4 units and the right has 2 units. For this particular example, the rule is

units and the right has 2 units. For this particular example, the rule is that if there are a total of *E* units of energy, *e* of which are assigned to the left block and (*E-e*) to the right, then there will be $\Omega(e)$ identical combinations where $\Omega(e)$ is calculated as:

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

$$\Omega(e) = \frac{E!}{e!(E-e)!} \qquad 2.36^{\dagger}$$

Here we use $\Omega(e)$ to denote the

	6 5 4 3		6 5 4 3		6 5 4 2		6 5 3 2		6 4 3 2		5 (4) (3) (2)
1	2	2	1	3	1	4	1	5	1	6	1

function that describes the number of states accessible to the system for a given value of *e*. In this particular example, "states accessible to the system" refers to a given dis-

Figure 2.7. There are 6 possible ways to distribute 6 energy units so that the left block has 1 unit and the right block has 5.

tribution of energy units between the two blocks. According to equation 2.36 there are 20 ways of distributing our 6 units of energy so that each block has three. There is, of course, only one way to distribute energy so that the left block has all of the energy and only one combination where the right block has all of it.

According to the basic postulate, any of the 64 possible distributions of energy are equally likely. The key observation, however, is that there are many ways to distribute energy for some values of *e* and only a few for other values. Thus the chances of the system being found in a state where each block has three units is 20/64 = 0.3125, whereas the chances of the system being in the state with the original distribution (1 unit to the left, 5 to the right) are only 6/64 = 0.0938. So it is much more likely that we will find the system in a state where energy is equally divided than in the original state.

Of course, two macroscopic blocks of copper at any reasonable temperature will have far more than 6 quanta of energy. Let's take a just slightly more realistic example and suppose that they have a total of 20 quanta, and compute the distribution. There will be 2^{20} possible distributions, far too many to consider individually, so let's do it the easy way and use equation 2.36 to produce a graph of the probability distribution. Equation 2.36 gives the number of identical states of the system for a given value of *e*. The other thing that we need to know is that the chances of any one of these states occurring, which is simply $(1/2)^{20}$. So to compute the probability of a particular distinguishable distribution of energy occurring, we multiply this probability by Ω . More generally the probability, \mathcal{P} , will be:

$$\mathcal{P}(e) = \frac{E!}{e!(E-e)!} p^e q^{E-e}$$
2.37

where *p* is the probability of an energy unit being in the left block and *q* is the probability of it being in the right. This equation is known as the *binomial distribution*^{*}. Since both *p* and *q* are equal to 0.5 in our case (if the blocks were of different mass or of different composition, *p* and *q* would not be equal), the product $p^e q^{E-e}$ is just p^E and 2.37 simplifies to:

$$\mathcal{P}(e) = \frac{E!}{e!(E-e)!} p^{E} = \Omega(e)p^{E}$$
2.38

Since p^E is a constant (for a given value of E and configuration of the system), the probability of the left block having *e* units of energy is directly proportional to $\Omega(e)$. It turns out that this is general relationship, so that for any system we may write:

$$\mathcal{P}(f) = \mathbf{C}\Omega(f) \tag{2.39}$$

$$\Omega = \frac{N!}{n_1! n_2! \dots n_m!} \quad \Omega = \frac{N!}{n_1! n_2! \dots n_m!}$$
 2.36a

⁺ This is the equation when there are two possible outcomes. A more general form for a situation where there are m possible outcomes (e.g., copper blocks) would be:

where there are n_1 outcomes of the first kind (i.e., objects assigned to the first block), n_2 outcomes of the second, etc. and $N = \sum n_i$ (i.e., N objects to be distributed).

^{*} If you have a spreadsheet program available to you, this equation may be a built-in function, which makes computing graphs such as Figure 2.8 much easier. In Microsoft ExcelTM, this is the BINOMDIST function.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

where *f* is some property describing the system and C is some constant (in this case 0.5^{20}). Figure 2.8a shows the probability of the left block having e units of energy. Clearly, the most likely situation is that both will have approximately equal energy. The chances of one block having 1 unit and the other 19 units is very small (2×10^{-5} to be exact). In reality of course, the number of quanta of energy available to the two copper blocks will be on the order of Avagadro's number. If one or the other block has 10 or 20 more units or even 10¹⁰ more quanta than the other, we wouldn't be able to detect it. Thus energy will always appear to be distributed evenly between the two, once the system has had time to adjust.

M. White

Figure 2.8b shows Ω as a function of *e*, the number of energy units in the left block. Comparing the two, as well as equation 2.38, we see that the most probable distribution of energy between the blocks corresponds to the situation where the system has the maximal number of states accessible to it, i.e., to where $\Omega(e)$ is maximum.

According to our earlier definition of equilibrium, the state ultimately reached by this system when we removed the constraint (the insulation) is the equilibrium one. We can see here that, unlike the ball on the hill, we cannot determine whether this system is at equilibrium or not simply from its energy: the total energy of the system remained constant. In general for a thermodynamic system, whether or not the system is at equilibrium depends not on its total energy, but on how that energy is internally distributed.

Clearly, it would be useful to have a function that could predict the internal distribution of energy at equilibrium. The function that does this is the *entropy*. To understand this, let's return to our copper blocks. Initially, the two copper blocks are separated by a thermal bar-



Figure 2.8. (a) Probability of one of two copper blocks of equal mass in thermal equilibrium having *e* units of energy when the total energy of the two blocks is 20 units. (b) Ω , number of states available to the system (combinations of energy distribution) as a function of ε .

rier and we can think of each as an isolated system. We assume that each has an internal energy distribution that is at or close to the most probable one; i.e., each is internally at equilibrium. Each block has its own function Ω (which we denote as $\Omega_{\rm l}$ and $\Omega_{\rm r}$ for the left and right block respectively) that gives the number of states accessible to it at a particular energy distribution. We assume that initial energy distribution is not the final one, so that when we remove the insulation, the energy distribution of system will spontaneously change. In other words:

$$\Omega_l^i \neq \Omega_l^f$$
 and $\Omega_r^i \neq \Omega_l^i$

where we use the superscripts *i* and *f* to denote initial and final respectively.

M. White

Geochemistry

2.41

2.45

2.48

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

When the left block has energy e_i , it can be in any one of $\Omega_1 = \Omega(e)$ possible states, and the right block can be in any one of $\Omega_r = \Omega(E-e)$ states. Both \mathcal{P} and Ω are multiplicative, so the total number of possible states after we remove the insulation, Ω , will be:

$$\Omega(e) = \Omega_l(e) \times \Omega_l(E - e)$$

 $\ln Q = \ln C + \ln \Omega$

To make \mathcal{P} and Ω additive we simply take the log:

$$\ln\Omega(e) = \ln\Omega_{1}(e) + \ln\Omega_{2}(E - e)$$
2.40

and

As additive properties, $\ln P$ and $\ln \Omega$ are consistent with our other extensive state variables (e.g., E, V).

We want to know which energy distribution, i.e., values of e and E-e, is the most likely one, because that corresponds to the equilibrium state of the system. This is the same as asking where the probability function, $\mathcal{P}(\mathbf{e})$, is maximum. Maximum values of functions have the useful property that they occur at points where the derivative of the function is 0. That is, a maximum of function f(x) will occur where $df(x)/d(x) = 0^{\dagger}$. Thus the maximum value of $\mathcal{P}(e)$ in Figure 2.8 occurs where $d\mathcal{P}/de = 0$. The most probably energy distribution will therefore occur at:

$$\frac{\partial \mathcal{P}(e)}{\partial e} = 0$$
 or equivalently $\frac{\partial \ln \mathcal{P}(e)}{\partial e} = 0$ 2.42

(we use the partial differential notation to indicate that, since the system is isolated, all other state variables are held constant). Substituting equation 2.41 into 2.42, we have:

$$\frac{\partial \ln \mathcal{P}(e)}{\partial e} = \frac{\partial (\ln C + \ln \Omega(e))}{\partial e} = \frac{\partial \ln \Omega(e)}{\partial e} = 0$$
 2.43

(since C is a constant). Then substituting 2.40 into 2.43 we have:

$$\frac{\partial \ln \Omega(e)}{\partial e} = \frac{\partial \ln \Omega_{l}}{\partial e} + \frac{\partial \Omega_{r}(E-e)}{\partial e} = 0$$
 2.44

$$\frac{\partial \ln \Omega_l}{\partial e} = -\frac{\partial \Omega_r (E - e)}{\partial e}$$

so the maximum occurs at:

The maximum then occurs where the function $\partial \ln \Omega / \partial e$ for the two blocks are equal (the negative sign will cancel because we are taking the derivative $\partial f(-e)/\partial e$). More generally, we may write:

$$\frac{\partial \ln \Omega_l^f(E_l^f)}{\partial E_l^f} = \frac{\partial \ln \Omega_r^f(E_r^f)}{\partial E_r^f}$$
 2.46

Notice two interesting things: the equilibrium energy distribution is the one where $\ln \Omega$ is maximum (since it proportional to \mathcal{P} and where the function $\partial \ln \Omega / \partial E$ of the two blocks are equal. It would appear that both are very useful functions. We define entropy, S, as:

$$S = k \ln \Omega$$

$$\beta = \frac{\partial \ln \Omega}{\partial E}$$
2.47^g
2.48

and a function
$$\beta$$
 such that:

where k is some constant (which turns out to be Boltzmann's constant or the gas constant; the choice depends on whether we work in units of atoms or moles). *The function S then has the property that it is naximum at equilibrium and*
$$\beta$$
 has the property that it is the same in every part of the system at equilibrium.

⁺ Either a maximum or minimum can occur where the derivative is 0, and a function may have several of both; so some foreknowledge of the properties of the function of interest is useful in using this property.

This equation, which relates microscopic and macroscopic variables, is inscribed on the tombstone of Ludwig Boltzmann (1844-1906), the Austrian physicist responsible for it.

Geochemistry

CHAPTER 7: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

Entropy also has the interesting property that in any spontaneous reaction, the total entropy of the system plus its surroundings must increase. In our example, this is a simple consequence of the observation that the final probability, $\mathcal{R}(E)$, and therefore also Ω , will be maximum and hence never be less than the original one. Because of that, the final number of accessible states must exceed the initial number and:

$$\ln \Omega_l^f(E_l^f) + \ln \Omega_r^f(E_r^f) \ge \ln \Omega_l^i(E_l^i) + \ln \Omega_r^i(E_r^i)$$

$$[\ln \Omega_l^f(E_l^f) - \ln \Omega_l^i(E_l^i)] \ge -[\ln \Omega_r^f(E_r^f) - \ln \Omega_r^i(E_r^i)]$$
2.49

rearranging:

The quantities in brackets are simply the entropy changes of the two blocks. Hence:

$$\Delta S_1 \ge -\Delta S_r \tag{2.50}$$

In other words, any decrease in entropy in one of the blocks must be at least compensated for by an increase in entropy of the other block.

For an irreversible process, that is, a spontaneous one such as thermal equilibrium between two copper blocks, we cannot determine exactly the increase in entropy. Experience has shown, however, that the increase in entropy will always exceed the ratio of heat exchanged to temperature. Thus the mathematical formulation of the second law is:

$$\left| \frac{\mathrm{d}\mathbf{Q}}{\mathrm{T}} \le \mathrm{d}\mathbf{S} \right| \qquad 2.51$$

Like the first law, equation 2.51 cannot be derived or formally proven; it is simply a postulate that has never been contradicted by experience. For a reversible reaction, i.e., one that is never far from equilibrium and therefore one where dQ is small relative to T,

$$dS = \frac{dQ_{rev}}{T}$$
 2.52

(see the box "the second law in the reversible case"). In thermodynamics, we restrict our attention to systems that are close to equilibrium, so equation 2.52 serves as an operational definition of entropy.

2.4.2.1 Microscopic Interpretation of Temperature

Let's now return to our function β . The macroscopic function having the property of our new function β is temperature. The relation of temperature to β is

$$kT = 1/\beta$$
 2.53

and

$$kT = 1/\beta 2.53
\frac{1}{kT} = \frac{\partial \ln \Omega}{\partial E} 2.54$$

Equation 2.53 provides a statistical mechanical definition of temperature. We can easily show that T is measure of the energy per degree of freedom. Do to this though, we need one other relationship, which we introduce without proof. This is that Ω increases roughly with E as:

$$\Omega \propto E^{\gamma}$$

where *f* is the number of degrees of freedom of the system (which in turn is proportional to the number of atoms or molecules in the system times the modes of motion, e.g., vibrational, rotational, translational, available to them). Hence:

> $\beta = \frac{\partial \ln \Omega}{\partial F} \propto \frac{f}{F}$ 2.55 $T \propto \frac{E}{kf}$

Substituting $T = 1/\beta$, then

2.4.2.2 Entropy and Volume

Our discussion of entropy might leave the impression that entropy is associated only with heat and temperature. This is certainly not the case. Our first example, that of the gases in the box, is a good demonstration of how entropy changes can also accompany isothermal processes. When the partition

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

is removed and the gases mix, there is an increase in the number of states accessible to the system. Before the partition is removed, there is only one state accessible to the system (here "accessible states" means distribution of red and black molecules between the two sides of the box), so $\Omega = 1$. Suppose that after we remove the partition, we find the system in a state is one where there is one molecule of each kind on each side (the most probable case). There are four such possible configurations, so $\Omega = 4$. The entropy change has thus been

$\Delta S = k(\ln 4 - \ln 1) = k \ln 4 = 2k \ln 2$

From the macroscopic perspective, we could say that the red gas, initially confined to the left volume, expands into the volume of the entire box, and the black gas expands from the right half to the entire volume. Thus entropy changes accompany volume changes.

2.4.2.3 SUMMARY

It is often said that entropy is a measure of the randomness of a system. From the discussion above, we can understand why. Entropy is a function of the number of states accessible to a system. Because there are more states available to a system when energy or molecules are "evenly" or "randomly" distributed than when we impose a specific constraint on a system (such as the thermal insulation between the blocks or the partition between the gases), there is indeed an association between randomness and entropy. When we remove the insulation between the copper blocks, we allow energy to be randomly distributed between them. In the example of the combustion of gasoline, before combustion all oxygen atoms are constrained to be associated with oxygen molecules. After combustion, oxygen is randomly distributed between water and CO_2 molecules.

More precisely, we may say that an increase in entropy of a system corresponds to a decrease in knowledge of it. In the example of our two gases in the box, before the partition is removed, we know all red molecules are located somewhere in the left half of the box and all black ones somewhere in the right half. After the partition is removed, we know only that the molecules are located somewhere within the combined volume. Thus our knowledge of the location of the molecules decreases in proportion to the change in volume. Molecules in ice are located at specific points in the crystal lattice. When ice melts, or evaporates, molecules are no longer constrained to specific locations: there is an increase in entropy of H_2O and a corresponding decrease in our knowledge of molecular positions. When we allowed the two copper blocks to come to thermal equilibrium, entropy increased. There were more possible ways to distribute energy after the blocks equilibrated than before. As a result, we knew less how energy was distributed after removing the insulation.

As a final point, we emphasize that the second law does not mean we cannot decrease the entropy of a "system". If that were so, the organization of molecules we call life would not be possible. However, if the entropy of a system is to decrease, the entropy of the surroundings must increase. Thus we can air condition a room, but the result is that the surroundings (the "outside") are warmed by more than the air in the room is cooled. Organisms can grow, but in doing so they inevitably, through consumption and respiration, increase the entropy of their environment. Thus we should not be surprised to find that the entropy of the manure is greater than that of hay plus oxygen.

2.4.3 Integrating Factors and Exact Differentials

A theorem of mathematics states that any inexact differential that is a function of only two variables can be converted to an exact differential. dW is an inexact differential, and dV is an exact differential. Since $dW_{rev} = -PdV$, dW_{rev} can be converted to a state function by dividing by P since

$$\frac{dW_{rev}}{P} = -dV \tag{2.56}$$

and V is a state function. Variables such as P which convert non-state functions to state functions are termed *integrating factor*. Similarly, for a *reversible* reaction, heat can be converted to the state function entropy by dividing by T:

Geochemistry

CHAPTER 7: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

$$\frac{dQ_{rev}}{T} = dS$$
 2.57

Thus temperature is the integrating factor of heat. Entropy is a state function and therefore an exact differential. Therefore equation 2.57 is telling us that although the heat gained or lost in the transformation from state 1 to state 2 will depend on the path taken, for a reversible reaction the ratio of heat gained or lost to temperature will always be the same, regardless of path.

If we return to our example of the combustion of gasoline above, the second law also formalizes our experience that we cannot build a 100% efficient engine: the transformation from state 1 to state 2 cannot be made in such a way that all energy is extracted as work; some heat must be given up as well. In this sense, the automobile radiator is necessitated by the second law.

Where P-V work is the only work of interest, we can combine the first and second laws as:

$$dU \le TdS - PdV$$

The implication of this equation is that equilibrium is approached at prescribed S and V, the energy of the system is minimized. For the specific situation of a reversible reaction where dS = dQ/T, this becomes

$$\mathbf{dU}_{rev} = \mathbf{T}\mathbf{dS}_{rev} - \mathbf{P}\mathbf{dV}$$
 2.58

This expresses energy in terms of its natural or *characteristic variables*, S and V. The characteristic variables of a function are those that give the simplest form of the exact differential. Since neither T nor P may have negative values, we can see from this equation that energy will always increase with increasing entropy (at constant volume) and that energy will decrease with increasing volume (at constant entropy). This equation also relates all the primary state variables of thermodynamics, U, S, T, P, and V. For this reason, it is sometimes called the *fundamental equation of thermodynamics*. We will introduce several other state variables derived from these 5, but these will be simply a convenience.

By definition, an adiabatic system is one where dQ = 0. Since $-dQ_{rev}/T = dS_{rev}$ (equation 2.52), it follows that for a reversible process, an adiabatic change is one carried out at constant entropy, or in other words,

Example 2.1. Entropy in Reversible and Irreversible Reactions

Air conditioners work by allowing freon contained in a closed system of pipes to evaporate in the presence of the air to be cooled, then recondensing the freon (by compressing it) on the warm or exhaust side of the system. Let us define our "system" as only the freon in the pipes. The system is closed since it can exchange heat and do work but not exchange mass. Suppose our system is contained in an air conditioner maintaining a room at 20° C or 293 K and exhausting to outside air at 303 K. Let's assume the heat of evaporation of the coolant (the energy required to transform it from liquid to gas), is 1000 joules. During evaporation, the heat absorbed by the coolant, dQ, will be 1000 J. During condensation –1000 J will be given up by the system. For each cycle, the *minimum* entropy change during these transformations is easy to calculate from equation 2.51:

Evaporation:

$$dS \ge \frac{dQ}{T} = 1000/293 = 3.413 \text{ J/K}$$
Condensation:

$$dS \ge \frac{dQ}{T} = -1000/303 = -3.300 \text{ J/K}$$

Condensation:

The minimum net entropy change in this cycle is the sum of the two or 3.413 - 3.300 = 0.113 J/K. This is a "real" process and irreversible, so the entropy change will be greater than this.

If we performed the evaporation and condensation isothermally at the equilibrium condensation temperature, i.e. reversibly, then this result gives the exact entropy change in each case. In this imaginary reversible reaction, where equilibrium is always maintained, there would be no net entropy change over the cycle. But of course no cooling would be achieved either, so it would be pointless from a practical viewpoint. It is nevertheless useful to assume this sort of reversible reaction for the purposes of thermodynamic calculations, because exact solutions are obtained.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

an isoentropic change. For adiabatic expansion or compression therefore, dU = -PdV.

2.5 ENTHALPY

We have now introduced all the fundamental variables of thermodynamics, T, S, U, P and V. Everything else can be developed and derived from these functions. Thermodynamicists have found it convenient to define several other state functions, the first of which is called enthalpy. Enthalpy is a composite function and is the sum of the internal energy plus the product PV:

$$\mathbf{H} \equiv \mathbf{U} + \mathbf{PV}$$
 2.59

As is the case for most thermodynamic functions, it is enthalpy changes rather than absolute enthalpy that are most often of interest. For a system going from state 1 to state 2, the enthalpy change is:

$$H_2 - H_1 = U_2 - U_1 + P_2 V_2 - P_1 V_1$$
 2.60

The First Law states:

$$H_2 - H_1 = \Delta Q + \Delta W + P_2 V_2 - P_1 V_1$$

then:
$$\Delta H = \Delta Q_P + \Delta W + P \Delta V$$
 2.61

so:

If pressure is constant, then:
$$\Delta H = \Delta Q_P + \Delta W + P \Delta V$$

 $U_2 - U_1 = \Delta Q + \Delta W$

(we use the subscript P in ΔQ_P to remind us that pressure is constant). In thermodynamics, PV work is often the only kind of work of interest. If the change takes place at constant pressure and P-V work is the only work done by the system, then the last two terms cancel and enthalpy is simply equal to the heat gained or lost by the system:

or in differential form:
$$\Delta H = \Delta Q_P$$

 $dH = dQ_p$ 2.62

H is a state function because it is defined in terms of state functions U, P, and V. Because enthalpy is a state function, dQ must also be a state function under the conditions of constant pressure and the only work done being PV work.

More generally, the enthalpy change of a system may be expressed as:

$$dH = dU + VdP + PdV$$

or at constant pressure as:

$$dH = dU + PdV 2.63$$

In terms of its characteristic variables, it may also be expressed as:

$$dH \le TdS + VdP \qquad 2.64$$

From this it can be show that H will be at a minimum at equilibrium when S and P are prescribed:

$$\mathbf{dH}_{rev} = \mathbf{TdS}_{rev} + \mathbf{VdP}$$
 2.65

The primary value of enthalpy is measuring the energy consumed or released in changes of state of a system. For example, how much energy is given off by the reaction:

$$2H_2 + O_2 \rightleftharpoons 2H_2O$$

To determine the answer we could place hydrogen and oxygen in a well-insulated piston-cylinder maintaining constant pressure. We would design it such that we could easily measure the temperature before and after reaction. Such an apparatus is known as a *calorimeter*. By measuring the temperature before and after the reaction, and knowing the *heat capacity* of the reactants and our calorimeter, we could determine the enthalpy of this reaction. This enthalpy value is often also called the *heat of reaction* or *heat of formation* and is designated ΔH_r (or ΔH_f). Similarly, we might wish to know how much heat is given off when NaCl is dissolved in water. Measuring temperature before and after reaction would allow us to calculate the *heat of solution*. The enthalpy change of a system that undergoes melting is known as the heat of fusion or heat of melting, ΔH_m (this quantity is sometimes denoted ΔH_i , we will use the subscript *m* to avoid confusion with heat of formation); that of a system undergoing boiling is known as the heat of vaporization, ΔH_{y} . As equation 2.65 suggests, measuring enthalpy change is also a convenient way of determining the entropy change.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

At this point, it might seem that we have wandered rather far from geochemistry. However, we shall shortly see that functions such as entropy and enthalpy and measurements of such things as heats of solutions and heats of melting are essential to predicting equilibrium geochemical systems.

2.6 НЕАТ САРАСІТУ

It is a matter of every day experience that the addition of heat to a body will raise its temperature. We also know that if we bring two bodies in contact, they will eventually reach the same temperature. In that state, the bodies are said to be in thermal equilibrium. However, thermal energy will not necessarily be partitioned equally between the 2 bodies. It would require half again as much heat to increase the temperature of 1g of quartz by 1° C as it would to increase the temperature of 1g of iron metal by 1° C. (We saw that temperature is a measure of the energy per degree of freedom. It would appear then that quartz and iron have different degrees of freedom per gram, something we will explore below.) *Heat capacity* is the amount of heat (in joules or calories) required to raise the temperature of a given amount (usually a mole) of a substance by 1 K. Mathematically, we would say:

$$C = \frac{dQ}{dT}$$
 2.66

However, the heat capacity of a substance will depend on whether heat is added at constant volume or constant pressure, because some of the heat will be consumed as work if the volume changes. Thus a substance will have two values of heat capacity: one for constant volume and one for constant pressure.

2.6.1 CONSTANT VOLUME HEAT CAPACITY

Recall that the first law states: dU = dQ + dWIf we restrict work to PV work, this may be rewritten as:

$$dU = dQ - PdV$$

If the heating is carried out at constant volume, i.e., dV = 0, then dU = dQ (all energy change takes the form of heat) and:

$$C_{v} = \left(\frac{\partial U}{\partial T}\right)_{V}$$
 2.67

The value of C_V for an ideal gas is $^3/_2R$ where R is the gas constant, as is shown in the boxed discussion below. Molecular gases, however, are not ideal. Vibrational and rotation modes also come into play, and heat capacity of real gases, as well as solids and liquids, is a function of temperature.

For solids, motion is vibrational and heat capacities depend on vibrational frequencies, which in turn depend on temperature and bond strength (for stronger bonds there is less energy stored as potential energy, hence less energy is required to raise temperature), for reasons discussed below. For incompressible substances such as solids, the difference between C_V and C_P is generally small.

2.6.2 Constant Pressure Heat Capacity

In geochemistry, constant volume situations are rare and temperature changes at constant pressure are of greater interest. Equation 2.61 states that $\Delta H = \Delta Q_p$. Substituting this expression in equation 2.66 we have:

$$C_{P} = \left(\frac{\partial H}{\partial T}\right)_{P}$$
 2.68

Thus enthalpy change at constant pressure may also be expressed as:

$$dH = C_P dT 2.69$$

In geochemistry, it will be mainly constant pressure heat capacity that will be of primary interest to us, because in natural situations, volumes do not usually remain fixed.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

2.6.3 Energy Associated with Volume and the Relationship between C_V and C_D

Constant pressure and constant temperature heat capacities are different because there is energy associated (work done) with expansion and contraction. Thus how much energy we must transfer to a substance to raise its temperature will depend on whether some of this energy will be consumed in this process of expansion. These energy changes are due to potential energy changes associated with changing the position of an atom or molecule in the electrostatic fields of its neighbors. The difference between C_v and C_p reflects this energy associated with volume. Let's now determine what this difference is.

We can combine relations 2.67 and 2.68 as:

$$C_{P} - C_{V} = \left(\frac{\partial H}{\partial T}\right)_{P} - \left(\frac{\partial U}{\partial T}\right)_{V}$$
 2.70

From this, we may derive the following relationship:

$$C_{P} - C_{V} = \left(\frac{\partial U}{\partial T}\right)_{P} - \left(\frac{\partial U}{\partial T}\right)_{V} + P\left(\frac{\partial V}{\partial T}\right)_{P}$$
 2.71

and further:

$$C_{P} - C_{V} = \left[\left(\frac{\partial U}{\partial V} \right)_{T} + P \right] \left(\frac{\partial V}{\partial T} \right)_{P}$$
 2.72

It can also be shown that, for a reversible process:

$$\left(\frac{\partial U}{\partial V}\right)_T = T\frac{\alpha}{\beta} - P \tag{2.73}$$

 $(\partial U/\partial V)_T$ is the energy associated with the volume occupied by a substance and is known as the *internal pressure* (P_{int}, which we introduced earlier in our discussion of the van der Waals law, e.g., equation 2.17). It is a measure of the energy associated with the forces holding molecules or atoms together. For real substances, energy changes associated with volume changes reflect potential energy increases associated with increase separation between charged molecules and/or atoms; there are no such forces in an ideal gas, so this term is 0 for an ideal gas. Substituting 2.73 into 2.72, we obtain:

$$C_P - C_v = T\overline{V}\frac{\alpha^2}{\beta}$$
 2.74

Thus the difference between C_p and C_v will depend on temperature and pressure for real substances. The terms on the right will always be positive, so that C_p will always be greater than C_v . This accords with our expectation, since energy will be consumed in expansion when a substance is heated at con-

Example 2.2: Measuring Enthalpies of Reaction

Sodium reacts spontaneously and vigorously with oxygen to form Na₂O. The heat given off by this reaction is the enthalpy of formation ΔH_f of Na₂O. Suppose that you react 23 g of Na metal with oxygen in a calorimeter that has the effective heat capacity of 5 kg of water. The heat capacity of water is 75.3 J/mol K. If the calorimeter has a temperature of 20°C before the reaction and a temperature of 29.9°C after the reaction, what is ΔH_f of Na₂O? Assume that the Na₂O contributes negligibly to the heat capacity of the system.

Answer: The heat capacity of the calorimeter is

$$75.3$$
 J/mol K × 5000 g ÷ 18 g/mol = 20917 J/K.

The heat required to raise its temperature by 9.9 K is then

$$0.9 \times 20917 = 207.08 \text{ kJ}$$

which is the enthaply of this reaction. Our experiment created 0.5 moles of Na₂O, so Δ H is –414.16 kJ/mol.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

stant pressure, whereas this will not be the case for heating at constant volume. For an ideal gas, $C_p-C_v = R$.

For gases, it is fairly straightforward to measure either C_p or C_v . It is impractical to measure C_v for solids and liquids. Thus only experimentally determined values of C_p are available for solids and liquids, and values of C_v must be obtained from equation 2.74 when required.

We found earlier that C_p is the variation of heat with temperature at constant pressure. How does this differ from the variation of energy with temperature at constant volume? To answer this question, we rearrange equation 2.71 and substitute C_v for $(\partial U/\partial T)_v$ and $V\alpha$ for $(\partial V/\partial T)_p$. After simplifying the result, we obtain (on a molar basis):

$$\left(\frac{\partial U}{\partial T}\right)_{P} = C_{P} - P\overline{V}\alpha \qquad 2.75$$

For an ideal gas, the term PV α reduces to R, so that $(\partial U/\partial T)_P = C_p - R$. $C_p - R$ may be shown to be equal to C_v , so the energy change with temperature for an ideal gas is the same for both constant pressure and constant volume conditions. This is consistent with the notion that the difference between C_p and C_v reflects the energy associated with and changing distances between atoms and molecules in the presence of attractive forces between them. In an ideal gas, there are no such forces, hence $(\partial U/\partial T)_P = (\partial U/\partial T)_V$.

2.6.4 HEAT CAPACITY OF Solids: A Problem in QUANTUM Physics

As we shall see, knowledge of the heat capacity of substances turns out to be critical to determining properties such as enthalpy and entropy, and, ultimately, to predicting chemical equilibrium. The heat capacity of a substance reflects the internal motion of its atoms. There are 3 kinds of motion available to atoms and molecules: translational, vibrational and rotational^t, but often one or more of these modes will not be available and not contribute to the energy of a substance. For gases at low temperature, only rotational and translational motions are important (for a monatomic gas, only translational modes are available), while only vibrational motions are important for solids (translational modes are available to solids, which is why solids have finite vapor pressures, but they are extremely improbable, which is why vapor pressures of solids are very small and can usually be neglected). Twice as much energy is typically required to raise the temperature of a vibrational mode by 1 K as for a translational mode. This is because vibration involves both kinetic and potential energy of two or more atoms. Also, vibrational modes do not accept much energy at low temperature. This latter phenomenon is not predicted by classical physics; as a result 19th century physicists were puzzled by the temperature dependence of heat capacity. In 1869 James Maxwell referred to the problem as "the greatest difficulty yet encountered in molecular theory." The solution required a more radical revision to physics than Maxwell imagined: the heat capacity problem turned out to be one of the first indications of the inadequacy of classical physics. An understanding of the dependence of heat capacity on temperature was only achieved in the 20th century with the aid of quantum physics. A complete theoretical treatment of heat capacity of real substances is beyond the scope of this book. However, even the few statements we will make will require us to make another excursion into statistical mechanics, a closely related field to discover the Boltzmann Distribution Law. What we learn will be of considerable use in subsequent chapters.

2.6.4.1 The Boltzmann Distribution Law

Consider a mineral sample, A, in a heat bath, B (B having much more mass than A), and assume they are perfectly isolated from their surroundings. The total energy of the system is fixed, but the energy of A and B will oscillate about their most probable values. The question we ask is *what is the probability that system A is in a state such that it has energy* E_A ?

 $^{^{\}pounds}$ R. Clausius recognized the possibility that molecules might have these three kinds of motion in 1855.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

We assume that the number of states accessible to A when it has energy E_A is some function of energy, i.e.,

$$\Omega(a) = \Omega(E_A)$$
 2.76

Following the basic postulate, we also assume that all states are equally probable and that the probability of a system having a given energy is simply proportional to the number of states the system can assume when it has that energy:

$$\mathcal{P} = C \ \Omega \tag{2.39}$$

where C is a constant. Thus the probability of A being in state a with energy E_A is:

$$\mathcal{P}_{a} = C_{A} \Omega(E_{A})$$
 2.77

Since the total energy of the two system is fixed, system B will have some fixed energy E_B when A is in state *a* with energy E_A , and:

$$E_{\rm B} = E - E_{\rm A}$$

where E is the total energy of the system. As we mentioned earlier, Ω is multiplicative, so the number of states available to the total system, A + B, is the product of the number of states available to A times the states available to B:

$$\Omega_{\text{Total}} = \Omega_{\text{A}} \Omega_{\text{B}}$$

If we stipulate that A is in state *a*, then Ω_A is 1 and the total number of states available to the system in that situation is just Ω_B :

$$\Omega_{\text{Total}} = 1 \times \Omega_{\text{B}} = \Omega(\text{E}_{\text{B}})$$

Thus the probability of finding A in state a is equal to the probability of finding B in one of the states associated with energy E_B , so that:

$$\mathcal{P}_{a} = C_{B}\Omega(E_{B}) = C_{B}\Omega(E - E_{A}) = C_{B}exp[\ln \Omega (E - E_{A})]$$
2.78

where E is the total energy of the system (i.e., $E = E_A + E_B$). We can expand $\ln \Omega$ (E-E_A) as a Taylor Series about E:

$$\ln\Omega(E - E_A) = \ln\Omega(E) - E_A \left(\frac{d\,\ln\Omega(E)}{dE}\right) + \dots$$
 2.79

and since B is much larger than A, $E >> E_A$, higher order terms may be neglected.

Substituting β for $\partial \ln \Omega(E)/dE$ (Equ 2.48), we have:

$$\Omega(E - E_A) = \exp(\ln \Omega(E) - E_A\beta) = \Omega(E)e^{-\beta E_A}$$

$$\mathcal{P}_a = C_B \Omega(E)e^{-\beta E_A} \qquad 2.80$$

and

Since the total energy of the system, E, is fixed, $\Omega(E)$ must also be fixed, i.e., constant, so:

$$\mathcal{P}_{a} = \mathrm{Ce}^{-\beta \mathrm{E}_{\mathrm{A}}} \qquad 2.81$$

Substituting 1/kT for β (Equ. 2.53), we have:

$$\mathcal{P}_{a} = Ce^{-E_{A}/kT}$$

 $C = 1/\sum e^{-\beta E_i}$

We can deduce the value of the constant C by noting that $\Sigma P_i = 1$, i.e., the probabilities over all energy levels must sum to one (because the system *must always* be in one of these states). Therefore:

$$\Sigma \mathcal{P}_1 = C \Sigma e^{-\beta E_i} = 1 \qquad 2.82$$

so that

Generalizing our result, the probability of the system being in state *i* corresponding to energy ε_i is:

2.83

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

$$\mathcal{P}_{i} = \frac{e^{-\varepsilon_{i}/kT}}{\sum_{n} e^{-\varepsilon_{n}/kT}} \qquad 2.84$$

This equation is the *Boltzmann distribution law*^{*}, and one of the most important equations in statistical mechanics. Though we derived it for a specific situation and introduced an approximation (the Taylor Series expansion), these were merely conveniences; the result is very general. If we define our "system" as an atom or molecule, then this equation tells us the probability of an atom having a given energy value, ε_i . This is the statistical mechanical interpretation of this equation. The basic tenet of quantum physics is that energy is quantized: only discrete values are possible. The Boltzmann distribution law tells the probability of an atom having the energy associated with quantum level *i*.

The Boltzmann distribution law says that the population of energy levels decreases exponentially as the energy of that level increases (*energy among atoms is like money among men: the poor are many and the rich few*). A hypothetical example is shown in Figure 2.9.



Figure 2.9. Occupation of vibrational energy levels calculated from the Boltzmann distribution. The probability of an energy level associated with the vibrational quantum number n is shown as a function of n for a hypothetical diatomic molecule at 273 K and 673 K.

2.6.4.2 The Partition Function

The denominator of Equ. 2.84, which is the probability normalizing factor or the *sum of the energy distribution over all accessible states*, is called the *partition function* and is denoted *Q*:

$$Q = \sum_{i} e^{-\varepsilon_i / kT}$$
 2.85

The partition function is a key variable in statistical mechanics and quantum physics. It is related to macroscopic variables with which we are already familiar, namely energy and entropy. Let's examine these relationships.

We can compute the total internal energy of a system, U, as the average energy of the atoms times the number of atoms, *n*. To do this we need to know how energy is distributed among atoms. Macroscopic systems have a very large number of atoms (~10²³, give or take a few in the exponent). In this case, the number of atoms having some energy ε_i is proportional to the probability of one atom having this energy. So to find the average, we take the sum over all possible energies of the product of energy times the possibility of an atom having that energy. Thus the internal energy of the system is just:

^{*} We now understand and interpret this law in terms of quantum physics, but Boltzmann formulated it 30 years before Planck and Einstein laid the foundations of quantum theory. Ludwig Boltzmann's work in the second half of the nineteenth century laid the foundations of statistical mechanics and paved the way for quantum theory in the next century. His work was heavily attacked by other physicists of the time, who felt physics should deal only with macroscopic observable quantities and not with atoms, which were then purely hypothetical constructs. These attacks contributed to increasingly frequent bouts of depression, which ultimately led to Boltzmann's suicide in 1906. Ironically and sadly, this was about the time that Perrin's experiments with Brownian motion, Millikan's oil drop experiment, and Einstein's work on the photoelectric effect confirmed the discrete nature of mass, charge, and energy, and thereby the enduring value of Boltzmann's work.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

$$U = n \sum_{i} \varepsilon_{i} \mathcal{P}_{i} = \frac{n \sum \varepsilon_{i} e^{-\varepsilon_{i}/kT}}{Q}$$
 2.86

The derivative of *Q* with respect to temperature (at constant volume) can be obtained from 2.85:

$$\left(\frac{\partial Q}{\partial T}\right)_{V} = \frac{1}{kT} \sum \varepsilon_{i} e^{-\varepsilon_{i}/kT}$$
 2.87

Comparing this with equation 2.86, we see that this is equivalent to:

$$\left(\frac{\partial Q}{\partial T}\right)_{V} = \frac{Q}{nkT^{2}}U$$
2.88

It is also easy to show that $\partial \ln Q / \partial T = 1 / Q \partial Q / \partial T$, so the internal energy of the system is:

$$U = nkT^{2} \left(\frac{\partial \ln Q}{\partial T}\right)_{V}$$
 2.89

For 1 mole of substance, *n* is equal to Avagadro's number, N_0 . Since $R = N_0 k$, equation 2.89, when expressed on a molar basis, becomes:

$$U = RT^2 \left(\frac{\partial \ln Q}{\partial T}\right)_V$$
 2.90

We should not be surprised to find that entropy is also related to *Q*. This relationship, the derivation of which is left to you (Problem 2.12), is:

$$S = \frac{U}{T} + R \ln Q \tag{2.91}$$

Since the partition function is a sum over all possible states, it might appear that computing it would be a formidable, if not impossible, task. As we shall see, however, *the partition function can very often be approximated to a high degree of accuracy by quite simple functions*. The partition function and Boltzmann distribution will prove useful to us in subsequent chapters in discussing several geologically important phenomena such as diffusion and the distribution of stable isotopes between phases, as well as in understanding heat capacities.

2.6.4.3 Energy Distribution in Solids

According to quantum theory, all modes of motion are quantized. Consider, for example, vibrations of atoms in a hydrogen molecule. Even at absolute zero temperature, the atoms will vibrate at a ground state frequency. The energy associated with this vibration will be:

$$\varepsilon_0 = \frac{1}{2} h v_0 \qquad 2.92$$

where *h* is Planck's constant and v is the vibrational frequency of the ground state. Higher quantum levels have higher frequencies (and hence higher energies) that are multiples of this ground state:

$$\varepsilon_{n} = (n + \frac{1}{2})hv_{0} \qquad 2.93$$

where *n* is the quantum number (an integer ≥ 0).

Now consider a monatomic solid, such as diamond, composed of N identical atoms arranged in a crystal lattice. For each vibration of each atom, we may write an atomic partition function, *q*. Since vibrational motion is the only form of energy available to atoms in a lattice, the atomic partition function may be written as:

$$q = \sum_{m} e^{-\epsilon_{m}/kT} = \sum_{n}^{\infty} e^{-(n+\frac{1}{2})hv_{0}/kT}$$
 2.94

September 16, 2005

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

We can rewrite 2.94 as:

$$q = e^{-hv_0/2kT} \sum_{n}^{\infty} e^{-nhv_0/kT}$$
 2.95

The summation term can be expressed as geometric series, $1 + x + x^2 + x^3 + ...$, where $x = e^{-hv_0/kT}$. Such a series is equal to 1/(1 - x) if x<1. Thus 2.95 may be rewritten in simpler form as:

$$q = \frac{e^{-hv_0/2KI}}{1 - e^{-hv_0/kT}}$$
 2.96

At high temperature, $hv_0/kT \ll 1$, and we may approximate $e^{-hv_0/kT}$ in the denominator of Equ. 2.96 by $1 - hv_0/kT$, so that at high temperature:

$$q \approx \frac{kTe^{-hv_0/2\,kT}}{hv_0}$$
 2.97

Using this relationship, and those between constant volume heat capacity and energy and between energy and the partition function it is possible to show that:

$$C_{\rm v} = 3R \qquad 2.98$$

This is called the *Dulong-Petit Limit* and it holds only where the temperature is high enough that the approximation $e^{-hv_0/kT} = 1-hv_0/kT$ holds. For a solid consisting of N different kinds of atoms, the predicted heat capacity is 3NR. Observations bear out these predictions. For example, at 25°C the observed heat capacity for NaCl, for which N is 2, is 49.7 J/K, whereas the predicted value is 49.9 J/K. Substances whose heat capacity agrees with that predicted in this manner are said to be *fully activated*. The temperature at which this occurs, called the *characteristic* or *Einstein temperature*, varies considerably from substance to substance (for reasons explained below). For most metals, it is in the range of 100 to 600 K. For diamond, however, the Einstein temperature is in excess of 2000 K.

Now consider the case where the temperature is very low. In this case, the $hv_0/kT >>1$ and the denominator of equation 2.99 therefore tends to 1, so that 2.99 reduces to:

$$q \cong e^{-hv_0/2kT}$$
 2.99

The differential with respect to temperature of ln*q* is then simply:

$$\left(\frac{\partial q}{\partial T}\right)_{V} = \frac{hv_{0}}{2kT}$$
 2.100

If we insert this into equation 2.90 and then into 2.80 and differentiate U with respect to temperature, we find that the predicted heat capacity is 0! In actuality, only a perfectly crystalline solid would have 0 heat capacity near absolute 0. Real solids have a small but finite heat capacity.

On a less mathematical level, the heat capacities of solids at low temperature are small because the spacings between the first few vibrational energy levels are large. As a result, energy transitions are large and therefore improbable. Thus at low temperature, relatively little energy will go into vibrational motions.

We can also see from equation 2.93 that the gaps between energy levels depend on the fundamental frequency v_0 . The larger the gap in vibrational frequency, the less likely will be the transition to higher energy states. The ground state frequency in turn depends on bond strength. Strong bonds have higher vibrational frequencies and as a result, energy is less readily stored in atomic vibrations. In general, covalent bonds will be stronger than ionic ones, which in turn are stronger than metallic bonds. Thus diamond, which has strong covalent bonds, has a low heat capacity until it is fully activated, and full activation occurs at very high temperatures. The bonds in quartz and alumina (Al₂O₃) are also largely covalent, and these substances also have low heat capacities until fully activated. Metals, on the other hand, tend to have weaker bonds and high heat capacities.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

Heat capacities are more difficult to predict at intermediate temperatures and require some knowledge of the vibrational frequencies. One simple assumption, used by Einstein[‡], is that all vibrations have the same frequency. The Einstein model provides reasonable predictions of C_v at intermediate and high temperatures, but does not work well at low temperatures. A somewhat more sophisticated assumption was used by Debye[§], who assumed a range of frequencies up to a maximum value, v_D , now called the *Debye frequency*, and then integrated the frequency spectrum. The procedure is too complex for us to treat here. At low temperature, the Debye theory predicts:

$$C_V = \frac{12\pi^4}{5} NR \left(\frac{T}{\theta_D}\right)^3 \qquad 2.101$$



Figure 2.10. Vibrational contribution to heat capacity as a function of kT/hv.

where $\theta_D = hv_D/k$ and is called the *Debye temperature*.

Nevertheless, geochemists generally use empirically determined heat capacities. Constant pressure heat capacities are easier to determine, and therefore more generally available and used. For minerals, which are relatively incompressible, the difference between C_v and C_p is small and can often be neglected. Empirical heat capacity data is generally in the form of the coefficients of polynomial expressions of temperature. The *Maier-Kelley formulation* is:

$$C_P = a + bT - \frac{c}{T^2} \tag{2.102}$$

where *a*, *b*, and *c* are the empirically determined coefficients. The *Haas-Fisher* formulation is:

$$C_{P} = a + bT + \frac{c}{T^{2}} + fT^{2} + gT^{-1/2}$$
 2.103

with *a*, *b*, *c*, *f*, and *g* as empirically determined constants.

Since these formulae and there associated constants are purely empirical (i.e., neither the equations nor constants have a theoretical basis), they should not be extrapolated beyond the calibrated rang.

2.6.5 Relationship of Entropy to Other State Variables

We can now use heat capacity to define the temperature dependency of entropy:

$$\left(\frac{\partial S}{\partial T}\right)_{V} = \frac{C_{V}}{T}$$
 2.104 $\left(\frac{\partial S}{\partial T}\right)_{P} = \frac{C_{P}}{T}$ 2.105

The dependencies on pressure and volume (at constant temperature) are:

[‡] Albert Einstein, though best know for his relativity theories, was also the founder, along with Max Planck, of quantum physics. His work on the quantum basis of heat capacity of solids was published in 1907. Einstein was born in 1879 in Ulm, Germany, and published some of his most significant papers while working as a patent clerk in Bern, Switzerland. He later joined the Prussian Academy of Sciences in Berlin. A dedicated and active pacifist, Einstein left Germany when Hitler came to power in 1933. He later joined the Center for Advanced Studies in Princeton, New Jersey. He died in Princeton in 1955.

[§] Peter Debye (1884-1966) was born in Maastricht, Netherlands (as Petrus Debije), but spent much of his early career in Germany, eventually become director of the Kaiser-Wilhelm-Institut in Berlin. While he was visiting Cornell University in 1940, Germany invaded Holland and Debye simply remained at Cornell, eventually becoming chairman of the Chemistry Department. Debye made numerous contributions to physics and physical chemistry; we shall encounter his work again in the next chapter.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

$$\left(\frac{\partial S}{\partial P}\right)_T = -\alpha V \qquad 2.106 \qquad \left(\frac{\partial S}{\partial V}\right)_T = \frac{\alpha}{\beta} \qquad 2.107$$

2.6.6 Additive Nature of Silicate Heat Capacities

For many oxides and silicates, heat capacities are approximately additive at room temperature. Thus, for example, the heat capacity of enstatite, $MgSiO_3$, may be approximated by adding the heat capacities of its oxide components, quartz (SiO_2) and periclase (MgO). In other words since:

$$SiO_2 + MgO \rightarrow MgSiO_2$$

then

$$C_{p-En} \approx C_{p-Qz} + C_{p-Pe}$$

Substituting values: $C_{p-En} \approx 10.62 + 9.03 = 19.65$ cal/mol-K

The observed value for the heat capacity of enstatite at 300 K is 19.62 cal/mol-K, which differs from our estimate by only 0.1%. For most silicates and oxides, this approach will yield estimates of heat capacities that are within 5% of the observed values. However, this is not true at low temperature. The same calculation for C_{p-En} carried out using heat capacities at 50 K differs from the observed value by 20%.

The explanation for the additive nature of oxide and silicate heat capacities has to do with the nature of bonding and atomic vibrations. The vibrations that are not fully activated at room temperature are largely dependent on the nature of the individual cation-oxygen bonds and not on the atomic arrangement in complex solids.

2.7 The Third Law and Absolute Entropy

2.7.1 STATEMENT OF THE THIRD LAW

The entropies of substances tend toward zero as absolute zero temperature is approached. Or as Lewis and Randall expressed it:

If the entropy of each element in some crystalline state may be taken as zero at the absolute zero of temperature, every substance has a finite positive entropy, but at absolute zero, the entropy may become zero, and does so become in the case of perfectly crystalline substances.

2.7.2 Absolute Entropy

We recall that entropy is proportional to the number of possible arrangements of a system: $S = k \ln \Omega$. At absolute zero, a perfectly crystalline substance has only one possible arrangement, namely the ground state. Hence $S = k \ln 1 = 0$.

The implication of this seemingly trivial statement is that we can determine the absolute entropy of substances. We can write the complete differential for S in terms of T and P as:

$$dS = \left(\frac{\partial S}{\partial T}\right)_{P} dT + \left(\frac{\partial S}{\partial P}\right)_{T} dP \qquad 2.108$$

Substituting equations 2.106 and 2.107, we have:

$$dS = \frac{C_P}{T} dT - \alpha V dP \qquad 2.108$$

The coefficient of thermal expansion is 0 at absolute 0; the choice of 1 atm for the heat capacity integration is a matter of convenience because C_P measurements are made at 1 atm.

Actually, the absolute entropies of real substances tend not to be zero at absolute zero, which is to say they are not 'perfectly crystalline' in the third law sense. A residual entropy, S₀, which reflects such things as mixing of two or more kinds of atoms (elements or even isotopes of the same element) at crystallographically equivalent sites, must also be considered. This *configurational entropy* is important for some geologically important substances such as feldspars and amphiboles. Configurational entropy can be calculated as

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

$$S_{conf} = -R \sum_{j} m_{j} \sum_{i} X_{ij} \ln X_{ij}$$
 2.110

where m_j is the total number of atoms in the *jth crystallographic* site (in atoms per formula unit) and $X_{i,j}$ is the mole fraction of the *i*th atom (element) in the *j*th site. We will return to this equation when we consider multicomponent systems.

2.8 Calculating Enthalpy and Entropy Changes

Example 2.3: Configurational Entropy

Olivine is an example of a solid solution, which we will discuss at length in Chapter 3. Fe and Mg may substitute for each other in the octahedral site. Assuming that the distribution of Fe and Mg within this site is purely random, what is the configurational entropy of olivine of the composition $(Mg_{0.8r}Fe_{0.2})_2SiO_4$?

Answer: To solve this problem, we need to apply equation 2.110. We need only consider the octahedral site containing Fe and Mg, because O and Si are the only kinds of atoms occupying the tetrahedral and anion sites. The values for X for these 2 sites will therefore be 1, and $\ln(1) = 0$, so there is no contribution to configurational entropy.

For the octahedral site, m = 2, $X_1 = X_{Mg} = 0.8$ and $X_2 = X_{Fe} = 0.2$. Therefore, the configurational entropy will be:

$$S_{conf} = -8.314 \times 2 (0.8 \ln(0.8) + 0.2 \ln(0.2)) = 8.32 \text{ J mol}^{-1} \text{K}^{-1}$$

2.8.1 Enthalpy Changes due to Changes in Temperature and Pressure

From equation 2.62, we can see that the temperature derivative of enthalpy is simply the isobaric heat capacity:

$$\left(\frac{\partial H}{\partial T}\right)_{p} = C_{p}$$

dH = C_P dT 2.111

and hence:

Thus the change in enthalpy over some temperature interval may be found as:

$$\Delta H = \int_{T_1}^{T_2} C_P dT \qquad 2.112$$

 $C_{\rm p}$ is often a complex function of temperature, so the integration is essential. Example 2.4 below illustrates how this is done.

Isothermal enthalpy changes refer are those occurring at constant temperature, for example, changes in enthalpy due to isothermal pressure changes. Though pressure changes at constant temperature are relatively rare in nature, hypothetical isothermal paths are useful in calculating energy changes. Since enthalpy is a state property, the net change in the enthalpy of a system depends only on the starting and ending state, i.e., the enthalpy change is pathindependent. Imagine a system consisting of a quartz crystal that undergoes a change in state from 25°C and 1 atm to 500° and 400 atm. How will the enthalpy of this system change? Though in actuality the pressure and temperature changes may have occurred simultaneously, because the enthalpy change is path-independent, we can treat the problem as an isobaric temperature change followed by an isothermal temperature change, as illustrated in Figure Knowing how to calculate isothermal enthalpy 2.11.



Figure 2.13. Transformations on a temperature-pressure diagram. Changes in state variables such as entropy and enthalpy are path independent. For such variables, the transformation paths shown by the solid line and dashed line are equivalent.

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

changes is useful for this reason.

We want to know how enthalpy changes as a function of pressure at constant temperature. We begin from equation 2.76, which expresses the enthalpy change as a function of volume and pressure:

$$dH = dU + VdP + PdV \tag{2.63}$$

By making appropriate substitutions for dU, we can derive the following of enthalpy on pressure:

$$dH = V(1 - \alpha T)dP \qquad 2.112$$

If changes are large, α , β , and V must be considered functions of T and P and and integration performed over the pressure change. The isothermal enthalpy change due to pressure change is thus given by:

$$\Delta H = \int_{P_1}^{P_2} V(1 - \alpha T) dP \qquad 2.113$$

Example 2.4: Calculating Isobaric Enthalpy Changes

How does the enthalpy of a 1 mol quartz crystal change if it is heated from 25° C to 300° C if the temperature dependence of heat capacity can be expressed as $C_p = a + bT - cT^{-2} J/K$ -mol, and a = 46.94, b = 0.0343, and c = 1129680? Assume pressure is constant.

The first step is to convert temperature to kelvins: all thermodynamic formulae assume temperature is in kelvins. So $T_1 = 298$ K and $T_2 = 573$ K. To solve this problem, we need to use equation 2.112. Substituting the expression for heat capacity into Equ. 2.112, we have:

$$\Delta H = \int_{298}^{573} \left(a + bT - cT^{-2}\right) dT = a \int_{298}^{573} dT + b \int_{298}^{573} T dT - c \int_{298}^{573} T^{-2} dT$$

Performing the integral, we have:

$$\Delta H = \left[aT + \frac{b}{2}T^2 + \frac{c}{T}\right]_{298}^{573} = \left[46.94 \times T + \frac{0.0343}{2}T^2 + \frac{1129680}{T}\right]_{298}^{573}$$

Now that we have done the math, all that is left is arithmetic. This is most easily done using a spreadsheet. Among other things, it is much easier to avoid arithmetical errors. In addition, we have a permanent record of what we have done. We might set up a spreadsheet to calculate this problem as follows:

	Values		Formulae & Results	
а	46.94	Н	(a*Temp)+(b*Temp^2/2)+(c_/1	emp)
b	0.0343	H1	19301.98	J/mol
с	1129680	H2	34498.98	J/mol
Temp1	298	$\Delta H =$	15.20	kJ/Mol
Temp2	573			

This example is from Microsoft ExcelTM. One the left, we have written down the names for the various constants in one column, and their values in an adjacent one. Using the Create Names command, we assigned the names in the first column to the values in the second (to avoid confusion with row names, we have named T_1 and T_2 Temp1 and Temp2 respectively; Excel automatically added an underline to 'c', so this constant appears as c_ in our formula). In the column on the right, we have written the formula out in the second row, then evaluated it at T_1 and T_2 in the third and forth rows respectively. The last row, in bold, contains our answer, 15.2 kJ/mole, determined simply by subtracting 'H1' from 'H2' (and dividing by 1000). *Hint*: we need to keep track of units. Excel won't do this for us.

2.8.3 Changes in Enthalpy due to Reactions and Change of State

We cannot measure the absolute enthalpy of substances, but we can determine the enthalpy *changes* resulting from transformations of a system, and they are of great interest in thermodynamics. For this purpose, a system of relative enthalpies of substances has been established. Since enthalpy is a function of both temperature and pressure, the first problem is to establish standard conditions of temperature

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

and pressure to which these enthalpies apply. These conditions, by convention, are 298.15 K and 0.1 MPa (25° C and 1 bar). Under these conditions the elements are assigned enthalpies of 0. Standard state *enthalpy of formation*, or heat of formation, from the elements, ΔH° , can then be determined for compounds by measuring the heat evolved in the reactions that form them from the elements (e.g., Example 2.2). For example, the heat of formation of water is determined from the energy released at constant pressure in the reaction: $H_2 + \frac{1}{2}O_2 \rightarrow H_2O$, which yields a ΔH° of -285.83 kJ/mol, where water is in the liquid state. The minus sign indicates heat is liberated in the reaction, i.e., the reaction is exothermic (a reaction that consumes heat is said to be *endothermic*).

Having established such a system, the enthalpy associated with a chemical reaction is easily calculated using Hess's Law, which is:

$$\Delta H_r = \sum_i v_i \Delta H_{f,i}^0$$
 2.114

where v_i is the stoichiometric coefficient for the *i*th species. In other words, the enthalpy of reaction is just the total enthalpy of the products less the total enthalpy of the reactants. The use of Hess's Law is illustrated in Example 2.5 below.

The *heat of vaporization* of a substance is the energy required to convert that substance from liquid to gas, i.e., to boil it. If the reaction $H_2 + 1/2 O_2 \rightarrow H_2O$ is run to produce water vapor, the ΔH° turns out to be -241.81 kJ/mol. The difference between the enthalpy of formation of water and vapor, 44.02 kJ/mol, is the heat consumed in going from liquid water to water vapor, or This is exactly the amount of energy that would be required to boil 1 mole of water, i.e., convert it from liquid to gas. Analogously, the heat of melting (or fusion) is the enthalpy change in the melting of a substance. Because reaction rates are often very slow, and some compounds are not stable at 298 K and 1 MPa, it is not possible to measure the enthalpy for every compound. However, the enthalpies of formation for these compounds can generally be calculated indirectly.

Example 2.4: Enthalpies (or Heats) of Reaction and Hess's Law

What is the energy consumed or evolved in the hydration of corundum (Al₂O₃) to form gibbsite (Al(OH)₃)? The reaction is:

$$\frac{1}{2}\operatorname{Al}_2\operatorname{O}_3 + \frac{3}{2} \operatorname{H}_2\operatorname{O} \to \operatorname{Al}(\operatorname{OH})_3$$

Answer: We use Hess's Law. To use Hess's Law we need the standard state enthalpies for water, corundum, and gibbsite. These are: Al_2O_3 : -1675.70 kJ/mol, H_2O : -285.83 and $Al(OH)_3$: -1293.13. The enthalpy of reaction is $\Delta H_r = -1293.13 - (0.5 \times -1675.70) - (1.5 \times -285.83) = -26.53 \text{ kJ}$

This is the enthalpy of reaction at 1 bar and 298° K. Suppose you were interested in this reaction under metamorphic conditions such as 300° C and 50 MPa. How would you calculate the enthalpy of reaction then?

2.8.4 Entropies of Reaction

Since

then at constant pressure:

$dS_{rev} = \frac{dQ}{T}$ and (2.57)

 $dS_{rev} = \frac{dH}{T}$ 2.115 Thus at constant pressure, the entropy change in a reversible reaction is simply the ratio of enthalpy

change to temperature.

Entropies are additive properties and entropies of reaction can be calculated in the same manner as for enthalpies, i.e., Hess's Law applies:

$$\Delta S_r = \sum_i v_i \Delta S_{f,i}^0 \qquad 2.116$$

The total entropy of a substance can be calculated as:

 $dH = dQ_{p}$ (2.62)

$$S_{298} = \int_0^{298} \frac{C_P dT}{T} + S_0 + \Delta S_\phi \qquad 2.117$$

September 16, 2005

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

where S_0 is the entropy at 0 K (configurational, or 'third law' entropy) and ΔS_{Φ} is the entropy change associated with any phase change. Compilations for S_{298} are available for many minerals. Table 2.2 lists some heat capacity constants for the power series formula as well as other thermodynamic data for a few geologically important minerals.

2.9 FREE ENERGY

We can now introduce two free energy functions, the Helmholtz Free Energy and the Gibbs Free En-

Phase/	Formula	ΔH_{f}^{o}	S°	ΔG_{f}^{o}	$\overline{\mathrm{V}}$		Cp	
Compound		(kJ/mol)	(J/K-mol)	(kJ/mol)	(cc/mol)*	а	b	с
H ₂ O _g	H ₂ O(gas)	-241.81	188.74	-228.57	24789.00	30.54	0.01029	0
H_2O_1	H ₂ O(liquid)	-285.84	69.92	-237.18	18.10	29.75	0.03448	0
CO_2	CO ₂	-393.51	213.64	-394.39	24465.10	44.22	0.00879	861904
Calcite	CaCO ₃	-1207.30	92.68	-1130.10	36.93	104.52	0.02192	2594080
Graphite	С	0	5.740		5.298			
Diamond	С	1.86	2.37		3.417			
Aragonite	CaCO ₃	-1207.21	90.21	-1129.16	34.15	84.22	0.04284	1397456
α-Qz	SiO ₂	-910.65	41.34	-856.24	22.69	46.94	0.03431	1129680
β-Qz	SiO ₂	-910.25	41.82	-856.24		60.29	0.00812	0
Cristobal.	SiO ₂	-853.10	43.40	-853.10	25.74	58.49	0.01397	1594104
Coesite	SiO ₂	-851.62	40.38	-851.62	20.64	46.02	0.00351	1129680
Periclase	MgO	-601.66	26.94	-569.38	11.25	42.59	0.00728	619232
Magnetite	Fe ₃ O ₄	-1118.17	145.73	-1014.93	44.52	91.55	0.20167	0
Spinel	$MgAl_2O_4$	-2288.01	80.63	-2163.15	39.71	153.86	0.02684	4062246
Ĥem	Fe_2O_3	-827.26	87.61	-745.40	30.27	98.28	0.07782	1485320
Corundum	Al_2O_3	-1661.65	50.96	-1568.26	25.58	11.80	0.03506	3506192
Kyanite	Al_2SiO_5	-2581.10	83.68	-2426.91	44.09	173.18	0.02853	5389871
Andalusite	Al_2SiO_5	-2576.78	92.88	-2429.18	51.53	172.84	0.02633	5184855
Sillimanite	Al_2SiO_5	-2573.57	96.78	-2427.10	49.90	167.46	0.03092	4884443
Almandine	Fe ₃ Al ₂ Si ₃ O ₁₂	-5265.5	339.93	-4941.73	115.28	408.15	0.14075	7836623
Grossular	$Ca_3Al_2Si_3O_{12}$	-6624.93	254.68	-6263.31	125.30	435.21	0.07117	11429851
Albite	NaAlSi ₃ O ₈	-3921.02	210.04	-3708.31	100.07	258.15	0.05816	6280184
K-feldspar	KAlSi ₃ O ₈	-3971.04	213.93	-3971.4	108.87	320.57	0.01804	12528988
Anorthite	$CaAl_2Si_2O_8$	-4215.60	205.43	-3991.86	100.79	264.89	0.06190	7112800
Jadeite	NaAlSi ₂ O ₆	-3011.94	133.47	-2842.80	60.44	201.67	0.04770	4966408
Diospide	CaMgSi ₂ O ₆	-3202.34	143.09	-3029.22	66.09	221.21	0.03280	6585616
Enstatite	MgSiO ₃	-1546.77	67.86	-1459.92	31.28	102.72	0.01983	2627552
Forsterite	Mg_2SiO_4	-2175.68	95.19	-2056.70	43.79	149.83	0.02736	3564768
Clinozo	$Ca_2Al_3Si_3O_{12}(OH)$	-68798.42	295.56	-6482.02	136.2	787.52	0.10550	11357468
Tremolite	$Ca_2MgSi_8O_{22}(OH)_2$	-12319.70	548.90	-11590.71	272.92	188.22	0.05729	4482200
Chlorite	$MgAl(AlSi_3)O_{10}(OH)_8$	-8857.38	465.26	-8207.77	207.11	696.64	0.17614	15677448
Pargasite	NaCa2Mg4Al3Si8O22(OH)2	-12623.40	669.44	-11950.58	273.5	861.07	0.17431	21007864
Phlogopite	KMg ₃ AlSi ₃ O ₁₀ (OH) ₂	-6226.07	287.86	-5841.65	149.66	420.95	0.01204	8995600
Muscovite	$KAl_3Si_3O_{10}(OH)_2$	-5972.28	287.86	-5591.08	140.71	408.19	0.110374	10644096
Gibbsite	Al(OH) ₃	-1293.13	70.08	-1155.49	31.96	36.19	0.19079	0
Boehmite	AlO(OH)	-983.57	48.45	-908.97	19.54	60.40	0.01757	0
Brucite	Mg(OH) ₂	-926.30	63.14	-835.32	24.63	101.03	0.01678	2556424

TABLE 2.2: STANDARD STATE THERMODYNAMIC DATA FOR SOME IMPORTANT MINERALS

Data for the standard state of 298.15 K and 0.1 MPa. ΔH_f^0 is the molar heat (enthalpy) of formation from the elements ; S° is the standard state entropy; V is the molar volume; a, b and c are constants for the heat capacity (C_p) computed as: C_p = a + bT - cT⁻² J/K-mol. Modified from Helgeson et al. (1978). *cc/mol = J/MPa/mol.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

ergy. The Gibbs Free Energy is one of the most useful functions in thermodynamics.

2.9.1 Helmholtz Free Energy

We can rearrange equation 2.58 to read dU - TdS = -PdV. The -PdV term is the work term and the TdS term is the heat function. TdS is the energy unavailable for work. Therefore dU - TdS is the amount of internal energy available for work, or the Free Energy. We define it as A, the Helmholtz Free Energy:

$$A \equiv U - TS \tag{2.118}$$

As usual, we are interested in the differential form (since we are more interested in changes than in absolutes):

$$dA = dU - d(TS) = dU - SdT - TdS$$
 2.119

Example 2.6: Calculating Enthalpy and Entropy Changes

If the heat capacity of steam can be represented by a three-term power series:

$$C_p = a + bT + cT^2$$

with constants a = 36.37 J/K-mol, $b = -7.84 \times 10^3 \text{ J/K}^2$ -mol, and $c = 9.08 \times 10^6 \text{ J/K}^3$ -mol, and the enthaply of vaporization at 100°C is 40.6 kJ/mol, calculate the S and H changes when 1 mol of liquid water at 100° C and 1 atm is converted to steam and brought to 200° C and 3 atm. Assume that with respect to volume, steam behaves as an ideal gas (which, in reality, it is certainly not).

Answer: We need to calculate entropy and enthalpy associated with three changes: the conversion of water to steam, raising the steam from 100° C to 200° C, and increasing the pressure from 1 atm to 3 atm. Since both S and H are state variables, we can treat these three processes separately; our answer will be the sum of the result for each of these processes and will be independent of the order in which we do these calculations.

1. Conversion of water to steam. This process will result in ΔH of 40.6 kJ. For entropy, $\Delta S = \Delta H/T = 40.6/373 = 109 \text{ J/K}$. We converted centigrade to Kelvin, or absolute, temperature.

2. Raising the steam from 100°C to 200°C (from 373 K to 473 K) isobarically. Since heat capacity is a function of temperature, we will have to integrate equation 2.81 over the temperature interval:

$$\int_{T_1}^{T_2} C_P dT = \int_{373}^{473} \left(a + bT + cT^2\right) dT = a \int_{373}^{473} dT + b \int_{373}^{473} T dT + c \int_{373}^{473} T^2 dT = \left[aT + \frac{b}{2}T^2 + \frac{c}{3}T^3\right]_{373}^{4/3}$$

Evaluating this, we find that $\Delta H = (17.20 - 0.88 + 0.32) - (13.57 - 0.55 + 0.16) = 3.469$ kJ. The entropy change is given by:

$$\Delta S = \int_{T_1}^{T_2} \frac{C_P}{T} dT = \int_{373}^{473} \frac{a}{T} dT + \int_{373}^{473} b dT + \int_{373}^{473} cT dT = \left[a \ln T + bT + \frac{c}{2} T^2 \right]_{373}^{473}$$

Evaluating this, we find that $\Delta S = (224.01-3.71+1.02) - (215.37-2.93+.63) = 8.24 \text{ J/K}$.

3. Increasing pressure from 1 atm to 3 atm (0.1 MPa to 0.3 MPa) isothermally. We can use equation 2.116 to determine the enthalpy change associated with the pressure change. On the assumption of ideal gas behavior, we can substitute 1/T for α . Doing so, we find the equation goes to 0; thus there is no enthalpy change associated with a pressure change for an ideal gas. This is in accord with assumption about an ideal gas: namely that there are no forces between molecules, hence no energy is stored as potential energy of attraction between molecules.

The isothermal pressure dependence of entropy is given by equation 2.106. We substitute 1/T for α and RT/P for V and integrate from P₁ to P₂:

$$\Delta S = \int_{P_1}^{P_2} -\frac{1}{T} \frac{RT}{P} dP = \int_{P_1}^{P_2} -\frac{R}{P} dP = -R \left[\ln P \right]_{0.1}^{0.3} = -8.315 \left[\ln \frac{0.3}{0.1} \right] J / K = -9.13 J / K$$

The whole enthalpy and entropy changes are the sum the changes in these three steps: $\Delta H = 40.6 + 3.5 + 0 = 44.1 \text{ kJ}$ $\Delta S = 108.8 + 8.2 - 9.1 = 107.9 \text{ J/K}$

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

or substituting 2.58 into 2.119:
$$dA = -SdT - PdV$$
 2.120

2.9.2 Gibbs Free Energy

2.9.2.1 Derivation

which reduces to:

The Gibbs Free Energy is perhaps misnamed. By analogy to the Helmholtz Free Energy, it should be called the Free Enthalpy (but enthalpy is an energy), because it is derived as follows:

$$\overline{G \equiv H - TS}$$

$$dG = d(H - TS) = dH - d(TS)$$
2.121
2.122

and

$$dG = TdS + VdP - d(TS) = TdS + VdP - SdT - TdS$$

or

$$\overline{dG = VdP - SdT}$$
 2.123

Notice the similarity to the Helmholtz Free Energy; in that case we subtracted the TS term from the internal energy; in this case we subtracted the TS term from the enthalpy. *The Gibbs Free Energy is the energy available for non-PV work (such as chemical work)*. It has two other important properties: its independent variables are T and P, generally the ones in which we are most interested in geochemistry, and it contains the entropy term (as does the Helmholtz free energy), and hence can be used as an indication of the direction in which spontaneous reactions will occur.

2.9.2.2 Gibbs Free Energy Change in Reactions

For a finite change at constant temperature, the Gibbs Free Energy change is:

$$\Delta G = \Delta H - T \Delta S$$
 2.124

The free energy change of formation, ΔG_{f} , is related to the enthalpy and entropy change of reaction:

$$\Delta G_f^o = \Delta H_f^o - T \Delta S_f^o \qquad 2.125$$

Like other properties of state, the Gibbs free energy is additive. Therefore:

$$\Delta G_r = \sum_i v_i \Delta G_{f,i}$$
 2.126

In other words, we can use Hess's Law to calculate the free energy change of reaction. Values for ΔG_f at the standard state are available in compilations.

2.9.3 Criteria for Equilibrium and Spontaneity

The Gibbs Free Energy is perhaps the single most important thermodynamic variable in geochemistry because it provides this criterion for recognizing equilibrium. This criterion is:

Products and reactants are in equilibrium when their Gibbs free energies are equal.

Another important quality of the Gibbs Free Energy is closely related:

At fixed temperature and pressure, a chemical reaction will proceed in the direction of lower Gibbs free energy (i.e., $\Delta G_r < 0$).

The reverse is also true: a reaction will not proceed if it produces an increase in the Gibbs Free Energy.

On an intuitive level, we can understand the Gibbs Free Energy as follows. We know that transformations tend to go in the direction of the lowest energy state (e.g., a ball rolls down hill). We also have learned that transformations go in the direction of increased entropy (if you drop a glass it breaks into pieces; if you drop the pieces they don't re-assemble into a glass). We must consider both the tendency for energy to decrease and the tendency for entropy to increase in order to predict the direction of a chemical reaction. This is what the Gibbs Free Energy does.

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

Example 2.7: Using Gibbs Free Energy to Predict Equilibrium

Using the thermodynamic data given in Table 2.2 calculate ΔG_r for the reaction:

 $CaAl_2Si_2O_8 + 2Mg_2SiO_4 \rightleftharpoons CaMgSi_2O_6 + MgAl_2O_4 + 2MgSiO_3$

(Anorthite + 2 Forsterite \rightleftharpoons Diopside + Spinel + 2 Enstatite)

at 298K and 1 MPa. Which mineral assemblage is more stable under these conditions (i.e., which side of the reaction is favored)? Which assemblage will be favored by increasing pressure? Why? Which side will be favored by increasing temperature? Why?

Answer: Since the data in Table 2.2 are listed in terms of ΔH_f and ΔS_f , it is most convenient to use equation 2.124, to calculate our answer:

$$\Delta G = \Delta H - T \Delta S$$

 ΔH is calculated as: $\Delta H_{f,Di} + \Delta H_{f,Sp} + 2 \times \Delta H_{f,En} - (\Delta H_{f,An} + 2 \times \Delta H_{f,Fo})$. ΔS is calculated in a similar manner. Our result is -6.08 kJ/mol. Because ΔG_r is negative, the reaction will proceed to the right, so that the assemblage on the right is more stable under the conditions of 298 K and 1 atm.

To find out which side will be favored by increasing pressure and temperature, we use equations 2.127 and 2.128 to see how ΔG will change. For temperature, $\partial \partial G/\partial T = -\Delta S$. ΔS for the reaction is -36.37 J/K-mol, and $\partial \partial G/\partial T = 36.37$. The result is positive, so that ΔG will increase with increasing T. Hence the left side will be favored by increasing temperature. Had we carried out the calculation at 1000° C and 0.1 MPa, a temperature appropriate for crystallization from magma, we would have found that the anorthite-forsterite assemblage is stable. For pressure, $\partial \partial G/\partial P = \Delta V$. ΔV for the reaction is -20.01 cc/mol (=J/MPa-mol). Thus ΔG will decrease with increasing pressure, so that the right side of the reaction will be favored.

Reassuringly, our thermodynamic result is consistent with geologic observation. The assemblage on the left, which could be called 'plagioclase peridotite' transforms to the assemblage on the right, 'spinel peridotite' as pressure increases in the mantle.

2.9.4 TEMPERATURE AND PRESSURE DEPENDENCE OF THE GIBbs FREE ENERGY

One reason the Gibbs Free Energy is particularly useful is that its characteristic variables are temperature and pressure, which are the 'external' variables of greatest interest in geochemistry. Since the Gibbs Free Energy is a state variable, we can deduce its temperature and pressure dependencies from equation 2.123, which are:

$$\left(\frac{\partial\Delta G}{\partial P}\right)_T = \Delta V$$
 2.127 $\left(\frac{\partial\Delta G}{\partial T}\right)_P = -\Delta S$ 2.128

Equations 2.127 and 2.128 allow us to predict how the Gibbs Free Energy of reaction will change with changing temperature and pressure. *Thus we can predict how the direction of a reaction will change if we change temperature and pressure.* To obtain the Gibbs Free Energy of Reaction at some temperature T' and pressure P', we integrate:

$$\Delta G_{T',P'} = \Delta G_{T_{ref},P_{ref}} + \int_{P_{ref}}^{P'} \Delta V_r dP - \int_{T_{ref}}^{T'} \Delta S_r dT \qquad 2.129$$

where ΔG° is the 'standard state' Gibbs Free Energy change. For solids we can often ignore the effects of temperature and pressure on ΔV so the first integral reduces to: $\Delta V(P-P_{ref})$. For liquids, and particularly for gases, effects of pressure and temperature of the volume change are important and cannot be ignored. The reference pressure is generally 0.1 MPa. On the other hand, we usually cannot ignore the temperature dependence of entropy. Hence we need to express ΔS_{rr} as a function of temperature. The temperature dependence of entropy is given by equation 2.105. Writing this in integral form, we have:

$$\Delta S(T) = \Delta S_{T_{ref}} + \int_{T_{ref}}^{T} \frac{\Delta C_p}{T} dT$$
 2.130

September 16, 2005

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

Substituting this into 2.129, the second integral becomes:

$$\Delta G_{T'} = -\Delta S_{ref}(T' - T_{ref}) - \int_{T_{ref}}^{T'} \int_{T_{ref}}^{T} \frac{\Delta C_p}{T} dT dT \qquad 2.131$$

Example 2.8. Predicting the Equilibrium Pressure of a Mineral Assemblage

Using the thermodynamic reaction and data as in Example 2.7:

 $CaAl_2Si_2O_8 + 2Mg_2SiO_4 \rightleftharpoons CaMgSi_2O_6 + MgAl_2O_4 + 2MgSiO_3$

(Anorthite + Forsterite \rightleftharpoons Diopside + Spinel + 2 Enstatite)

determine the pressure at which these two assemblages will be in equilibrium at 1000° C. Assume that the volume change of the reaction is independent of pressure and temperature (i.e., α and $\beta = 0$).

Answer: These two assemblages will be in equilibrium if and only if the Gibbs Free Energy of reaction is 0. Mathematically, our problem is to solve equation 2.129 for *P* such that $\Delta G_{1273,P} = 0$.

Our first step is to find ΔG_r for this reaction at 1000°C (1273 K) using equation 2.131. Heat capacity data in Table 2.2 is in the form: $C_p = a + bT - cT^2$. Substituting for ΔC_{pr} we have:

$$\Delta G_T = -\Delta S_{T_{ref}}(T' - T_{ref}) - \int_{T_{ref}}^{T'} \int_{T_{ref}}^{T} \left(\frac{\Delta a}{T} + \Delta b - \frac{\Delta c}{T^3}\right) dT dT \qquad 2.132$$

Performing the double integral and collecting terms, substituting ΔT for T' – T_{ref}, this simplifies to:

$$\Delta G_{T'} = -\Delta T \left[\Delta S_{T_{ref}} - \Delta a + \frac{\Delta b}{2} \Delta T - \frac{\Delta c \Delta T}{2T' T_{ref}^2} \right] - \Delta a T' \ln \frac{T}{T_{ref}}$$
2.133

Equation 2.133 is a general solution to equation 2.131 when the Maier-Kelley heat capacity is used.

We found ΔS_{Tref} to be -36.37 J/K-mol in Example 2.7. Computing Δa as $(a_{Di} + a_{Sp} + 2a_{En}) - (a_{An} + 2a_{Fo})$, we find $\Delta a = 15.96$ J/mol. Computing Δb and Δc similarly, they are -0.01732 J/K-mol and 1.66×10^6 J-K²/mol respectively. Substituting values into equation 2.135, we find $\Delta G_T = 36.74$ kJ/mol.

Since we may assume the phases are incompressible, the solution to the pressure integral is:

$$\Delta G_P = \int_{P_{ref}}^{P} \Delta V_r dP = \Delta V_r (P' - P_{ref})$$
 2.134

Equation 2.129 may now be written as:

$$\Delta G_{T',P'} = 0 = \Delta G^{\circ} + \Delta G_T + \Delta V_r(P' - P_{ref})$$

Let $\Delta G_{1273,0.1} = \Delta G^{\circ} + \Delta G_T$. ΔG° is -6.95 kJ/mol (calculated from values in Table 2.2), so $\Delta G_{1273,0.1} = 29.86$ kJ/mol. $\Delta G_{1273,0.1}$ is positive, meaning that the left side of the reaction is favored at 1000° C and atmospheric pressure, consistent with our prediction based on $\partial G/\partial T$.

Solving for pressure, we have

$$P' = \frac{-\Delta G_{T'P_{ref}}}{\Delta V_r} + P_{ref}$$
 2.135

With $\Delta V = -20.01$ cc/mol, we obtain a value of 1.49 GPa (14.9 kbar). Thus assemblages on the right and left will be in equilibrium at 1.49 GPa and 1000°C. Below that pressure, the left is stable, above that pressure, the right side is the stable assemblage, according to our calculation.

The transformation from 'plagioclase peridotite' to 'spinel peridotite' actually occurs around 1.0 GPa in the mantle. The difference between our result and the real world primarily reflects differences in mineral composition: mantle forsterite, enstatite and diopside are solids solutions containing Fe and other elements. The difference does not reflect our assumption that the volume change is independent of pressure. When available data for pressure and temperature dependence of the volume change is included in the solution, the pressure obtained is only marginally different: 1.54 GPa.

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

Example 2.9: Volume and Free Energy Changes for finite compressibility

The compressibility (β) of forsterite (Mg₂SiO₄) is 8.33 x 10⁻⁶ MPa⁻¹. Using this and the data given in Table 2.2, what is the change in molar volume and Gibbs Free Energy of forsterite at 100 MPa and 298K? *Answer:* Let's deal with volume first. We want to know how the molar volume (43.79 cc/mol) changes as the pressure increases from the reference value (0.1 MPa) to 1GPa. The compressibility is defined as:

$$\beta = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T \tag{2.12}$$

So the change in volume for an incremental increase in pressure is given by:

$$dV = -V\beta dP \qquad 2.136$$

To find the change in volume over a finite pressure interval, we rearrange and integrate:

$$\int_{V^o}^{V} \frac{dV}{V} = -\int_{P^o}^{P} \beta dP$$

Performing the integral, we have:

$$\ln\frac{V}{V^o} = -\beta(P - P^o)$$
 2.137

where P is the pressure interval, P'-P°. This may be rewritten as:

$$V = V^{o} e^{-\beta(P - P^{o})}$$
 2.138

However, te value of β is of the order of 10^{-2} , and in this case, the approximation $e^x \approx x + 1$ holds, so that 2.138 may be written as:

$$V \cong V^{\circ}(1 - \beta(P - P^{\circ}))$$
 2.139

Equation 2.139 is a general expression that expresses volume as a function of pressure when β is known and is independent of temperature and pressure. Furthermore, in situations where $P >> P^{\circ}$, this can be simplified to:

$$V \cong V^o(1 - \beta P) \tag{2.140}$$

Using equation 2.140, we calculate a molar volume of 43.54 cc/mol (identical to the value obtained using 2.138). The volume change, ΔV is 0.04 cc/mol.

The change in Free Energy with volume is given by:

$$\left(\frac{\partial G}{\partial P}\right)_T = V$$

so that the free energy change as a consequence of a finite change is pressure can be obtained by integrating:

$$\Delta G = \int_{P^o}^P V dP$$

Into this we may substitute equation 2.140:

$$\Delta G = \int_{V^{o}}^{V} V^{o} (1 - \beta P) dP = V^{o} \left[P - \beta P^{2} \right]_{P^{o}}^{P}$$
2.141

Using 2.141 we calculate a value of ΔG of 4.37 kJ/ mol.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

2.10 The Maxwell Relations

The reciprocity relationship, which we discussed earlier, leads to a number of useful relationships. These relationships are known as the Maxwell Relations^{*}. Consider the equation:

$$dU = TdS - PdV \tag{2.58}$$

If we write the partial differential of U in terms of S and V we have:

$$dU = \left(\frac{\partial U}{\partial S}\right)_V dS + \left(\frac{\partial U}{\partial V}\right)_S dV \qquad 2.142$$

From a comparison of these two equations, we see that:

$$\left(\frac{\partial U}{\partial S}\right)_{V} = T$$
 2.137 and $\left(\frac{\partial U}{\partial V}\right)_{S} = -P$ 2.143

And since the cross differentials are equal, it follows that:

$$\left(\frac{\partial T}{\partial V}\right)_{S} = \left(\frac{\partial P}{\partial S}\right)_{V}$$
 2.144

The other Maxwell Relations can be derived in an exactly analogous way from other state functions. They are:

from *dH* (Equ. 2.65) $\left(\frac{\partial T}{\partial P}\right)_{S} = \left(\frac{\partial V}{\partial S}\right)_{P}$ 2.145

from *dA* (Equ. 2.120)
$$\left(\frac{\partial P}{\partial T}\right)_{V} = \left(\frac{\partial S}{\partial V}\right)_{T}$$
 2.146

from dG (Equ. 2.123)
$$\left(\frac{\partial V}{\partial T}\right)_{P} = -\left(\frac{\partial S}{\partial P}\right)_{T}$$
 2.147

References and Suggestions for Further Reading

Anderson, G. M. and D. A. Crerar. 1993, *Thermodynamics in Geochemistry*, New York: Oxford Univ. Press.

Broecker, W. S. and V. M. Oversby, 1971. Chemical Equilibria in the Earth. New York: McGraw-Hill.

Castellan, G. W. 1964. Physcial Chemistry. Reading: Addison-Wesley.

Chandler, D. 1987. Introduction to Modern Statistical Mechanics. New York: Oxford.

Denbigh, K. G., 1981. The Principles of Chemical Equilibrium, Cambridge: Cambridge Univ. Press.

Feynman, R., R. B. Leighton, and M. L. Sands, 1989. *The Feynman Lectures on Physics, Vol. I.* Pasadena: Calif. Inst. Tech.

Fletcher, P. 1993. Chemical Thermodynamics for Earth Scientists, Essex: Logman Scientific and Technical.

Haas, J. L. and J. R. Fisher, 1976. Simultaneous evaluation and correlation of thermodynamic data, *Am. J. Sci.* 276: 525-545.

Helgeson, H. C., J. M. Delany, H. W. Nesbitt, and D. K. Bird. 1978. Summary and critique of the thermodynamic properties of rock-forming minerals. *Am. J. Sci.* 278A: 1-229.

Nordstrom, D. K. and J. L. Munoz, 1986. Geochemical Thermodynamics, Palo Alto: Blackwell Scientific.

Reif, F. 1965. Fundamentals of Statistical and Thermal Physics. New York: McGraw-Hill.

Swalin, R. A. 1962. Thermodynamics of Solids. New York: John Wiley and Sons.

^{*} The Maxwell Relations are named for Scottish physicist James Clerk Maxwell (1831-1879), perhaps the most important figure in 19th century physics. He is best known for his work on electromagnetic radiation, but he also made very important contributions to statistical mechanics and thermodynamics.

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

Problems

1. For a pure olivine mantle, calculate the adiabatic temperature gradient $(\partial T/\partial P)_S$ at 1 atm and 1000°C. Use the thermodynamic data in Table 2.2 for forsterite (Mg-olivine, Mg₂SiO₄) and $\alpha = 44 \times 10^{-6} \text{ K}^{-1}$, and $\beta = 0.8 \times 10^{-6}$.

The following conversions may be useful: 1 cal = 4.184 joules, 1 liter-atm = 24.21 cal.

2. For an ideal gas, show that:

$$\frac{\partial \alpha V}{\partial P} = -\frac{\partial \beta V}{\partial T}$$

3. A quartz crystal has a volume of 7.5 ml at 298 K and 0.1 MPa. What is the volume of the crystal at 840K and 12.3 MPa if

a.) α = 1.4654 × 10⁻⁵ K⁻¹ and β = 2.276 × 10⁻¹¹ Pa⁻¹ and α and β are independent of T and P.

b.) $\alpha = 1.4310 \times 10^{-5} \text{ K}^{-1} + 1.1587 \times 10^{-9} \text{ K}^{-2}\text{T}$

 $\beta = 1.8553 \times 10^{-11} Pa^{-1} + 7.9453 \times 10^{-8} P^{-1}$

4. One mole of an ideal gas is allowed to expand against a piston at constant temperature of 0°C. The initial pressure is 1 MPa and the final pressure is 0.04 MPa. Assuming the reaction is reversible,

a. What is the work done by the gas during the expansion?

b. What is the change in the internal energy and enthalpy of the gas?

c. How much heat is gained/lost during the expansion?

5. Lava lakes generally have a temperature of about 1100°C. Assuming that basaltic magma travels from its place of origin quickly enough so that negligible heat is lost to wall rocks, calculate the temperature of the magma at a depth of 40 km. The density of basaltic magma at 1200°C is 2.61 g/cc; coefficient of thermal expansion is about 1×10^4 /K. Assume a heat capacity of 0.2 cal/g-deg and that pressure is related to depth as 1km = 0.33kb (kb is kilobars; assume 1bar = 1 atm.). (*HINT: "negligible heat loss" means the system may be treated as adiabatic.*)

6 Show that the C_p of an ideal monatomic gas is 5/2 R.

7. Show that:

 $\left(\frac{\partial P}{\partial T}\right)_{V} = \frac{\alpha}{\beta}$

8. Show that for a reversible process:

 $\left(\frac{\partial U}{\partial V}\right)_{T} = T\frac{\alpha}{\beta} - P \qquad (\text{equ. 2.73})$

(*Hint: begin with the statement of the first law (equation 2.58), make use of the Maxwell relations, and your proof in problem 7.*)

9. Imagine that there are 30 units of energy to distribute among 3 copper blocks.

a.) If the energy is distributed completely randomly, what is the probability of the first block having all the energy?

b.) If n_1 is the number of units of energy of the first block, construct a graph (a histogram) showing the probability of a given value of n_1 occurring as a function of n_1 . (*HINT: use equation 2.37, but modify it for the case where there are 3 blocks*).

9. Consider a box partitioned into equal volumes, with the left half containing 1 mole of Ne and the right half containing 1 mole of He. When the partition is removed, the gases mix. Show, using a classi-

Geochemistry

CHAPTER Z: FUNDAMENTAL CONCEPTS OF THERMODYNAMICS

cal thermodynamic approach (i.e., macroscopic) that the entropy change of this process is $\Delta S = 2R \ln 2$. Assume that He and Ne are ideal gases and that temperature is constant.

- 10. Find expressions for C_p and C_v for a van der Waals gas.
- 11. Show that β (the compressibility, defined in equation 2.12) of an ideal gas is equal to 1/P.

12. Show that
$$S = \frac{\overline{U}}{T} + R \ln Q$$

Hint: Start with equations 2.47 and 2.36a using the approximation that $\ln N! = N \ln N - N$.

13. Show that $\Delta H = \int_{P_1}^{P_2} V(1 - \alpha T) dP$

Hint: Begin with equation 2.63 and express dU as a function of temperature and volume change.

14. Helium at 298K and 1 atm has $S^{\circ} = 30.13 \text{ cal/K-mole}$. Assume He is an ideal gas.

- a. Calculate V, H, G, α , β , Cp, Cv, for He at 298K and 1 atm.
- b. What are the values for these functions at 600K and 100 atm?
- c. What is the entropy at 600 K and 100 atm?

15. Given the following standard data:

	$Mg + 3/2 O_2 + Si \rightarrow MgSiO_3$	$\Delta H = -1497.4 \text{ kJ/mol}$
	$Si + O_2 \rightarrow SiO_2$	$\Delta H = -859.4 \text{ kJ/mol}$
	$2Mg + 2O_2 + Si \rightarrow Mg_2SiO_4$	$\Delta H = -2042.6 \text{ kJ/mol}$
find ΔH in Joules for:	$Mg_2SiO_4 + SiO_2 \rightarrow 2MgSiO_3$	

16. Using the data in Table 2.2, calculate the enthalpy and entropy change of diopside as it is heated at constant pressure from 600 K to 1000 K.

17. Calculate the total enthalpy upon heating of 100g of quartz from 25° C to 900° C. Quartz undergoes a phase transition from α -quartz to β -quartz at 575° C. The enthalpy of this phase transition is $\Delta H_{tr} = 0.411 \text{ kJ/mol}$. Use the Maier-Kelly heat capacity data in Table 2.2.

18. Calcite and aragonite are two forms of $CaCO_3$ that differ only their crystal lattice structure. The reaction between them is thus simply:

Using the data in Table 2.2,

- a.) Determine which of these forms is stable at the surface of the earth (25° C and 0.1 MPa).
- b.) Which form is favored by increasing temperature?
- c.) Which form is favored by increasing pressure?

19. Use the data in Table 2.2 to determine the pressure at which calcite and aragonite are in equilibrium at 300°C.

20. Suppose you found kyanite and andalusite coexisting in the same rock, that you had reason to believe this was an equilibrium assemblage, and that you could independently determine the temperature of equilibrium to be 400°C. Use the data in Table 2.2 to determine the pressure at which this rock equilibrated. CHAPTER 7: SOLUTIONS

CHAPTER 3: Solutions and Thermodynamics of Multicomponent Systems

3.1 INTRODUCTION

In the previous chapter, we introduced thermodynamic tools that allow us to predict the equilibrium mineral assemblage under a given set of conditions. For example, having specified temperature, we were able to determine the pressure at which the assemblage anorthite+forsterite is in equilibrium with the assemblage diopside+spinel+enstatite. In that reaction the minerals had unique and invariant compositions. In the Earth, things are not quite so simple: these minerals are present as solid solutions^{*}, with substitutions of Fe for Mg, Na for Ca, and Cr and Fe³⁺ for Al, among others. Indeed, most natural substances are solutions; that is, their compositions vary. Water, which is certainly the most interesting substance at the surface of the Earth and perhaps the most important, inevitably has a variety of substances dissolved in it. These dissolved substances are themselves often of primary geochemical interest. More to the point, they affect the chemical behavior of water. For example, the freezing temperature of an aqueous NaCl solution is lower than that of pure water. You may have taken advantage of this phenomenon by spreading salt to de-ice sidewalks and roads, or adding salt to ice to make ice cream.

In a similar way, the equilibrium temperature and pressure of the plagioclase+olivine \rightleftharpoons clinopyroxene+spinel+orthopyroxene reaction depends on the composition of these minerals. To deal with this compositional dependence, we need to develop some additional thermodynamic tools, which is the objective of this chapter. This may seem burdensome at first: if it were not for the variable composition of substances, we would already know most of the thermodynamics we need. However, as we will see in Chapter 4, we can use this compositional dependence to advantage in reconstructing conditions under which a mineral assemblage or a hydrothermal fluid formed.

A final "difficulty" is that the valance state of many elements may vary. Iron, for example, may change from its Fe^{2+} state to Fe^{3+} when an igneous rock weathers. The two forms of iron have very different chemical properties; for example Fe^{2+} is considerably more soluble in water than is Fe^{3+} . Another example of this kind of reaction is photosynthesis, the process by which CO_2 is converted to organic carbon. These kinds of reactions are called "oxidation–reduction", or "redox" reactions. The energy your brain uses to process the information you are now reading comes from oxidation of organic carbon — carbon originally reduced by photosynthesis in plants. To fully specify the state of a system, we must specify its "redox" state. We treat redox reactions in the final section of this chapter.

Though Chapter 4 will add a few more tools to our geochemical toolbox, and treat a number of advanced topics in thermodynamics, it is designed to be optional. With completion of this chapter, you will have a sufficient thermodynamic background to deal with a wide range of phenomena in the Earth, and most of the topics in the remainder of this book.

3.2 Phase Equilibria

3.2.1 Some Definitions

3.2.1.1 Phase

Phases are real substances that are homogeneous, physically distinct, and (in principle) mechanically separable. For example, the phases in a rock are the minerals present. Amorphous substances are also phases, so glass or opal would be phases. The sugar that won't dissolve in your ice tea is a distinct phase from the tea, but the dissolved sugar is not. *Phase* is not synonymous with *compound*. Phases

^{*}The naturally occurring minerals of varying composition are referred to as plagioclase rather than anorthite, olivine rather than forsterite, clinopyroxene rather than diopside, and orthopyroxene rather than enstatite.

Geochemistry

CHAPTER 7: SOLUTIONS

need not be chemically distinct: a glass of ice water has two distinct phases: water and ice. Many solid compounds can exist as more than one phase. Nor need they be compositionally unique: plagioclase, clinopyroxene, olivine, etc., are all phases even though their composition can vary. Thus a fossil in which the aragonite (CaCO₃) is partially retrograded into calcite (also CaCO₃) consists of 2 phases. Systems, and reactions occurring within them, consisting of a single phase are referred to as *homogenous*; those systems consisting of multiple phases, and the reactions occurring within them, are referred to as *heterogeneous*.

3.2.1.2 Species

Species is somewhat more difficult to define than either *phase* or *component*. A species is a chemical entity, generally an element or compound (which may or may not be ionized). The term is most useful in the context of gases and liquids. A single liquid phase, such as an aqueous solution, may contain a number of species. For example, H_2O , H_2CO_3 , HCO_3^- , CO_3^{2+} , H^+ , and OH^- are all species commonly present in natural waters. The term species is generally reserved for an entity that actually exists, such as a molecule, ion, or solid on a microscopic scale. This is not necessarily the case with components, as we shall see. The term species is less useful for solids, although it is sometimes applied to the pure end-members of solid solutions and to pure minerals.

3.2.1.3 Component

In contrast to a species, a *component* need not be a real chemical entity, rather it is simply an algebraic term in a chemical reaction. The *minimum number of components*^{*} of a system is rigidly defined as *the minimum number of independently variable entities necessary to describe the composition of each and every phase of a system*. Unlike species and phases, components may be defined in any convenient manner: what the components of your system are and how many there are depend on your interest and on the level of complexity you will be dealing with. Consider our aragonite-calcite fossil. If the only reaction occurring in our system (the fossil) is the transformation of aragonite to calcite, one component, CaCO₃, is adequate to describe the composition of both phases. If, however, we are also interested in the precipitation of calcium carbonate from water, we might have to consider CaCO₃ as consisting of 2 components: Ca^{2+} and CO_3^2 .

There is a rule to determine the minimum number of components in a system once you decide what your interest in the system is; the hard part is often determining your interest. The rule is:

$$\mathbf{c} = \mathbf{n} - \mathbf{r} \tag{3.1}$$

where *n* is the number of species, and *r* is the number of independent chemical reactions possible between these species. Essentially, this equation simply states that if a chemical species can be expressed as the algebraic sum of other components, we need not include that species among out minimum set of components. Let's try the rule on the species we listed above for water. We have 6 species: H_2O , H_2CO_3 , HCO_3^- , CO_3^{2+} , H^+ , and OH^- . We can write 3 reactions relating them:

$$HCO_{3} = H^{+} + CO_{3}^{2-}$$

 $H_{2}CO_{3} = H^{+} + HCO_{3}^{-}$
 $H_{2}O = H^{+} + OH^{-}$

Equation 3.1 tells us we need 3 = 6 - 3 components to describe this system: CO_3^2 +, H⁺, and OH⁻. Put another way, we see that carbonic acid, bicarbonate, and water can all be expressed as algebraic sums the hydrogen, hydroxyl, and carbonate ions, so they need not be among our minimum set of components.

In igneous and metamorphic petrology, components are often the major oxides (though we may often chose to consider only a subset of these). On the other hand, if we were concerned with the isotopic

^cCaution: some books use the term *number of components* as synonymous with *minimum number of components*.

Geochemistry

CHAPTER 7: SOLUTIONS



equilibration of minerals with a hydrothermal fluid, ¹⁸O would be considered as a different component than ¹⁶O.

Perhaps the most straightforward way of determining the number of components is a graphical approach. If all phases can be represented on a one-dimensional diagram (that is, a straight line representing composition), we are dealing with a two component system. For example, consider the hydration of Al_2O_3 (corundum) to form boehmite (AlO(OH)) or gibbsite $Al(OH)_3$. Such a system would contain 4 phases (corundum, boehmite, gibbsite, water), but is nevertheless a two component system because all phases may be represented in one-dimension of composition space, as is shown in Figure 3.1. Because there are two polymorphs of gibbsite, one of boehmite, and two other possible phases of water, there are 9 phases possible phases in this two-component system. Clearly, a system may have many more phases than components.

Similarly, if a system may be represented in 2 dimensions, it is a three-component system. Figure 3.2



Figure 3.2. Phase diagram for the system Al₂O₃–H₂O–SiO₂. The lines are called *joins* because they join phases. In addition to the end-members, or components, phases represented are *g*: gibbsite, *by*: bayerite, *n*: norstrandite (all polymorphs of Al(OH)₃), *d*: diaspore, *bo*: boehmite (polymorphs of AlO(OH)), *a*: andalusite, *k*: kyanite, *s*: sillimanite (all polymorphs of Al₂SiO₅), *ka*: kaolinite, *ha*: halloysite, *di*: dickite, *na*: nacrite (all polymorphs of Al₂Si₂O₅(OH)₄), and *p*: pyrophyllite (Al₂Si₄O₁₀(OH)₂). There are also 6 polymorphs of quartz (coesite, stishovite, tridymite, cristobalite, *α*-quartz, and β-quartz).

is a three-component system. Figure 3.2 is a ternary diagram illustrating the system Al_2O_3 – H_2O –SiO₂. The graphical representation approach reaches it practical limit in a four component system because of the difficulty of representing more than 3 dimensions on paper. A four component system is a quaternary one, and can be represented with a three-dimensional quaternary diagram.

It is important to understand that a component may or may not have chemical reality. For example in the exchange reaction:

NaAlSi₃O₈ + K⁺ = KAlSi₃O₈ + Na⁺ we could alternatively define the *exchange operator* KNa₋₁ (where Na₋₁ is -1 mol of Na ion) and write the equation as:

 $NaAlSi_{3}O_{8} + KNa_{-1} = KAlSi_{3}O_{8}$

In addition, we can also write the reaction:

$K - Na = KNa_{-1}$

Here we have 4 species and 2 reactions and thus a minimum of only 2 components. You can see that *a component is merely an algebraic term.*

There is generally some freedom in choosing components. For example, in the ternary (i.e., 3 component) system $SiO_2 - Mg_2SiO_4 - MgCaSi_2O_6$, we could choose our components to be quartz, diopside, and forsterite, or we could choose them to be SiO_2 , MgO, and CaO. Either way, we are dealing with a ternary

Geochemistry

CHAPTER 3: SOLUTIONS

system (which contains MgSiO₃ as well as the three other phases).

3.2.1.4 Degrees of Freedom

The number of degrees of freedom in a system is equal to the sum of the number of independent *intensive* variables (generally T & P) and independent concentrations (or activities or chemical potentials) of components in phases that must be fixed to define uniquely the state of the system. A system that has no degrees of freedom (i.e., is uniquely fixed) is said to be invariant, one that has one degree of freedom is univariant, etc. Thus in an univariant system, for example, we need specify only the value of one variable, for example, temperature or the concentration of one component in one phase, and the value of pressure and all other concentrations are then fixed, i.e., they can be calculated (assuming the system is at equilibrium).

3.2.2 The Gibbs Phase Rule

The Gibbs[‡] Phase Rule is a rule for determining the *degrees of freedom*, or *variance*, of a system *at equilibrium*. The rule is:

$$f = \mathbf{c} - \phi + \mathbf{2} \tag{3.2}$$

where *f* is the degrees of freedom, *c* is the number of components, and ϕ is the number of phases. The mathematical analogy is that the degrees of freedom are equal to the number of variables less the number of equations relating those variables. For example, in a system consisting of just H₂O, if two phases of coexist, for example, water and steam, then the system in univariant. Three phases coexist at the triple point of water, so the system is said to be invariant, and T and P are uniquely fixed: there is only one temperature and one pressure at which the three phases of water can coexist (273.15 K and 0.006 bar). If only one phase is present, for example just liquid water, then we need to specify 2 variables to describe completely the system. It doesn't matter which two we pick. We could specify molar volume and temperature and from that we could deduce pressure. Alternatively, we could specify pressure are fixed. It is important to remember this applies to intensive parameters. To know volume, an extensive parameter, we would have to fix one additional extensive variable (such as mass or number of moles). And again, we emphasize that all this applies only to systems at equilibrium.

Now consider the hydration of corundum to form gibbsite. There are 3 phases, but there need be only two components. If these 3 phases (water, corundum, gibbsite) are at equilibrium, we have only 1 degree of freedom, i.e., if we know the temperature at which these 3 phases are in equilibrium, the pressure is also fixed.

Rearranging equation 3.2, we also can determine the *maximum* number of phases that can coexist at equilibrium in any system. The degrees of freedom cannot be less than zero, so for an invariant, one component system, a maximum of three phases can coexist at equilibrium. In a univariant one-component system, only 2 phases can coexist. Thus sillimanite and kyanite can coexist over a range of temperatures, as can kyanite and andalusite. But the three phases of Al_2SiO_5 coexist only at one unique temperature and pressure.

Let's consider the example of the three-component system $Al_2O_3-H_2O-SiO_2$ in Figure 3.2. Although many phases are possible in this system, for any given composition of the system only three phases can coexist at equilibrium over a range of temperature and pressure. Four phases, e.g., a, k, s and q, can coexist only along a one-dimensional line or curve in P-T space. Such points are called univariant lines (or curves). Five phases can coexist at invariant points at which both temperature and pressure are uniquely fixed. Turning this around, if we found a metamorphic rock whose composition fell within

[‡] J. Williard Gibbs (1839-1903) is viewed by many as the father of thermodynamics. He received the first doctorate in engineering granted in the U. S., from Yale in 1858. He was Professor of Mathematical Physics at Yale from 1871 until his death. He also helped to found statistical mechanics. The importance of his work was not widely recognized by his American colleagues, though it was in Europe, until well after his death.

Geochemistry

CHAPTER 3: SOLUTIONS

the $Al_2O_3-H_2O-SiO_2$ system, and if the rock contained 5 phases, it would be possible to determine uniquely the temperature and pressure at which the rock equilibrated.

3.2.3 The Clapeyron Equation

A common problem in geochemistry is to know how a phase boundary varies in P-T space, e.g., how a melting temperature will vary with pressure. At a phase boundary, two phases must be in equilibrium, i.e., ΔG must be 0 for the reaction Phase 1 \rightleftharpoons Phase 2. The phase boundary therefore describes the condition:

$$d(\Delta G_r) = \Delta V_r dP - \Delta S_r dT = 0.$$

Thus the slope of a phase boundary on a temperature-pressure diagram is:

$$\frac{dT}{dP} = \frac{\Delta V_r}{\Delta S_r}$$
3.3

where ΔV_r and ΔS_r are the volume and entropy changes associated with the reaction. Equation 3.3 is known as the *Clausius-Clapeyron Equation*, or simply the *Clapeyron Equation*. Because ΔV_r and ΔS_r are functions of temperature and pressure, this, of course, is only an instantaneous slope. For many reactions, however, particularly those involving only solids, the temperature and pressure dependencies of ΔV_r and ΔS_r will be small and the Clapeyron slope will be relatively constant over a large T and P range.

Because $\Delta S = \Delta H/T$, the Clapeyron equation may be equivalently written as:

$$\frac{dT}{dP} = \frac{T\Delta V_r}{\Delta H_r}$$
 3.4

Slopes of phase boundaries in P-T space are generally positive, implying that the phases with the largest volumes also generally have the largest entropies (for reasons that become clear from a statistical mechanical treatment). This is particularly true of solid-liquid phase boundaries, although there is one very important exception: water. How do we determine the pressure and temperature dependence of ΔV_r and why is ΔV_r relatively T and P independent in solids?

We should emphasize that application of the Clapeyron equation is not limited to reactions between two phases in a one-component system, but may be applied to any univariant reaction.

3.3 Solutions

Solutions are defined as homogeneous phases produced by dissolving one or more substances in another substance. In geochemistry we are often confronted by solutions: as gases, liquids, and solids. Free energy depends not only on T and P, but also on composition. In thermodynamics it is generally most convenient to express compositions in terms of mole fractions, X_i , the number of moles of *i* divided by the total moles in the substance (moles are weight divided by atomic or molecular weight). The sum of all the X_i must, of course, total to 1.

Solutions are distinct from purely mechanical mixtures. For example, salad dressing (oil and vinegar) is not a solution. Similarly, we can grind anorthite $(CaAl_2Si_2O_8)$ and albite $(NaAlSi_3O_8)$ crystals into a fine powder and mix them, but the result is not a plagioclase solid solution. The Gibbs Free Energy of mechanical mixtures is simply the sum of the free energy of the components. If, however, we heated the anorthite-albite mixture to a sufficiently high temperature that the kinetic barriers were overcome, there would be a reordering of atoms and the creation of a true solution. Because this reordering is a spontaneous chemical reaction, there must be a decrease in the Gibbs Free Energy associated with it. This solution would be stable at 1 atm and 25°C. Thus we can conclude that the solution has a lower Gibbs Free Energy than the mechanical mixture. On the other hand, vinegar will never dissolve in oil at 1 atm and 25°C because the Gibbs Free Energy of that solution is greater than that of the mechanical mixture.

CHAPTER 7: SOLUTIONS

Example 3.1: The Graphite-Diamond Transition

At 25°C the graphite-diamond transition occurs at 1600 MPa (megapascals, 1 MPa =10 b). Using the standard state (298 K, 0.1 MPa) data below, predict the pressure at which the transformation occurs when temperature is 1000°C.

	Graphite	Diamond
α (K ⁻¹)	1.05×10^{-05}	7.50×10^{-06}
β (MPa ⁻¹)	3.08×10^{-05}	2.27 ×10 ⁻⁰⁶
S° (J/K-mol)	5.74	2.38
V (cm^3/mol)	5.2982	3.417

Answer: We can use the Clapeyron equation to determine the slope of the phase boundary. Then, assuming that ΔS and ΔV are independent of temperature, we can extraplate this slope to 1000° C to find the pressure of the phase transition at that temperature.

First, we calculate the volumes of graphite and diamond at 1600 MPa as (equ. 2.139):

$$V = V^{\circ}(1 - \beta \Delta \vec{P}) \qquad 3.5$$

where ΔP is the difference between the pressure of interest (1600 MPa in this case) and the reference pressure (0.1 MPa). Doing so, we find the molar volumes to be 5.037 for graphite and 3.405 for diamond, so ΔV_r is -1.6325 cc/mol. The next step will be to calculate ΔS at 1600 MPa. The pressure dependence of entropy is given by equation 2.143: $\partial S / \partial P _T = -\alpha V$. Thus to determine the effect of pressure we integrate:

$$S_{P} = S^{\circ} + \int_{P_{ref}}^{P_{1}} \left(\frac{\partial S}{\partial P} \right)_{T} dP = S^{\circ} + \int_{P_{ref}}^{P_{1}} -\alpha V dP \quad 3.6$$

(We use S_p to indicate the entropy at the pressure of interest and S° the entropy at the reference pressure.) We need to express V as a function of pressure, so we substitute 3.5 into 3.6:

$$S_{P} = S^{\circ} + \int_{P_{ref}}^{P_{1}} -\alpha V^{\circ} (1 - \beta P) dP$$

= $S^{\circ} - \alpha V^{\circ} \left[\Delta P - \frac{\beta}{2} (P_{1}^{2} - P_{ref}^{2}) \right]$ 3.7

The reference pressure, $P_{ref'}$ is negligible compared to P_1 (0.1 MPa vs 1600 PMa), so that this simplifies to:

 $S_{\rm P} = S^{\circ} - \alpha V^{\circ} \left[\Delta P - \frac{\beta}{2} P_1^2 \right]$

For graphite, S_p is 5.66 J/K-mol, for diamond, it is 2.34 J/K-mol, so ΔS_r at 1600 MPa is -3.32 J-K⁻¹-mol⁻¹.

The Clapeyron slope is therefore:

$$\frac{\Delta S}{\Delta V} = \frac{-3.322}{-1.63} = 2.035 \text{ JK}^{-1} \text{ cm}^{-3}$$

One distinct advantage of the SI units is that $cm^3 = J/MPa$, so the above units are equivalent to K/MPa. From this, the pressure of the phase change at 1000° C can be calculated as:

$$P_{1000} = P_{298} + \Delta T \times \frac{\Delta S}{\Delta V}$$

=1600 + 975 × 2.035 = 3584 MI

The Clapeyron slope we calculated (solid line) is compared with the experimentally determined phase boundary in Figure 3.3. Our calculated phase boundary is linear whereas the experimental one is not. The curved nature of the observed phase boundary indicates ΔV and ΔS are pressure and temperature dependent. This is indeed the case, particularly for graphite. A more accurate estimate of the volume change requires β be expressed as a function of pressure.



Figure 3.3. Comparison of the graphite-diamond phase boundary calculated from thermodynamic data and the Clapeyron slope (solid line) with the experimentally observed phase boundary (dashed line).

Geochemistry

CHAPTER 7: SOLUTIONS

3.3.1 RAOULT'S LAW

Working with solutions of ethylene bromide and propylene bromide, Raoult^{II} noticed that the vapor pressures of the components in a solution were proportional to the mole fractions of those components:

$$P_i = X_i P_i^o$$
 3.8

where P_i is the vapor pressure of component *i* above the solution, X_i is the mole fraction of *i* in solution, and P_i^o is the vapor pressure of pure *i* under standard conditions. Assuming the partial pressures are additive and the sum of all the partial pressures is equal to the total gas pressure ($\Sigma P_i = P_{total}$):

$$P_i = X_i P_{total}$$
3.9

Thus partial pressures are proportional to their mole fractions. *This is the definition of the partial pressure of the ith gas in a mixture.*

Raoult's Law holds only for *ideal solutions*, i.e., substances where there are no intermolecular forces. It also holds to a good approximation where the forces between like molecules are the same as between different molecules. The two components Raoult was working with were very similar chemically, so that this condition held and the solution was nearly ideal. As you might guess, not all solutions are ideal. Fig. 3.4 shows the variations of partial pressures above a mixture of water and dioxane. Significant deviations from Raoult's Law are the rule except where X_i approaches 1.



Figure 3.4. Vapor pressure of water and dioxane in a water-dioxane mixture showing deviations from ideal mixing. Shaded areas are areas where Raoult's Law (dashed lines). Henry's Law slopes are shown as dot-dashed lines. After Nordstrom and Munoz (1986).

[¶] Francois Marie Raoult (1830-1901), French chemist.
Geochemistry

CHAPTER 7: SOLUTIONS

3.3.2 Henry's Law

Another useful approximation occurs when X_i approaches 0. In this case, the partial pressures are not equal to the mole fraction times the vapor pressure of the pure substance, but they do vary linearly with X_i . This behavior follows Henry's Law[£], which is:

$$P_i = hX_i \quad for X_i << 1$$
3.10

where *h* is known as the Henry's Law constant.

3.4 CHEMICAL POTENTIAL

3.4.1 Partial Molar Quantities

Free energy and other thermodynamic properties are dependent on composition. We need a way of expressing this dependence. For any extensive property of the system, such as volume, entropy, energy, or free energy, we can define a *partial molar value*, which expresses how that property will depend on changes in amount of one component. For example, we can define partial molar volume of component *i* in phase ϕ as:

$$v_i^{\phi} = \left(\frac{\partial V}{\partial n_i}\right)_{T,P,n_{j,j\neq i}}$$
 such that $V = \sum_i n_i v_i$ 3.11

(we will use small letters to denote partial molar quantities; the superscript refers to the phase and the subscript refers to the component). The English interpretation of equation 3.11 is that *the partial molar volume of component* i *in phase* ϕ *tells us how the volume of phase* ϕ *will vary with an infinitesimal addition of component* i, *if all other variables are held constant*. For example, the partial molar volume of Na in an aqueous solution such as seawater would tell us how the volume of that solution would change for an infinitesimal addition of Na. In this case *i* would refer to the Na component and ϕ would refer to the aqueous solution phase. In Table 2.2, we see that the molar volumes of the albite and anorthite end-members of the plagioclase solid solution are different. We could define v_{Ab}^{Pl} as the partial molar volume of albite in plagioclase, which would tell us how the volume of plagioclase would vary for an infinitesimal addition of albite. (In this example, we have chosen our component as albite rather than Na. While we could have chosen Na, the choice of albite simplifies matters because the replacement of Na with Ca is accompanied by the replacement of Si by Al.)

The second expression in 3.11 says that the volume of a phase is the sum of the partial molar volumes of the components times the number of moles of each component present. Thus the volume of plagioclase would be sum of the partial molar volumes of the albite and anorthite components weighted by the number of moles of each.

Another example might be a solution of water and ethanol. The variation of the partial molar volumes of water and ethanol in a binary solution is illustrated in Figure 3.5. This system illustrates very clearly why the qualification "for an infinitesimal addition" is always added: the value of a partial molar quantity of a component may vary with the amount of that component present.

Equation 3.11 can be generalized to all partial molar quantities and also expresses an important property of partial molar quantities: *an extensive variable of a system or phase is the sum of it's partial molar quantities for each component in the system*. In our example above, this means that the volume of plagio-clase is the sum of the partial molar volume of the albite and anorthite components.

Generally, we find it more convenient to convert extensive properties to intensive properties by dividing by the total number of moles in the system, Σn . Dividing both sides of equation 3.11 by Σn we have:

$$\overline{V} = \sum_{i} X_{i} v_{i} \qquad 3.12$$

 $^{^{\}pounds}$ named for English chemist William Henry (1775-1836), who formulated it.

Geochemistry

W. M. White

CHAPTER 3: SOLUTIONS

This equation says that the molar volume of a substance is the sum of the partial molar volumes of its components times their mole fractions. For a pure phase, the partial molar volume equals the molar volume since X=1.

3.4.2 Definition of Chemical Potential and Relationship to Gibbs Free Energy

We define μ as the *chemical potential*, which is simply the partial molar Gibbs Free Energy:

$$\mu_{i} = \left(\frac{\partial G}{\partial n_{i}}\right)_{P,T,n_{j,j\neq i}} \qquad 3.13$$

The chemical potential thus tells us how the Gibbs Free Energy will vary with the number of moles, n_i , of component *i* holding temperature, pressure, and the number of moles of all other components constant. We said that the Gibbs Free Energy of a system is a measure of the capacity of the system to do chemical work. Thus the chemical potential of component *i* is the amount by which this capacity to do chemical work is changed for an infinitesimal addition of component *i* at constant temperature and pressure. In a NiCd battery (common rechargeable batteries) for example, the chemical potential of Ni in the battery (our system



Figure 3.5. Variation of the partial molar volumes of water and ethanol as a function of the mole fraction of ethanol in a binary solution. This figure also illustrates the behavior of a very non-ideal solution.

chemical potential of Ni in the battery (our system) is a measure of the capacity of the battery to provide electrical energy per mole of additional Ni for an infinitesimal addition.

The total Gibbs Free Energy of a system will depend upon composition as well as on temperature and pressure. The equations we introduced for Gibbs Free Energy in Chapter 2 fully describe the Gibbs Free Energy only for single components systems or systems containing only pure phases. The Gibbs Free Energy change of a phase of variable composition is fully expressed as:

$$dG = VdP - SdT + \sum_{i} \mu_{i} dn_{i}$$

$$3.14$$

3.4.3 Properties of the Chemical Potential

We now want to consider two important properties of the chemical potential. To illustrate these properties, consider a simple two-phase system in which an infinitesimal amount of component *i* is transferred from phase β to phase α , under conditions where T, P, and the amount of other components is held constant in each phase. One example of such a reaction would be the transfer of Pb from a hydrothermal solution to a sulfide mineral phase. The chemical potential expresses the change in Gibbs Free Energy under these conditions:

$$dG = dG^{\alpha} + dG^{\beta} = \mu_i^{\alpha} dn_i^{\alpha} + \mu_i^{\beta} dn_i^{\beta}$$
3.15

since we are holding everything else constant, what is gained by α must be lost by β , so $-dn_i^{\alpha} = dn_i^{\beta}$ and:

$$dG = \left(\mu_i^{\alpha} - \mu_i^{\beta}\right) dn_i \qquad 3.16$$

At equilibrium, dG = 0, and therefore $\mu_i^{\alpha} = \mu_i^{\beta}$ 3.17

Equation 3.17 reflects a very general and very important relationship, namely:

Geochemistry

CHAPTER 7: SOLUTIONS

In a system at equilibrium, the chemical potential of every component in a phase is equal to the chemical potential of that component in every other phase in which that component is present.

Equilibrium is the state toward which systems will naturally transform. The Gibbs Free Energy is the chemical energy available to fuel these transformations. *We can regard differences in chemical potentials as the forces driving transfer of components between phases*. In this sense, the chemical potential is similar to other forms of potential energy, such as gravitational or electromagnetic. Physical systems spontaneously transform so as to minimize potential energy. Thus for example, water on the surface of the Earth will move to a point where it's gravitational potential energy is minimized, i.e., downhill. Just as gravitational potential energy drives this motion, the chemical potential drives chemical reactions, and just as water will come to rest when gravitational energy is minimized, chemical reactions will cease when chemical potential is minimized. So in our example above, the spontaneous transfer of Pb between a hydrothermal solution and a sulfide phase will occur until the chemical potentials of Pb in the solution and in the sulfide are equal. At this point, there is no further energy available to drive the transfer.

We defined the chemical potential in terms of the Gibbs Free Energy. However, in his original work, Gibbs based the chemical potential on the internal energy of the system. As it turns out, however, the quantities are the same:

$$\mu_{i} = \left(\frac{\partial G}{\partial n_{i}}\right)_{P,T,n_{j,j\neq i}} = \left(\frac{\partial U}{\partial n_{i}}\right)_{S,V,n_{j,j\neq i}} 3.18$$

It can be further shown (but we won't) that:

$$\mu_{i} = \left(\frac{\partial G}{\partial n_{i}}\right)_{P,T,n_{j,j\neq i}} = \left(\frac{\partial U}{\partial n_{i}}\right)_{S,V,n_{j,j\neq i}} = \left(\frac{\partial H}{\partial n_{i}}\right)_{S,P,n_{j,j\neq i}} = \left(\frac{\partial A}{\partial n_{i}}\right)_{T,V,n_{j,j\neq i}}$$

3.4.4 THE GIDDS-DUHEM RELATION

Since μ is the partial molar Gibbs Free Energy, the Gibbs Free Energy of a system is the sum of the chemical potentials of each component:

$$G = \sum_{i} n_{i} \left(\frac{\partial G}{\partial n_{i}} \right)_{P,T,n_{j,j\neq i}} = \sum_{i} n_{i} \mu_{i}$$

$$3.19$$

The differential form of this equation (which we get simply by applying the chain rule) is:

$$dG = \sum_{i} n_i d\mu_i + \sum_{i} \mu_i dn_i$$
 3.20

Equating this with equation 3.14, we obtain:

$$\sum_{i} n_i d\mu_i + \sum_{i} \mu_i dn_i = VdP - SdT + \sum_{i} \mu_i dn_i$$
 3.21

Rearranging, we obtain the *Gibbs-Duhem Relation*:

$$VdP - SdT - \sum_{i} n_i d\mu_i = 0$$
3.22

The Gibbs-Duhem Equation describes the relationship between simultaneous changes in pressure, temperature and composition in a single-phase system. In a closed system at equilibrium, net changes in chemical potential will occur only as a result of changes in temperature or pressure. *At constant temperature and pressure*, there can be no net change in chemical potential at equilibrium:

$$\sum_{i} n_i d\mu_i = 0 \tag{3.23}$$

Geochemistry

CHAPTER 3: SOLUTIONS

This equation further tells us that the chemical potentials do not vary independently, but change in a related way. In a closed system, only one chemical potential can vary independently. For example, consider a two component system. Then we have $n_1d\mu_1 + n_2d\mu_2 = 0$ and $d\mu_2 = -(n_1/n_2)d\mu_1$. If a given variation in composition produces a change in μ_1 then there is a concomitant change in μ_2 .

For multi-phase systems, we can write a version of the Gibbs-Duhem relation for each phase in the system. For such systems, the Gibbs-Duhem relation allows us to reduce the number of independently variable components in each phase by one. We will return to this point later in the chapter.

We can now state an additional property of chemical potential:

In spontaneous processes, components or species are distributed between phases so as to minimize the chemical potential of all components.

This allows us to make one more characterization of equilibrium: *equilibrium is point where the chemical potential of all components is minimized*.

3.4.5 Derivation of the Phase Rule

Another significant aspect of the Gibbs-Duhem Equation is that the phase rule can be derived from it. We begin by recalling that the variance of a system (the number of variables that must be fixed or independently determined to determine the rest) is equal to the number of variables less the number of equations relating them. In a multicomponent single-phase system, consisting of *c* components, there are *c* +2 unknowns required to describe the equilibrium state of the system: T, P, μ_1 , μ_2 , ... μ_c . But in a system of ϕ phases at equilibrium, we can write ϕ versions of equation 3.23, which reduces the independent variables by ϕ . Thus the number of independent variables that must be specified to describe a system of *c* components and ϕ phases is:

$$f = c + 2 - \phi$$

which is the Gibbs phase rule.

Specification of f variables will completely describe the system, at least with the qualification that in thermodynamics we are normally uninterested in the size of the system, that is, in extensive properties such as mass, volume, etc. (though we are interested in their intensive equivalents) and outside forces or fields such as gravity, electric or magnetic fields, etc. Nevertheless, the size of the system is described as well, provided only that one of the f variables is extensive.

3.5 Ideal Solutions

Having placed another tool, the chemical potential, in our thermodynamic toolbox, we are ready to continue our consideration of solutions. We will begin with *ideal solutions*, which, like ideal gases, are fictions that avoid some of the complications of real substances. For an ideal solution, we make an assumption similar to one of those made for an ideal gas, namely that there were no forces between molecules. In the case of ideal solutions, which may be gases, liquids, or solids, we can relax this assumption somewhat and require only that *the interactions between different kinds of molecules in an ideal solution are the same as those between the same kinds of molecules*.

3.5.1 Chemical Potential in Ideal Solutions

How does chemical potential vary in an ideal solution? Consider the vapor pressure of a gas. The derivative of G with respect to pressure at constant temperature is volume:

$$\left(\frac{\partial G}{\partial P}\right)_T = V$$

Written in terms of partial molar quantities:

$$\left(\frac{\partial\mu}{\partial P}\right)_T = v$$

Geochemistry

W. M. White

CHAPTER 7: SOLUTIONS

If the gas is ideal, then:

$$\left(\frac{\partial \mu}{\partial P}\right)_{T,ideal} = \frac{RT}{P}$$
 3.24

and if we integrate from P° to P we obtain:

$$\mu^{P} - \mu^{P^{o}} = RT \ln \frac{P}{P^{o}} \qquad 3.25$$

where $\mu^{P^{\circ}}$ is the chemical potential of the pure gas at the reference (standard state) pressure P°. This is the standard-state chemical potential and is written as μ° . If we let P° be the vapor pressure of pure *i* and P be the vapor pressure of *i* in an ideal solution, then we may substitute X for P/P° into Raoult's Law (Equation 3.8) and to obtain the following:

$$\mu_{i,ideal} = \mu_i^o + RT \ln X_i$$
 3.26

This equation describes the relationship between the chemical potential of component i and its mole fraction in an ideal solution.

3.5.2 Volume, Enthalpy, Entropy, and Free Energy Changes in Ideal Solutions

We will be able to generalize a form of this equation to non-ideal cases a bit later. Let's first consider some other properties of ideal mixtures. For real solutions, any extensive thermodynamic property such as volume can be considered to be the sum of the volume of the components plus a volume change due to mixing:

$$\overline{V} = \sum_{i} X_{i} \overline{V_{i}} + \Delta V_{mixing}$$
 3.27

The first term on the right reflects the volume resulting from mechanical mixing of the various components. The second term reflects volume changes associated with solution. For example, if we mixed 100 ml of ethanol and 100 ml of water (Figure 3.5), the volume of the resulting solution would be 193 ml. Here, the value of the first term on the right would be 200 ml, the value of the second term would be -7 ml. We can write similar equations for enthalpy, etc. But the volume change and enthalpy change due to mixing are both 0 in the ideal case. This is true because both volume and enthalpy changes of mixing arise from intermolecular forces, and, by definition, such intermolecular forces are absent in the ideal case. Thus:

therefore:

and

$$\begin{split} \Delta V_{ideal\ mixing} &= 0\\ \overline{V}_{ideal} &= \sum_{i} X_{i} v_{i} = \sum_{i} X_{i} \overline{V}_{i}\\ \Delta H_{ideal\ mixing} &= 0\\ \overline{H}_{ideal} &= \sum_{i} X_{i} h_{i} = \sum_{i} X_{i} \overline{H}_{i} \end{split}$$

and therefore:

This, however, is not true of entropy. You can imagine why: if we mix two substances on an atomic level, the number of possible arrangements of our system increases even if they are ideal substances. The entropy of ideal mixing is (compare equation 2.110):

$$\Delta \overline{S}_{ideal\,mixing} = -R \sum_{i} X_{i} \ln X_{i}$$
3.28

$$\overline{S}_{ideal \, solution} = \sum_{i} X_{i} \overline{S}_{i} - R \sum_{i} X_{i} \ln X_{i}$$

$$3.29$$

Because $\Delta G_{mixing} = \Delta H_{mixing} - T\Delta S_{mixing}$ and $\Delta H_{mixing} = 0$, it follows that:

$$\Delta \overline{G}_{ideal\ mixing} = RT \sum_{i} X_{i} \ln X_{i}$$
3.30

We stated above that the total expression for an extensive property of a solution is the sum of the partial molar properties of the pure phases (times the mole fractions) plus the mixing term. The partial

September 26, 2005

Geochemistry

CHAPTER 7: SOLUTIONS

molar Gibbs Free Energy is the chemical potential, so the full expression for the Gibbs Free Energy of an ideal solution is:

$$\overline{G}_{ideal} = \sum_{i} X_{i} \mu_{i}^{o} + RT \sum_{i} X_{i} \ln X_{1}$$

$$3.31$$

Rearranging terms, we can re-express equation 3.31 as:

$$\overline{G}_{ideal} = \sum_{i} X_{i} (\mu_{i}^{o} + RT \ln X_{1})$$

$$3.32$$

The term in parentheses is simply the chemical potential of component *i*, $\mu_{i\nu}$ as expressed in equation 3.26. Substituting equation 3.26 into 3.32, we have

$$\overline{G}_{ideal} = \sum_{i} X_{i} \mu_{i}$$

$$3.33$$

Note that for an ideal solution, μ_i is always less than or equal to μ_i^o because the term RTln X_i is always negative (because the log of a fraction is always negative).

Let's consider ideal mixing in the simplest case, namely binary mixing. For a two component (binary) system, $X_1 = (1 - X_2)$, so we can write equation 3.30 for the binary case as:

$$\Delta \overline{G}_{ideal\ mixing} = RT [(1 - X_2) \ln(1 - X_2) + X_2 \ln X_2]$$
3.34

Since X_2 is less than 1, ΔG is negative and becomes increasingly negative with temperature, as illustrated in Figure 3.6. The curve is symmetrical with respect to X; i.e., the minimum occurs at $X_2 = 0.5$.

Now let's see how we can recover information on μ_i from plots such as Figure 3.6, which we will call G-bar—X plots. Substituting $X_1 = (1 - X_2)$ into equation 3.33, it becomes:

$$\Delta \bar{G}_{ideal \ solution} = \mu_1 (1 - X_2) + \mu_2 X_2 = \mu_1 + (\mu_2 - \mu_2) X_2$$
3.35

This is the equation of a straight line on such a plot with slope of $(\mu_2 - \mu_1)$ and intercept μ_1 . This line is illustrated in Figure 3.7. The curved line is described by equation 3.31. The dashed line is given by equation 3.35. Both equation 3.31 and 3.35 give the same value of \overline{G} for a given value of X_2 , such as X'_2 . Thus the straight line and the curved one in Figure 3.7 much touch at X'_2 . In fact, the straight line is the tangent to the curved one at X'_2 . The intercept of the tangent at $X_2 = 0$ is μ_1 and the intercept at X_2 .

= 1 is μ_2 . The point is, on a plot of molar free energy vs. mole fraction (a G-X diagram), we can determine the chemical potential of component *i* in a two component system by extrapolating a tangent of the free energy curve to $X_i = 1$. We see that in Figure 3.7, as X_1 approaches 1 (X_2 approaches 0), the intercept of the tangent approaches μ_1° , i.e., μ_1 approaches μ_1° . Looking at equation 3.26, this is exactly what we expect. Figure 3.7 illustrates the case of an ideal solution, but the intercept method applies to non-ideal solutions as well, as we shall see.

Finally, the solid line connecting the $\mu^{o's}$ is the Gibbs Free Energy of a mechanical mixture of components 1 and 2, which we may express as:

$$\Delta \bar{G}_{mixture} = \sum_{i} X_{i} \mu_{i}^{o} \qquad 3.36$$



Figure 3.6. Free energy of mixing as a function of temperature in the ideal case.

Geochemistry

CHAPTER 3: SOLUTIONS



Figure 3.7. Molar free energy in an ideal mixture and graphical illustration of equation 3.31. After Nordstrom & Munoz, 1986.

You should satisfy yourself that the ΔG_{mixing} is the difference between this line and the free energy curve:

$$\overline{G}_{ideal\,mix.} = \overline{G}_{ideal\,sol.} - \overline{G}_{mixture} \quad 3.37$$

3.6 REAL SOLUTIONS

We now turn our attention to real solutions, which are somewhat more complex than ideal ones, as you might imagine. We will need to introduce a few new tools to help us deal with these complexities.

3.6.1 CHEMICAL POTENTIAL IN REAL Solutions

Let's consider the behavior of a real solution in view of the two solution models we have already introduced: Raoult's Law and Henry's Law. Figure 3.8 illustrates the variation of chemical potential as a function of composition in a hypothetical real solution. We can identify 3 regions where the behavior of the chemical potential is distinct:

1.) The first is where the mole fraction of component X_i is close to 1 and Raoult's Law holds. In this case, the amount of solute dissolved in *i* is trivially small, so molecular interactions involving solute molecules do not significantly affect the thermodynamic properties of the solution, and the behavior of μ_i is close to that in an ideal solution:

$$\mu_{i,ideal} = \mu_i^o + RT \ln X_i \tag{3.26}$$

2.) At the opposite end is the case where X_i is very small. Here interactions between two component i molecules are extremely rare, and the behavior of μ_i is essentially controlled by interactions between i and those of the solvent. While the behavior of μ_i is not ideal, it is nonetheless a linear function of ln X_i . This is the region where Henry's Law holds. The compositional dependence of the chemical potential in this region can be expressed as:

$$\mu_i = \mu_i^o + RT \ln h_i X_i \qquad 3.38$$

where *h* is the Henry's Law constant defined in equation 3.10. This equation can be rewritten as:

$$\mu_i = \mu_i^o + RT \ln X_i + RT \ln h_i \tag{3.39}$$

By definition, *h* is independent of composition at constant T and P and can be regarded as adding a fixed amount to the standard state chemical potential (a fixed amount to the intercept in Fig. 3.8). By independent of composition, we mean it is independent of X_{i} , the mole fraction of the component of interest. *h* will, of course depend on the nature of the solution. For example, if Na is our component of interest, h_{Na} will not be the same for an electrolyte solution as for a silicate melt. We can define a new term, μ^* , as:

$$\mu_i^* \equiv \mu_i^o + RT \ln h_i \qquad 3.40$$

Substituting 3.40 into 3.39 we obtain: $\mu_i = \mu_i^* + RT \ln X_i$ 3.41

When plotted against $\ln X_{i}$ the chemical potential of *i* in the range of very dilute solutions is given by a straight line with slope RT and intercept μ^* (the intercept is at $X_i = 1$ and hence $\ln X_i = 0$ and $\mu_i = \mu^*$).

CHAPTER 7: SOLUTIONS

Thus μ^* can be obtained by extrapolating the Henry's Law slope to X = 1. We can think of μ^* as the chemical potential in the hypothetical standard state of Henry's Law behavior at X = 1.

3.) The third region of the plot is that region of real solution behavior between the regions where Henry's Law and Raoult's Law apply. In this region, μ is not a linear function of ln X. We will introduce a new parameter, *activity*, to deal with this region.

3.6.2 Fugacities

The tools we have introduced to deal with ideal solutions and infinitely dilute ones are based on observations of the gaseous state: Raoult's Law and Henry's Law. We will continue to make reference to gases in dealing with real solutions that follow neither law. While this approach has a largely historical basis, it is nevertheless a consistent one. So following this pattern, we will first introduce the concept of fugacity, and derive from it a more general parameter, activity.

In the range of intermediate concentrations, the partial pressure of the vapor of component i

above a solution is generally not linearly related to the mole fraction of component i in solution. Thus chemical potential of i cannot be determined from equations such as 3.26, which we derived on the assumption that the partial pressure was proportional to the mole fraction. To deal with this situation, chemists invented a fictitious partial pressure, *fugacity*. Fugacity may be thought of as the 'escaping tendency' of a real gas from a solution. It was defined to have the same relationship to chemical potential as the partial pressure of an ideal gas:

$$\mu_i = \mu_i^o + RT \ln \frac{f_i}{f_i^o}$$
3.42

where f° is the standard-state fugacity, which is analogous to standard-state partial pressure. We are free to choose the standard state, but the standard state for f° and μ° must be the same. f° is analogous to the standard state partial pressure, P°, of an ideal gas. If we chose our standard state to be the pure substance, then f° is identical to P°, but we may wish to choose some other standard state where this will not be the case. Since the behavior of real gases approaches ideal at low pressures, the fugacity will approach the partial pressure under these circumstances. Thus the second part of the definition of fugacity is:

$$\lim_{P \to 0} \frac{f_i}{P_i} = 1 \tag{3.43}$$

For an ideal gas, fugacity is identical to partial pressure. Since, as we stated above, fugacity bears the same relationship to chemical potential (and other state functions) of a non-ideal substance as pressure of a non-ideal gas, we substitute fugacity for pressure in thermodynamic equations.

The relationship between pressure and fugacity can be expressed as:

$$f = \phi P \tag{3.44}$$

where ϕ is the *fugacity coefficient*, which will be a function of temperature and pressure differ for each real gas. The fugacity coefficient expresses the difference in the pressure between a real gas and an



Figure 3.8. Schematic plot of the chemical potential of component *i* in solution as a function of $\ln X_i$. Here μ° is the chemical potential of pure i at the pressure and temperature of the diagram. After Nordstrom and Munoz (1986).

Geochemistry

CHAPTER 7: SOLUTIONS

ideal gas under comparable conditions. Kerrick and Jacobs (1981) fitted the Redlich-Kwong equation (equation 2.15) to observations on the volume, pressure and volume of H_2O and CO_2 to obtain values for the coefficients *a* and *b* in equation 2.15. From these, they obtained fugacity coefficients for these gases at a series of temperatures and pressures. These are given in Table 3.1.

3.6.3 Activities and Activity Coefficients

Fugacities are thermodynamic functions that are directly related to chemical potential and can be calculated from measured P-T-V properties of a gas, though we will not discuss how. However, they have meaning for solids and liquids as well as gases since solids and liquids have finite vapor pressures. Whenever a substance exerts a measurable vapor pressure, a fugacity can be calculated. Fugacities are relevant to the equilibria between species and phase components, because if the vapor phases of the components of some solid or liquid solutions are in equilibrium with each other, and with their respective solid or liquid phases, then the species or phases components in the solid or liquid must be in equilibrium. One important feature of fugacities is that we can use them to define another thermodynamic parameter, the *activity*, *a*:

$$a_i \equiv \frac{f_i}{f_i^o}$$
 3.45

 f° is the standard state fugacity. Its value depends on the standard state you choose. You are free to choose a standard state convenient for whatever problem you are addressing.

If we substitute equation 3.45 into equation 3.42, we obtain the important relationship:

$$\mu_i = \mu_i^o + RT \ln a_i \tag{3.46}$$

The 'catch' on selecting a standard state for f° , and hence for determining a_i in equation 3.46, is that this state must be the same as the standard state for μ° . Thus we need to bear in mind that standard states are implicit in the definition of activities and that those standard states are tied to the standard-state chemical potential. Until the standard state is specified, activities have no meaning.

Comparing equation 3.46 with 3.26 leads to:

$$a_{i,ideal} = X_i \tag{3.47}$$

Thus in ideal solutions, the activity is equal to the mole fraction.

Chemical potentials can be thought of as driving forces that determine the distribution of components between phases of variable composition in a system. Activities can be thought of as the effective concentration or the availability of components for reaction. In real solutions, it would be convenient to relate all non-ideal thermodynamic parameters to the composition of the solution, because composition is generally readily and accurately measured. To relate activities to mole fractions, we define a new parameter, the rational activity coefficient, λ . The relationship is:

$$a_i = X_i \lambda_i \qquad 3.48$$

TADLE 3.1.	H ₂ O and	CO ₂ Fuqad	сіту Со	efficients
		TOC		

H_2O		1°C		
P, MPa	400	600	800	1000
50	0.4	0.78	0.91	
200	0.2	0.52	0.79	0.94
400	0.21	0.54	0.84	1.03
600	0.28	0.67	1.01	1.22
800	0.4	0.89	1.27	1.49
CO ₂		T ℃		
P, MPa	377	577	777	977
50	1.02	1.1	1.12	1.12
200	1.79	1.86	1.82	1.75
400	4.91	4.18	3.63	3.22
600	13.85	9.48	7.2	5.83
800	38 73	21 33	14 15	10 44

From Kerrick and Jacobs (1981).

Geochemistry

CHAPTER 7: SOLUTIONS

Example 3.2. Using Fugacity to Calculate Gibbs Free Energy

The minerals brucite (Mg(OH)₂) and periclase (MgO) are related by the reaction:

$Mg(OH)_2 \rightleftharpoons MgO + H_2O$

Which side of this reaction represent the stable phase assemblage at 600° C and 200 MPa? *Answer:* We learned how to solve this sort of problem in Chapter 2: the side with the lowest Gibbs Free Energy will be the stable assemblage. Hence, we need only to calculate ΔG_r at 600° C and 200 MPa. To do so, we use equation 2.129:

$$\Delta G_{T'P'} = \Delta G^{o} - \int_{T_{ref}}^{T'} \Delta S_r dT + \int_{T_{ref}}^{T'} \Delta V_r dP \qquad (2.129)$$

Our earlier examples dealt with solids, which are incompressible to a good approximation, and we could simply treat ΔV_r as being independent of pressure. In that case, the solution to the first integral on the left was simply $\Delta V_r(P'-P_{ref})$. The reaction in this case, like most metamorphic reactions, involves H₂O, which is certainly not incompressible: the volume of H₂O, as steam or a supercritical fluid, is very much a function of pressure. Let's isolate the difficulty by dividing ΔV_r into two parts: the volume change of reaction due to the solids, in this case the difference between molar volumes of periclase and brucite, and the volume change due to H₂O. We will denote the former as ΔV^s and assume that it is independent of pressure. The second integral in 2.131 then becomes:

$$\int_{P_{ref}}^{P} \Delta V_r dP = \Delta V^s (P' - P_{ref}) + \int_{P_{ref}}^{P'} V_{H_2O} dP$$
 3.49

How do we solve the pressure integral above? One approach is to assume that H_2O is an ideal gas.

For an ideal gas:

so that the pressure integral becomes:
$$\int_{P}^{P'} \frac{RT}{P} dP = RT \ln \frac{P'}{P_{ref}}$$

Steam is a very non-ideal gas, so this approach would not yield a very accurate answer. The concept of fugacity provides us with an alternative solution. For a non-ideal substance, fugacity bears the same relationship to volume as the pressure of an ideal gas. Hence we may substitute fugacity for pressure so that the pressure integral in equation 2.129 becomes:

$$\int_{f_{ref}}^{f'} \frac{RT}{f} df = RT \ln \frac{f'}{f_{ref}}$$

where we take the reference fugacity to be 0.1 MPa. Equation 3.47 thus becomes:

$$\int_{P_{ref}}^{P'} \Delta V_r dP = \Delta V^s (P' - P_{ref}) + \int_{f_{ref}}^{f'} V_{H_2O} df = \Delta V^s (P' - P_{ref}) + RT \ln \frac{f'}{f^{ref}} \qquad 3.50$$

We can then compute fugacity using equation 3.44 and the fugacity coefficients in Table 3.1.

Using the data in Table 2.2 and solving the temperature integral in 2.129 in the usual way (equation 2.131), we calculate the $\Delta G_{T,P}$ is 3.29 kJ. Since it is positive, the left side of the reaction, i.e., brucite, we predict that brucite is stable.

The ΔS of this reaction is positive, however, implying that at some temperature, periclase plus water will eventually replace brucite. To calculate the actual temperature of the phase boundary requires a trial and error approach: for a given pressure, we must first guess a temperature, then look up a value of ϕ in Table 2.1 (interpolating as necessary), and calculate ΔG_r . Depending on our answer, we make a revised guess of T and repeat the process until ΔG is 0. Using a spreadsheet, however, this goes fairly quickly. Using this method, we calculate that brucite breaks down at 660° C at 200 MPa, in excellent agreement with experimental observations.

 $V = \frac{RT}{R}$

Geochemistry

CHAPTER 3: SOLUTIONS

The rational activity coefficient differs slightly in definition from the *practical activity coefficient*, γ , used in aqueous solutions. λ is defined in terms of mole fraction, whereas γ is defined in terms of moles of solute per moles of solvent. Consider for example the activity of Na in an aqueous sodium chloride solution. For $\lambda_{Na'}$ X is computed as:

$$X_{Na} = \frac{n_{Na}}{n_{Na} + n_{Cl} + n_{H_2O}}$$
$$\frac{n_{Na}}{n_{H_2O}}$$

whereas for $\gamma_{Na'} X_{Na}$ is:

where *n* indicates moles of substance. In very dilute solution, the difference is trivial. γ is also used for other concentration units that we will introduce in section 3.7.

3.6.4 Excess Functions

The ideal solution model provides a useful reference for solution behavior. Comparing real solutions with ideal ones leads to the concept of *excess functions*, for example:

$$G_{excess} \equiv G_{real} - G_{ideal} \tag{3.51}$$

which can be resolved into contributions of excess enthalpy and entropy:

$$G_{excess} = H_{excess} - TS_{excess}$$
 3.52

The excess enthalpy is a measure of the heat released during mixing the pure end-members to form the solution, and the excess entropy is a measure of all the energetic effects resulting from a nonrandom distribution of species in solution. We can express excess enthalpy change in the same way as excess free energy, i.e.:

$$H_{excess} \equiv H_{real} - H_{ideal} \tag{3.53}$$

But since $\Delta H_{ideal mixing} = 0$, $\Delta H_{excess} = \Delta H_{real}$; in other words, the enthalpy change upon mixing is the excess enthalpy change. Similar expressions may, of course, be written for volume and entropy (bearing in mind that unlike volume and enthalpy, ΔS_{ideal} is not zero).

Combining equation 3.46 with equation 3.48 leads to the following:

$$\mu_i = \mu_i^o + RT \ln X_i \lambda_i \qquad 3.54$$

$$\mu_i = \mu_i^o + RT \ln X_i + RT \ln \lambda_i \qquad 3.55$$

Equation 3.55 shows how activity coefficients relate to Henry's and Raoult's Laws. Comparing equation 3.55 with equation 3.39, we see that in the region where Henry's Law holds, that is dilute solutions, the activity coefficient is equal to Henry's law constant. In the region where Raoult's Law holds, the activity coefficient is 1 and equation 3.55 reduces to equation 3.26 since RT ln $\lambda_i = 0$.

Since we know that

We can rewrite this as:

$$\left(\frac{\partial G}{\partial n_i}\right)_{T,P,n_{j,j\neq i}} = \mu_i = \mu_i^o + RT \ln X_i \lambda_i$$

comparing equations 3.51 and 3.55, we find that:

$$\left(\frac{\partial G_{excess}}{\partial n_i}\right)_{T,P,n_{j,j\neq i}} = RT \ln \lambda_i$$

$$\bar{G}_{excess i} = RT \ln \lambda_i \qquad 3.56$$

which is the same as:

So that the molar excess free energy associated with component i is simply RT times the log of the activity coefficient. The total molar excess free energy of the solution is then:

$$\overline{G}_{excess} = RT \sum_{i} X_{i} \ln \lambda_{i}$$
3.57

Geochemistry

CHAPTER 7: SOLUTIONS

We will see the usefulness of the concept of excess free energy shortly when we consider activities in electrolyte solutions. It will also prove important in our treatment of non-ideal solid solutions and exsolution phenomena in the next chapter.

Depression of the Melting Point

In northern climates such as Ithaca, NY salting road and sidewalks to melt snow and ice is a common practice in winter. We have now acquired the thermodynamics tools to show why salt melts ice and that this effect does not depend on any special properties of salt or water. Depression of the melting point by addition of a second component to a pure substance is a general phenomenon. Suppose that we have an aqueous solution containing sodium chloride coexisting with pure ice. If the two phases are at equilibrium, then the chemical potential of water in ice must equal that of water in the solution, i.e.:

$$\mu_{H_{2}O}^{ice} = \mu_{H_{2}O}^{aq} \qquad 3.58$$

(we are using subscripts to denote the component, and superscripts to denote the phase; aq denotes the liquid aqueous solution). We define our standard state as that of the pure substance. According to equ. 3.48, the chemical potential of water in the solution can be expressed as:

$$\mu_{H_{2O}}^{aq} = \mu_{H_{2O}}^{o} + RT \ln a_{H_{2O}}^{aq} \qquad 3.59$$

 $\mu_{H_{2O}}^{o}$ denotes the chemical potential of pure liquid water. Substituting 3.59 into 3.58 and rearranging, we have:

$$\mu_{H_{2O}}^{ice} - \mu_{H_{2O}}^{o,aq} = RT \ln a_{H_{2O}}^{aq} \qquad 3.60$$

Ice will incorporate very little salt; if we assume it is a pure phase, we may write 3.60 as:

or

$$\mu_{H_{2O}}^{o,ice} - \mu_{H_{2O}}^{o,aq} = RT \ln a_{H_{2O}}^{aq} \qquad 3.60a$$

$$\mu_{H_{2}O}^{o,aq} - \mu_{H_{2}O}^{o,ice} = -RT \ln a_{H_{2}O}^{aq} \qquad 3.61$$

(The order is important: equation 3.60a describes the freezing process, 3.61 the melting process. These processes will have equal and opposite entropies, enthalpies, and free energies). The left hand side of 3.61 is the Gibbs Free Energy of melting for pure water, which we denote as ΔG_m° (ΔG_m° is 0 at the melting temperature of pure water, which we denote T_m° , but non-zero at any other temperature). We may rewrite 3.61 as:

$$\Delta G_m^o = -RT \ln a_{H_2O}^{aq} \qquad 3.62$$

If we assume that ΔH and ΔS are independent of temperature (which is not unreasonable over a limited temperature range) and we assume pressure is constant as well, the left hand side of the equation may also be written as:

$$\Delta G_m^o = \Delta H_m^o - T \Delta S_m^o \qquad 3.63$$

Substituting 3.63 into 3.62:

$$\Delta H_m^o - T\Delta S_m^o = -RT \ln a_{H_2O}^{aq} \qquad 3.64$$

At the melting temperature of pure water, ΔG_m° is zero, so that:

$$\Delta H_m^o = T_m^o \Delta S_m^o$$

Substituting this into 3.64 and rearranging:

$$\Delta S_m^o (T_m^o - T) = -RT \ln a_{H_2O}^{aq} \qquad 3.65$$

Further rearrangement yields:

$$\frac{T_m^o}{T} - 1 = \frac{-R}{\Delta S_m^o} \ln a_{H_2O}^{aq}$$

For a reasonably dilute solution, the activity of water will approximately equal its mole fraction, so that:

$$\frac{T_m^o}{T} - 1 = \frac{-R}{\Delta S_m^o} \ln X_{H_2O}^{aq} \qquad 3.66$$

The entropy of melting is always positive and since X is always less than 1, the left hand side of 3.66 must always be positive. Thus the ratio T_m^o /T must always be greater than 1. So the temperature at which an aqueous solution will freeze will always be less than the melting point of pure water. Salting of roads is not a question of geochemical interest, but there are many examples of depression of the freezing point of geological interest. For example, the freezing point of the ocean is about -2° C. And this phenomenon is important in igneous petrology, as we shall see in the next chapter. A related phenomenon of geological interest is elevation of the boiling point of a liquid: for example hydrothermal solutions boil at temperatures significantly above that of pure water. Can you demonstrate that elevation of the boiling point of an ideal solution depends only on the mole fraction of the solute?

Geochemistry

W. M. White

CHAPTER 7: SOLUTIONS

3.7 Electrolyte Solutions

Electrolyte solutions are solutions in which the solute dissociates to form ions, which facilitate electric conduction. Seawater is an obvious example of a natural electrolyte solution, but all natural waters are also electrolytes, though generally more dilute ones. These solutions, which Lavoisier^{*} called the "rinsings of the Earth" are of enormous importance in many geologic processes.

3.7.1 THE NATURE OF WATER AND WATER-Electrolyte Interaction

There is perhaps no compound more familiar to us than H₂O. Common place though it might be, H₂O is the most remarkable compound in nature. Its unusual properties include: the highest heat capacity of all solids and liquids except ammonia, the highest latent heat of vaporization of all substances, the highest surface tension of all liquids, its maximum density is at 4° C, with density decreasing below that temperature (negative coefficient of thermal expansion), the solid form is less dense than the liquid (negative Clapeyron slope), and finally, it is the best solvent known, dissolving more substances and in greater quantity than any other liquid. We will digress here briefly to consider the structure and properties of H₂O and the nature of water-electrolyte interactions from a microscopic perspective.

Many of the unusual properties of water arise from its non-linear polar structure, which is illus-



Figure 3.9. (a.) Structure of the water molecule. Bond angle in the liquid phase is 108°, 105° in the gas. The hydrogens retain a partial positive charge and the oxygen retains a partial positive charge. (b.) Partial structure present in liquid water. Lines connecting adjacent molecules illustrate hydrogen bonds.

trated in Figure 3.9a. The polar nature of water gives rise to van der Waals forces and the hydrogen bond discussed in Chapter 1. The hydrogen bond, which forms between hydrogens and the oxygens of adjacent molecules, imposes a dynamic partial structure on liquid water (Fig. 3.9b). These bonds continually break and new ones reform, and there is always some fraction of unassociated molecules. On average, each water molecule is coordinated by 4 other water molecules. When water boils, all hydrogen bonds are broken. The energy involved in breaking these bonds accounts for the high heat of vaporization.

The dissolving power of water is due to its *dielectric* nature. A dielectric substance is one that reduces the forces acting between electric charges. When placed between two electrically charged plates (a capacitor), water molecules will align themselves in the direction of the electric field. As a result, the molecules oppose the charge on the plates and effectively reduce the transmission of the electric field. The *permittivity*, ε , of a substance is the measure of this effect. The *relative permittivity*, or *dielectric constant*, ε_r , of a substance is defined as the ratio of the capacitance observed when the substance is placed between the plates of a capacitor to the capacitance of the same capacitor when a vacuum is present between the plates:

$$\varepsilon_r = \frac{\varepsilon}{\varepsilon_o} \tag{3.67}$$

^{*}Antoine Lavoisier (1743-1794) laid the foundations of modern chemistry in his book, *Traité de Elémentaire de Chemie,* published in 1789. He died at the guillotine during the French Revolution.

Geochemistry

W. M. White

CHAPTER 7: SOLUTIONS

where ε_0 is the permittivity of a vacuum (8.85 × 10⁻¹² C²/J m). The relative permittivity of water is 78.54 at 25°C and 1 atm. For comparison, the relative permittivity of methane, a typical non-polar molecule, is 1.7.

Water molecules surrounding a dissolved ion will tend to align themselves to oppose the charge of the ion. This insulates the ion from the electric field of other ions. This property of water accounts in large measure for its dissolving power. For example, we could easily calculate that the energy required dissociate NaCl (i.e., the energy required to move Na⁺ and Cl⁻ ions from their normal interatomic distance in a lattice, 2.36Å to infinite separation) is about 585 kJ/mol. Because water has a dielectric constant of about 80, this energy is reduced by a factor of 80, so only 7.45 kJ are required for dissociation.

The charged nature of ions and the polar nature of water result in the *solvation* of dissolved ions. Immediately adjacent to the ion, water molecules align themselves to oppose the charge on the ion, e.g., the oxygen of the water molecule will be closest to a cation (Figure 3.10). These water molecules are called the *first solvation shell* or layer and they are effectively bound to the ion, moving with it as it moves. Beyond the first solvation shell is a region of more loosely bound molecules that are only partially oriented,



Figure 3.10. Solvation of a cation in aqueous solution. In the first solvation shell, water molecules are bound to the cation and oriented so that the partial negative charge on the oxygen faces the cation. In the second solvation shell molecules are only loosely bound and partially oriented.

called the second solvation shell or layer. The boundary of this latter shell is diffuse: there is no sharp transition between oriented and unaffected water molecules. The energy liberated, called the *solvation energy*, in this process is considerable. For NaCl, for example, it is -765 kJ/mol (it is not possible to deduce the solvation energies of Na⁺ and Cl⁻ independently). The total number of water molecules bound to the ion is called the *solvation number*. Solvation effectively increases the electrostatic radius of cations by about 0.9 Å and of anions by about 0.1Å per unit of charge.

An additional effect of solvation is *electrostriction*. Water molecules in the first solvation sphere are packed more tightly than they would otherwise be. This is true, to a lesser extent, of molecules in the secondary shell. In addition, removal of molecules from the liquid water structure causes partial collapse of this structure. The net effect is that the volume occupied by water in an electrolyte solution is less than in pure water, which can lead to negative apparent molar volumes of solutes, as we shall see. The extent of electrostriction depends strongly on temperature and pressure.

A final interesting property of water is that some fraction of water molecules will *autodissociate*. In pure water at standard state conditions, one in every 10^{-7} molecules will dissociate to form H⁺ and OH⁻ ions. Although in most thermodynamic treatments the protons produced in this process are assumed to be free ions, most will combine with water molecules to form H₃O⁺ ions. OH⁺ is called the *hydroxl* ion; the H₃O⁺ is called *hydronium*.

3.7.2 Some Definitions and Conventions

The first two terms we need to define are solvent and solute. *Solvent* is the substance present in greatest abundance in a solution; in the electrolyte solutions the we will discuss here, water is always the solvent. *Solute* refers to the remaining substances present in solution. Thus in seawater, water is the solvent and NaCl, CaSO₄, etc., are the solutes. We may also refer to the individual ions as solutes.

Geochemistry

CHAPTER 7: SOLUTIONS

3.7.2.1 Concentration Units

Geochemists concerned with aqueous solutions commonly use a variety of concentration units other than mole fraction. The first is *molality* (abbreviated as lower-case *m*), which is *moles of solute per kg of solvent* (H₂O). Molality can be converted to moles solute per moles solvent unit by dividing by 55.51 mol/kg. A second unit is *molarity* (abbreviated as uppercase *M*), which is *moles of solute per liter* of solution. To convert molality to mole fraction, we would divide by the molecular weight of solvent and use the rational activity coefficient. Natural solutions are often sufficiently dilute that the difference between molality and molarity is trivial (seawater, a relatively concentrated natural solution, contains only 3.5 weight percent dissolved solids). Another common unit is weight fraction (i.e., grams per gram solution), which may take several forms, such as weight percentage, parts per thousand or parts per million (abbreviated %, ppt or ‰, ppm or mg/kg). To convert to mole fraction, one simply divides the weight of solute and H₂O by the respective molecular weights.

3.7.2.2 pH

One of the most common parameters in aqueous geochemistry is pH. pH is defined as the negative logarithm of the hydrogen ion activity:

$$pH \equiv -\log a_{H^+}$$
 3.68

3.7.2.3 Standard State and Other Conventions

The first problem we must face in determining activities in electrolyte solutions is specifying the standard state. With gases, the standard state is generally the pure substance (generally at 298 K and 1 atm), but this is generally not a reasonable choice for electrolytes. A NaCl solution will become saturated at about 0.1 X_{NaCl} , and crystalline NaCl has very different properties from NaCl in aqueous solution. By convention, a hypothetical standard state of unit activity at 1 molal concentration is chosen: $a^{\circ} = m = l$ 3.69

Activity is generally given units of molality in this case (it is dimensionless as we defined it in equation 3.45), so that in this hypothetical standard state, activity equals molality. The standard state is hypothetical because, for most electrolytes, the activity will be less than 1 in a 1 m (molal) solution. Because the *standard state* generally is unattainable in reality, we must also define an attainable *reference state*, for most electrolytes are be pathened.

from which experimental measurements can be extrapolated. *By convention, the reference state is that of an infinitely dilute solution,* i.e., the Henry's Law state. For multicomponent solutions, we also specify that the concentrations of all other components be held constant. Hence the reference state is:

$$\lim_{m \to 0} \frac{a_i}{m_i} = 1 \ (m_j \ constant) \qquad 3.70$$

This convention is illustrated in Figure 3.11. In such solutions, the activity coefficient can be shown to depend on the charge of the ion, its concentration, and the concentration of other ions in the solution as well as temperature and other parameters of the solute. Comparing 3.70 with equations 3.46 and 3.48, we see that under these conditions, the activity coefficient is 1. By referring to infinite dilution, we are removing the effect of solute-solute interactions. The standard state properties of an electrolyte solution therefore only take account of solvent-solute interactions.



Figure 3.11. Relationship of activity and molality, reference state, and standard state for aqueous solutions. After Nordstrom and Munoz (1986).

Geochemistry

W. M. White

CHAPTER 7: SOLUTIONS

Clearly, it is impossible to measure the properties of the solute, such as chemical potential or molar volume, at infinite dilution. In practice, this problem is overcome by measuring properties at some finite dilution and extrapolating the result to infinite dilution. Indeed, even at finite concentrations, it is not possible to measure directly many properties of electrolytes. Volume is a good example. One cannot measure the volume of the solute, but one can measure the volume change of the solution as a function of concentration of the solute. Then by assuming that the partial molar volume of water does not change, a partial molar volume of the solute can be calculated. This is called the apparent molar *volume*, V_A . The apparent molar volume of NaCl as a function of molarity is shown in Figure 3.12. In essence, this convention assigns all deviations from non-ideality to the solute, and allow us to use the partial molar volume of pure water in the place of the true, but unknown, molar volume of water in the solution. Thus the volume of NaCl solution is given by:

$$V = n_w \overline{V}_w + n_{NaCl} \overline{V}_{NaCl}^{aq}$$
 3.71

This convention leads to some interesting effects. For example, the apparent molar volume of magnesium sulfate increases with pressure, and many other salts, including NaCl (Fig. 3.13), exhibit the same behavior. Just as curiously, the apparent molar

volume of sodium chloride in saturated aqueous solution becomes negative above ~200° C (Figure 3.13). Many other salts show the same effect. These examples emphasize the "apparent" nature of molar volume when defined in this way. Of course, the molar volume of NaCl does not actually become negative; rather this is result of the interaction between Na⁺ and Cl⁻ and H₂O (electrostriction) and the convention of assigning all non-ideality to sodium chloride.

The concentration of a salt consisting of v_A moles of cation A and v_B moles of cation B is related to the concentration of its constituent ionic species as:

$$m_A = V_A m_{AB}$$
$$m_B = V_B m_{AB}$$

and

By convention, the thermodynamic properties of ionic species A and B are related to those of the salt AB by:

$$\Psi_{AB} \equiv v_A \Psi_A + v_B \Psi_B \tag{3.73}$$

where Ψ is some thermodynamic property. Thus the chemical potential of MgCl₂ is related to that of Mg²⁺ and Cl⁻ as:

$$\mu_{MgCl_2} = \mu_{Mg^+} + 2 \times \mu_{Cl^-}$$

3.72

The same holds for enthalpy of formation, entropy, molar volume, etc.

A final important convention is that the partial molar properties and energies of formation for the proton are taken to be zero under all conditions.



Figure 3.12. Apparent molar volume of NaCl in aqueous solution as a function of molality. The standard molar volume, V°, is the apparent molar volume at infinite dilution.



Figure 3.13. Standard molar volume of NaCl in aqueous solution as a function of temperature and pressure.

Geochemistry

CHAPTER 7: SOLUTIONS

3.7.3 Activities in Electrolytes

The assumption we made for ideal solution behavior was that interactions between molecules (species might be a better term in the case of electrolyte solutions) of solute and molecules of solvent were not different from those interactions between solvent ions only. In light of the discussion of aqueous solutions earlier, we can see this is clearly not going to be the case for an electrolyte solution. We have seen significant deviations from ideality even where the components have no net charge (e.g., waterethanol); we can expect greater deviations due to electrostatic interactions between charged species.

The nature of these interactions suggests that a purely macroscopic viewpoint, which takes no account of molecular and ionic interactions, may have severe limitations in predicting equilibria involving electrolyte solutions. Thus chemists and geochemists concerned with the behavior of electrolytes have had to incorporate a microscopic viewpoint into electrolyte theory. On the other hand, they did not want to abandon entirely the useful description of equilibria based on thermodynamics. We have already introduced concepts, the activity and the activity coefficient, which allow us to treat non-ideal behavior within a thermodynamic framework. *The additional task imposed by electrolyte solutions, and indeed all real solutions, therefore, is not to rebuild the framework, but simply to determine activities from readily measurable properties of the solution.* The dependence of all partial molar properties of a solute on concentration can be determined once the activity coefficient and its temperature and pressure dependence are known.

3.7.3.1 The Debye-Hückel and Davies Equations

Both solvent-solute and solute-solute interactions in electrolytes give rise to excess free energies and non-ideal behavior. By developing a model to account for these two kinds of interactions, we can develop an equation that will predict the activity of ions in electrolyte solution.

In an electrolyte solution, each ion will exert an electrostatic force on every other ion. These forces will decrease with the increase square of distance between ions. The forces between ions will be reduced by the presence of water molecules, due to its dielectric nature. As total solute concentration increases, the mean distance between ions will decrease. Thus we can expect that activity will depend on the total ionic concentration in the solution. The extent of electrostatic interaction will also obviously depend on the charge of the ions involved: the force between Ca^{2+} and Mg^{2+} ions will be greater at the same distance than between Na^+ and K^+ ions.

In the Debye-Hückel Theory (Fig. 3.14), a given ion is considered to be surrounded by an atmo-

sphere or cloud of oppositely charged ions (this atmosphere is distinct from, and unrelated to, the solvation shell). If it were not for the thermal motion of the ions, the structure would be analogous to that of a crystal lattice, though considerably looser. Thermal motion, however, tends to destroy this structure. The density of charge in this ion atmosphere increases with the *square root* of the ionic concentrations, but increases with the *square of* the charges on those ions. The dielectric effect of intervening water molecules will tend to reduce the interaction between ions. Debye–Hückel Theory also assumes that:

- all electrolytes are completely dissociated into ions,
- the ions are spherically symmetrical charges (hard spheres),
- the solvent is structureless; the sole property is its permittivity,
- the thermal energy of ions exceeds the electrostatic interaction energy.



Figure 3.14. An ion surrounded by a cloud of oppositely charged ions, as assumed in Debye-Hückel Theory.

Geochemistry

CHAPTER 3: SOLUTIONS

W. M. White

With these assumptions, Debye and Hückel (1923) used the Poisson–Boltzmann equation, which describes the electrostatic interaction energy between ion and a cloud of opposite charges, to derive the following relationship (see Morel and Hering, 1993 for the full derivation):

$$\log_{10} \gamma_i = \frac{-Az_i^2 \sqrt{I}}{1 + B\dot{a}_i \sqrt{I}}$$
3.74

I is *ionic strength*, in units of molality or molarity, calculated as:

$$I = \frac{1}{2} \sum_{j} m_j z_j^2 \qquad 3.75$$

where m is the concentration and z the ionic charge. The parameter å is known as the *hydrated ionic radiu, or effective radius* (significantly larger than the radius of the same ion in a crystal). A and B constants known as solvent parameters and are functions of T and P. Equation 3.74 is known as the *Debye-*

Hückel Extended Law; we will refer to it simply as the *Debye-Hückel Equation*. Table 3.2a summarizes the Debye-Hückel solvent parameters over a range of temperature and Table 3.2b gives values of å for a number of ions.

For very dilute solutions, the denominator of equation 3.74 approaches 1 (because I approaches 0), hence equation 3.74 becomes:

$$\log_{10} \gamma_i = -A z_i^2 \sqrt{I} \qquad 3.76$$

This equation is known as the *Debye-Hückel Limiting Law* (so called because it applies in the limit of very dilute concentrations).

Davies (1938, 1962) introduced an empirical modification of the Debye-Hückel equation. Davies equation is:

$$\log_{10} \gamma_i = -Az_i^2 \left[\frac{\sqrt{I}}{1 + \sqrt{I}} - bI \right] \quad 3.77$$

where A is the same as in the Debye-Hückel equation and *b* is an empirically determined parameter with a value of around 0.3. It is instructive to see how the activity coefficient of Ca²⁺ would vary according to Debye-Hückel and Davies equations if we vary the ionic strength of the solution. This variation is shown in Figure 3.15. The Davies equation predicts that activity coefficients begin to increase above ionic strengths of about 0.5 m. For reasons discussed below and in greater detail in Chapter 4, activity coefficients do actually increase at higher ionic strengths. On the whole, the Davies equation is slightly more accurate for many solutions at ionic strengths in the range of 0.1 to 1 m. Be-

TAble 3.2a Debye-Hückel Solvent Parameters

T ℃	А	$B(10^8 \text{ cm})$	
0	0.4911	0.3244	
25	0.5092	0.3283	
50	0.5336	0.3325	
75	0.5639	0.3371	
100	0.5998	0.3422	
125	0.6416	0.3476	
150	0.6898	0.3533	
175	0.7454	0.3592	
200	0.8099	0.3655	
225	0.8860	0.3721	
250	0.9785	0.3792	
275	1.0960	0.3871	
300	1.2555	0.3965	

from Helgeson and Kirkham (1974).

TAble 3.2b Debye-Hückel Effective Radii

Ion	å (10⁻° cm)
Rb ⁺ , Cs ⁺ , NH ⁺ , Ag	2.5
K ⁺ , Cl ⁻ , Br ⁻ , I ⁻ , NO ⁻ ₃	3
OH^- , F^- , HS^- , $BrO_{\overline{3}}$, $IO_{\overline{4}}$, $MnO_{\overline{4}}$	3.5
$Na^{+},HCO_{3}^{-},H_{2}PO_{4}^{-},HSO_{3}^{-},SO_{4}^{2-},HPO_{3}^{2-},PO_{3}^{3-}$	4.0-4.5
Pb ²⁺ , CO ₃ ²⁻ , SO ₃ ²⁻ ,	4.5
Sr ²⁺ , Ba ²⁺ , Cd ²⁺ , Hg ²⁺ , S ²⁻	5
$Li^+, Ca^{2+}, Cu^{2+}, Zn^{2+}, Sn^{2+}, Mn^{2+}, Fe^{2+}, Ni^{2+}$	6
Mg^{2+} , Be^{2+}	8
H^+ , Al^{3+} , trivalent rare earths	9
$Th^{4+}, Zr^{4+}, Ce^{4+}$	11

from Garrels and Christ (1982).

Geochemistry

W. M. White

CHAPTER 3: SOLUTIONS

cause of this, as well as its simplicity, the Davies equation is widely used.

3.7.3.2 Limitations to the Debye-Hückel Approach

None of the assumptions made by Debye and Hückel hold in the absolute. Furthermore, the Poisson-Boltzmann equation provides only an approximate description of ion interactions, and Debye and Hückel used an approximate solution of this equation. Thus we should not expect the Debye-Hückel equations to provide an exact prediction of activity coefficients under all conditions.

Perhaps the greatest difficulty is the assumption of complete dissociation. When ions approach each other closely, the electrostatic interaction energy exceeds the thermal energy, which violates the assumption made in the ap-



Figure 3.15. Variation of the Ca^{2+} activity coefficient with ionic strength according to the Debye-Hückel (black solid line) and Davies equations (red dashed line).

proximate solution of the Poisson-Boltzmann equation. In this case, the ions are said to be associated. Furthermore, the charge on ions is not spherically symmetric and this asymmetry becomes increasingly important at short distances. Close approach is obviously more likely at high ionic strength, so not surprisingly the Debye-Hückel equation breaks down at high ionic strength.

We can distinguish two broad types of ion associations: ion pairs and complexes. These two classes actually form a continuum, but we will define a complex as an association of ions in solution that involves some degree of covalent bonding (i.e., electron sharing). Ion pairs, on the other hand, are held together purely by electrostatic forces. We will discuss formation of ion pairs and complexes in greater detail in subsequent chapters. Here we will attempt to convey only a very qualitative understanding of these effects.

An ion pair⁺ can be considered to have formed when ions approach closer than some critical distance where the electrostatic energy, which tends to bind them, exceeds twice the thermal energy, which tends to move them apart. When this happens, the ions are electrostatically bound and their motions are linked. This critical distance depends on the charge of the ions involved and is therefore much greater for highly charged ions than for singly charged ones. As we will show in Chapter 4, ion pairs involving singly charged ions will never form, even at high ionic strengths. On the other hand, multiply charged ions will tend to form ion pairs even at very low ionic strengths.

Formation of ion pairs will cause further deviations from ideality. We can identify two effects. First, the effective concentration, or activity, of an ionic species than forms ionic associations will be reduced. Consider, for example, a pure solution of $CaSO_4$. If some fraction, α , of Ca^{2+} and SO_4^{2-} ions forms ion pairs, then the effective concentration of Ca^{2+} ions is:

$$[Ca^{2+}]_{eff} = [Ca_{2+}]_{tot} (1 - \alpha)$$

(here we follow the usual convention of using brackets to denote concentrations). The second effect is on ionic strength. By assuming complete dissociation, we similarly overestimate the effective concentration in this example by a factor of $(1 - \alpha)$.

⁺ The term *ion pair* is a bit of a misnomer because such associations can involve more than two ions. In concentrated solutions, ion pairs may consist of a cation plus several anions.

M. White

Geochemistry

CHAPTER 7: SOLUTIONS



Figure 3.16. Ionic strength of natural electrolyte solutions and the applicability of the Debye-Hückel and Davies equations.

A second phenomenon that causes deviations from ideality not predicted by Debye-Hückel is solvation. As we noted, an ion in aqueous solution are surrounded by a sphere of water molecules that are bound to it. Since those water molecules bound to the ion are effectively unavailable for reaction, the activity of water is reduced by the fraction of water molecules bound in solvation shells. This fraction is trivial in dilution solutions, but is important at high ionic strength. The result of this effect is to increase the activity of ions.

Despite these problems, Debye-Hückel has proved to be remarkably successful in predicting activity coefficients in dilute solution. The extended Debye-Hückel Equation (Equation 3.74) is most useful at concen-

trations less than 0.1 M, which includes many natural waters and provides adequate approximation for activity coefficients up to ionic strengths of about 1 M, which would include most solutions of geological interest, including seawater. As we noted above, the Davies equation is slightly more accurate in the range of 0.1 to 1 M ionic strength. Above these concentrations, both the Davies and Debye-Hückel equations are increasingly inaccurate. There are thus a variety of geological solutions for which the Debye-Hückel and Davies equations cannot be used, including hydrothermal solutions, highly saline lakes, formation brines, and aerosol particles. Figure 3.16 summarizes the typical ionic strengths of natural solutions and the applicability of these equations. The Debye-Hückel Limiting Law is useful only for very dilution solutions, less than 10^{-5} mol/kg, which is more dilute than essentially all solutions of geological interest. We will consider several methods of estimating activities in higher ionic strength solutions in Chapter 4.

3.8 Ideal Solid Solutions and Their Activities

When we deal with solid solutions, we are again faced with the inadequacy of the purely macro-

Τ	A	B_{1}	C	D	Example 3.3: Calculating Activities
lon	g/kg	$mol/kg \times 10^{\circ}$	Z-	$mz^2 \times 10^2$	Using the Debye-Hückel Equation
Cl⁻	0.0078	0.2201	1	0.2201	Given the composition for the average
SO_4^2	0.0112	0.1167	4	0.4667	river water in column A in the adjacent table,
$HCO_{\overline{3}}$	0.0583	0.9557	1	0.9557	calculate the activity of the Ca^{2+} ion at 25° C
$M e^{2+}$	0.0041	0.1687	4	0.6746	using the Debye-Hückel equation.
Ca^{2+}	0.015	0.3742	4	1.4970	Answer: Our first step is to convert these
K^+	0.0023	0.0588	1	0.0588	concentrations to molality by dividing by the
$N a^+$	0.0041	0.1782	1	0.1782	respective molecular weights. We obtain the
					motal concentrations in column B. We also

need to compute z^2 (column C), and the product z^2m (column D). Using equation 3.75, we calculate the ionic strength to be 0.00202 m. (note, one must use the ionic strength in molal or molar, and not millimolar, units in the Debye-Hückel Equation.

We substitute this value for I, then find a = 6, A = 0.5092, and B = 0.3283 in Table 3.1, and obtain a value for the activity coefficient of 0.8237, and an activity of 0.308×10^{-3} m. If we did the calculation for other temperatures, we would see that for a dilute solution such as this, the activity coefficient is a only weak function of temperature, decreasing to 0.625 at 300° C.

Geochemistry

CHAPTER 7: SOLUTIONS

scopic approach of classical thermodynamics. There is little disadvantage to this approach for gases, where the arrangement of molecules is chaotic. But the crystalline state differs from that of gases in that the arrangement of atoms in the crystal lattice is highly ordered, and the properties of the crystal depend strongly on the nature of the ordering. For this reason, we cannot afford to ignore the arrangement of atoms in solids, particularly with respect to solutions.

Solid solutions differ from those of gases and liquids in several respects. First, solution in the solid state inevitably involves substitution. While we can increase the concentration of HCl in water simply by adding HCl gas, we can only increase the concentration of Fe in biotite solid solution if we simultaneously remove Mg. Second, solid solutions involve substitution at crystallographically distinct sites. Thus in biotite a solid solution between phlogopite (KMg₃AlSi₃O₁₀(OH)₂) and annite (KFe₃AlSi₃O₁₀(OH)₂) occurs as Fe²⁺ replaces Mg²⁺ in the octahedral site; the tetrahedral Si site and the anion (O) sites remain unaffected by this substitution. Third, substitution is often coupled. For example, the solid solution between anorthite (CaAl₂Si₂O₈) and albite (NaAlSi₃O₈) in plagioclase feldspar involves not only the substitution of Na⁺ for Ca²⁺, but also the substitution of Al³⁺ for Si⁴⁺. The anorthite albite solution problem is clearly simplified if we choose anorthite and albite as our components rather than Na⁺, Ca²⁺, Al³⁺ and Si⁴⁺. Such components are known as *phase components*. Choosing pure phase end members as components is not always satisfactory either because substitution on more than one site is possible, leading to an unreasonably large number of components, or because the pure phase does not exist and hence its thermodynamic properties cannot be measured.

However we choose our components, we need a method of calculating activities that takes account of the ordered nature of the crystalline state. Here we will discuss two ideal solution models of crystalline solids. We tackle the problem of non-ideal solid solutions in Chapter 4.

3.8.1 Mixing on Site Model

Many crystalline solids can be successfully treated as ideal solutions. Where this is possible, the thermodynamic treatment and assessment of equilibrium are greatly simplified. A simple and often successful model that assumes ideality but takes account of the ordered nature of the crystalline state is the *mixing on site model*, which considers the substitution of species in any site individually. In this model, the activity of an individual species is calculated as:

$$a_{i,ideal} = (X_i)^{\nu} \tag{3.78}$$

where X is the mole fraction of the *i*th atom and *v* is the number of sites per formula unit on which mixing takes place. For example, v=2 in the Fe-Mg exchange in olivine, $(Mg,Fe)_2SiO_4$. One trick to simplifying this equation is to pick the formula unit such that v = 1. For example, we would pick $(Mg,Fe)Si_{1/2}O_2$ as the formula unit for olivine. We must then consistently choose all other thermodynamic parameters to be 1/2 those of $(Mg,Fe)_2SiO_4$.

The entropy of mixing is given by:

$$\Delta S_{ideal\ mixing} = -R \sum_{j} \left(n_j \sum_{i} X_{i,j} \ln X_{i,j} \right)$$
3.79

where the subscript *j* refers to sites and the subscript *i* refers to components, and n is the number of sites per formula unit. The entropy of mixing is the same as the configurational entropy, residual entropy, or 'third law entropy', i.e., entropy when T = 0 K. For example, in clinopyroxene, there are two exchangeable sites, a sixfold-coordinated M1 site, (Mg, Fe⁺², Fe⁺³, Al⁺³), and an eightfold-coordinated M2 site (Ca⁺², Na⁺). Here *j* ranges from 1 to 2 (e.g., 1 = M1, 2 = M2), but n = 1 in both cases (because both sites accept only one atom). *i* must range over all present ions in each site, so in this example, *i* ranges from 1 to 4 (1 = Mg, $2 = Fe^{2+}$, etc.) when *j*=1 and from 1 to 2 when *j* = 2. Since we have assumed an ideal solution, $\Delta H = 0$ and $\Delta G_{ideal} = -T\Delta S$. In other words, all we need is temperature and equ. 3.91 to calculate the free energy of solution.

In the mixing-on-site model, the activity of a phase component in a solution, for example pyrope in garnet, is the product of the activity of the individual species in each site in the phase:

Geochemistry

CHAPTER 7: SOLUTIONS

$$a_{\phi} = \prod_{i} X^{v_i} \tag{3.80}$$

where a_f is the activity of phase component ϕ , *i* are the components of pure ϕ , and v_i is the stoichiometric proportion of *i* in pure ϕ . For example, to calculate the activity of aegirine (NaFe³⁺Si₂O₆) in aegirine-augite ([Na,Ca][Fe³⁺,Fe²⁺,Mg]Si₂O₆), we would calculate the product: $X_{Na}X_{Fe^{3+}}$. Note that it would not be necessary to include the mole fractions of Si and O, since these are 1.

A slight complication arises when more than one ion occupies a structural site in the pure phase. For example, suppose we wish to calculate the activity of phlogopite (KMg₃Si₃AlO₁₀(OH)₂) in a biotite of composition K_{0.8}Ca_{0.2}(Mg_{0.17}Fe_{0.83})₃Si_{2.8}Al_{1.2}O₁₀(OH)₂. The tetrahedral site is occupied by Si and Al in the ratio of 3:1 in the pure phase end members. If we were to calculate the activity of phlogopite in pure phlogopite using equation 3.80, the activities in the tetrahedral site would contribute only $X_{Si}^{-4} X_{A1}^{-4} = (0.75)^3(0.25)^1 = 0.1055$ in the pure phase. So we would obtain an activity of 0.1055 instead of 1 for phlogopite in pure phlogopite. Since the activity of a phase component must be one when it is pure, we need to normalize the result. Thus we apply a correction by multiplying by the raw activity we obtain from 3.92 by 1/(0.1055) = 9.481, and thus obtain an activity of phlogopite of 1.

3.8.2 Local Charge Balance Model

Yet another model for the calculation of activities in ideal solid solutions is the *local charge balance* model. A common example is the substitution of Ca for Na in the plagioclase solid solution

Example 3.4. Calculating Activities Using the Mixing on Site Model

Sometimes it is desirable to calculate the activities of pure end member components in solid solutions. Garnet has the general formula $X_3Y_2Si_3O_{12}$. Calculate the activity of pyrope, Mg₃Al₂Si₃O₁₂, in a garnet solid solution of composition:

$$(Mg_{.382}Fe_{2,316}^{2+}Mn_{.167}Ca_{.156})(Al_{1.974}Fe_{.044}^{3+})Si_{3}O_{12}$$

Answer: The chemical potential of pyrope in garnet contains mixing contributions from both Mg in the cubic site and Al in the octahedral site:

$$\mu_{py}^{gt} = \mu_{py}^{o} + 3RT \ln X_{Mg} + 2RT \ln X_{Al} = RT \ln(X_{Mg}^3 X_{Al}^2)$$

The activity of pyrope is thus given by:

$$a_{py}^{gt} = X_{py}^{gt} = X_{Mg}^{3} X_{Al}^{2}$$

In the example composition above, the activity of Mg is:

$$a_{Mg} = X_{Mg}^{3} = \left(\frac{[Mg]}{[Mg] + [Fe^{2+}] + [Mn] + [Ca]}\right) = 0.126^{3} = 0.002$$

is: $a_{Al} = X_{Al}^{2} = \left(\frac{[Al]}{[Al] + [Fe^{3+}]}\right) = 0.976^{2} = 0.956$

and that of Al is:

The activity of pyrope in the garnet composition above is $0.002 \times 0.956 = 0.00191$. There is, of course, no mixing contribution from the tetrahedral site because it is occupied only by Si in both the solution and the pure pyrope phase.

(NaAlSi₃O₈—CaAl₂Si₂O₈). To maintain charge balance, the substitution of Ca²⁺ for Na⁺ in the octahedral site requires substitution of Al³⁺ for Si⁴⁺ in the tetrahedral site to maintain local charge balance. In this model, the activity of the end-member of phase component is equal to the mole fraction of the component (see Example 3.5).

Geochemistry

CHAPTER 7: SOLUTIONS

Example	3.5: Астіvіті	ies Using the Local Ch4	arge B	Alance Mod	εl	
Plagio Analys	CLASE SIS	Given the adjacent ar tivities of albite and a	nalysis anorthi	of a plagioclas te in the soluti	se crystal, ion?	what are the ac-
Oxide	Wt. percent	Answer: According to albite will be equal to	the <i>lo</i>	cal charge bal ole fraction of I	<i>ance mode</i> Na in the o	<i>l</i> , the activity of ctahedral site. To
SiO ₂	44.35	calculate this, we first	st must	convert the v	veight perc	ent oxides to for-
Al_2O_3	34.85	mula units of cation.	The fir	st step is to c	alculate th	e moles of cation
CaO	18.63	from the oxide weigh	t perce	ents. First, we	e can conve	rt weight percent
N a ₂ O	0.79	oxide to weight percei	$\frac{1}{2}$ wt $\frac{1}{2}$	n using the for	mula: oxide X	
K ₂ O	0.05	atomic wt.	cation	\times formula uni	its cation in	oxide
			mo	lecular wt. oxi	ide	
				wt % cation		
Next, we	calculate the mo	oles of cation: moles cation	$on = \overline{a}$	tomic wt. catio	on	
Combinin	ng these two equ	ations, the 'atomic wt. cati	on' terr	ns cancel and v	ve have:	
		forn	nula un	its cation in or	xide	
	moles ca	ation = wt % oxide \times	molecu	ılar wt. oxide		
Next, we	want to calculat	e the number of moles of ea	ach cati	on per formula	unit. A ge	eneral formula for
feldspar i	s: XY ₄ O ₈ , where	X is Na, K, or Ca in the 'A	A' site a	nd Y is Al or Si	i in the tetr	ahedral site. So
to calculat	te formula units	in the 'A' site, we divide	the nur	nber of moles of	of Na, K, a	nd Ca by the sum
of moles o	of Na, K, and Ca	. To calculate formula un	its in tl	ne tetrahedral	site, we d	ivide the number
of moles of	of AI and Si by	the sum of moles of A1	-			T
and Si and	d multiply by 4,	since there are 4 ions in		CATION	ORMULAL	NITS
uns site.	For to those quar	atition of the moles per 8		Mol. wt.	moles	moles per
oxygens.	The table belo	w shows the results of		oxide	cation	8 oxygens
these calo	culations.		Si	60.06	0.7385	2.077
The ac	ctivity of albit	e is equal to the mole	Al	101.96	0.6836	1.923
fraction o	of Na, 0.07; the	activity of anorthite is	Ca	56.08	0.3322	0.926
0.93.			Na	61.98	0.0255	0.071
			Κ	94.2	0.0011	0.003

3.9 Equilibrium Constants

Now that we have introduced the concepts of activity and activity coefficients, we are ready for one of the most useful parameters in physical chemistry: the equilibrium constant. Though we can predict the equilibrium state of a system, and therefore the final result of a chemical reaction, from the Gibbs Free Energy alone, the equilibrium constant is a convenient and succinct way express this. As we shall see, it is closely related to, and readily derived from, the Gibbs Free Energy.

3.9.1 Derivation and Definition

Consider a chemical reaction such as:

$$aA + bB \rightleftharpoons cC + dD$$

carried out under isobaric and isothermal conditions. The Gibbs Free Energy change of this reaction can be expressed as:

$$\Delta G = c\mu_c + d\mu_d - a\mu_a - b\mu_b \tag{3.81}$$

At equilibrium, ΔG must be zero. A general expression then is:

Geochemistry

CHAPTER 3: SOLUTIONS

$$\Delta G = \sum_{i} v_{i} \mu_{i} = 0 \qquad 3.82$$

where v_i is the stoichiometric coefficient of species *i*. Equilibrium in such situations need not mean that all the reactants (i.e., those phases on the left side of the equation) are consumed to leave only products. Indeed, this is generally not so. Substituting 3.46 into 3.82 we obtain:

$$\sum_{i} \mathbf{v}_{i} \mathbf{\mu}_{i}^{o} + RT \sum_{i} \mathbf{v}_{i} \ln a_{i} = 0$$

$$3.83$$

or:

$$\sum \mathbf{v}_i \boldsymbol{\mu}_i^o + RT \ln \prod_i a_i^{\mathbf{v}_i} = 0$$
 3.84

The first term is simply the standard state Gibbs Free Energy change, ΔG° , for the reaction. There can be only one fixed value of ΔG° for a fixed standard state pressure and temperature, and therefore of the activity products. The activity products are therefore called the *equilibrium constant* K, familiar from elementary chemistry:

$$\mathbf{K} = \prod_{i} a_{i}^{\mathbf{v}_{i}}$$
 3.85

Substituting 3.85 into 3.84 and rearranging, we see that the equilibrium constant is related to the Gibbs Free Energy change of the reaction by the equation:

$$\Delta G_r^o = -RT \ln \mathbf{K} \tag{3.86}$$

At this point, it is worth saying some more about 'standard states'. We mentioned that one is free to choose a standard state, but there are pitfalls. In general, there are two kinds of standard states, fixed pressure-temperature standard states and variable P-T standard states. If you chose a fixed temperature standard state, then equation 3.86 is only valid at that standard-state temperature. If you chose a variable-temperature standard state, then 3.86 is valid for all temperatures, but ΔG° is then a function of temperature. The same goes for pressure.

3.9.2 The Law of Mass Action

Let's attempt to understand the implications of equation 3.85. Consider the dissociation of carbonic acid, an important geological reaction:

$$H_2CO_3 = HCO_3^- + H^+$$

For this particular case, equation 3.85 is expressed as:

$$\mathbf{K} = \frac{a_{HCO_3^-} a_{H^+}}{a_{H_2CO_3}}$$

The right side of the equation is a quotient, the product of the activities of the products divided by the product of the activities of the reactants, and is called the *reaction quotient*. At equilibrium, the reaction quotient is equal to the equilibrium constant. The equilibrium constant therefore allows us to predict the relative amounts of products and reactants that will be present when a system reaches equilibrium.

Suppose now that we prepare a beaker of carbonic acid solution; it's not hard to prepare: we just allow pure water to equilibrate with the atmosphere. Let's simplify things by assuming that this is an ideal solution. This allows us to replace activities with concentrations (the concentration units will dictate how we define the equilibrium constant; see below). When the solution has reached equilibrium, just enough carbonic acid will have dissociated so that the reaction quotient will be equal to the equilibrium constant. Now let's add some H⁺ ions, perhaps by adding a little HCl. The value the reaction quotient increases above that of the equilibrium constant and the system is no longer in equilibrium.

Geochemistry

CHAPTER 3: SOLUTIONS

Systems will always respond to disturbances by moving toward equilibrium (how fast they respond is another matter, and one that we will address in Chapter 5). The system will respond by adjusting the concentrations of the 3 species until equilibrium is again achieved, in this case, hydrogen and bicarbonate ions will combine to form carbonic acid until the reaction quotient again equals the equilibrium constant. We can also see that had we reduced the number of hydrogen ions in the solution (perhaps by adding a base), the reaction would have been driven the other way: i.e., hydrogen ions would be produced by dissociation. Equation 3.85 is known as the *Law of Mass Action*, which we can state more generally as: *changing the concentration of one species to a system undergoing reaction will cause the reaction to be driven in a direction that minimizes that change*.

3.9.2.1 Le Chatelier's Principle

We can generalize this principle to the effects of temperature and pressure as well. Recall that:

$$\left(\frac{\partial\Delta G_r}{\partial P}\right)_T = \Delta V$$
 (2.129) and $\left(\frac{\partial\Delta G_r}{\partial T}\right)_T = -\Delta S_r$ (2.130)

and that systems respond to changes imposed on them by minimizing G. Thus a system undergoing reaction will respond to an increase in pressure by minimizing volume. Similarly, it will respond to an increase in temperature by maximizing entropy. The reaction ice \rightarrow water illustrates this. If the pres-

Example 3.6. Manipulating Reactions and Equilibrium Constant Expressions

Often we encounter a reaction for which we have no value of the equilibrium constant. In many cases, however, we can derive an equilibrium constant by considering the reaction of interest to be the algebraic sum of several reactions for which we do have equilibrium constant values. For example, the concentration of carbonate ion is often much lower than that of the bicarbonate ion. In such cases, it is more convenient to write the reaction for the dissolution of calcite as:

$$CaCO_3 + H_2O \rightleftharpoons Ca^{2+} + HCO_3 + OH^-$$
 3.87

Given the following equilibrium constants, what is the equilibrium constant expression for the above reaction?

$$\mathbf{K}_{2} = \frac{a_{H^{+}}a_{CO_{3}^{2-}}}{a_{HCO_{3}^{-}}} \qquad \qquad \mathbf{K}_{cal} = \frac{a_{Ca^{2+}}a_{CO_{3}^{2-}}}{a_{CaCO_{3}}} \qquad \qquad \mathbf{K}_{H_{2}O} = \frac{a_{H^{+}}a_{OH^{-}}}{a_{H_{2}O}}$$

Answer: Reaction 3.87 can be written as the algebraic sum of three reactions:

+
$$CaCO_3 \rightleftharpoons Ca^{2+} + CO_3^{2-}$$

+ $H_2O \rightleftharpoons H^+ + OH^-$
- $HCO_3^- \rightleftharpoons H^+ + CO_3^{2-}$
 $\overline{CaCO_3 - HCO_3^- + H_2O \rightleftharpoons Ca^{2+} + OH^-}$

The initial inclination might be to think that if we can sum the reactions, the equilibrium constant of the resulting reaction is the sum the equilibrium constants of the components ones. However, this is not the case. Whereas we sum the reactions, we take the product of the equilibrium constants. Thus our new equilibrium constant is:

$$\mathbf{K} = \frac{\mathbf{K}_{cal} \mathbf{K}_{H_2 O}}{\mathbf{K}_2}$$

For several reasons (chief among them is that equilibrium constants can be very large or very small numbers), it is often more convenient to work with the log of the equilibrium constant. A commonly used notation is pK. pK is the negative logarithm (base 10) of the corresponding equilibrium constant (note this notation is analogous to that used for pH). The pK's sum and our equilibrium constant expression is: $pK = pK_{cal} - pK_{H_2O} - pK_2$

Geochemistry

CHAPTER 7: SOLUTIONS

sure is increased on a system containing water and ice, the equilibrium will shift to favor the phase with the least volume, which is water (recall that water is unusual in that the liquid has a smaller molar volume than the solid). If the temperature of that system is increased, the phase with the greatest molar entropy is favored, which is also water.

Another way of looking at the effect of temperature is to recall that:

$$\Delta S \ge \frac{\Delta Q}{T}$$

Combining this with equation 2.130, we can see that if a reaction $A + B \rightarrow C + D$ generates heat, then increasing the temperature will retard formation of the products, i.e., the reactants will be favored.

A general statement that encompasses both the law of mass action and the effects we have just discussed is then:

When perturbed, a system reacts to minimize the effect of the perturbation.

This is known as Le Chatelier's Principle.

3.9.3 K_D Values, Apparent Equilibrium Constants, and the Solubility Product

It is often difficult to determine activities for phase components or species, and therefore it is more convenient to work with concentrations. We can define a new 'constant', the distribution coefficient, K_{Dr} as:

$$K_{\rm D} = \prod_i X_i^{v_i}$$
 3.88

 K_D is related to the equilibrium constant K as:

$$K_{\rm D} = \frac{K_{eq}}{K_{\lambda}}$$
 3.89

where K_{λ} is simply the ratio of activity coefficients:

$$\mathbf{K}_{\lambda} = \prod_{i} \lambda_{i}^{\mathbf{v}_{i}} \tag{3.90}$$

Distribution coefficients are functions of temperature and pressure, as are the equilibrium constants, though the dependence of the two may differ. The difference is that K_D values are also functions of composition.

An alternative to the distribution coefficient is the apparent equilibrium constant, which we define as:

$$\mathbf{K}^{app} = \prod_{i} m_{i}^{\mathbf{v}_{i}} \qquad 3.91 \qquad \mathbf{K}^{app} = \frac{\mathbf{K}_{eq}}{\mathbf{K}_{\gamma}} \qquad 3.92$$

with K_{γ} defined analogously to K_{λ} . The difference between the apparent equilibrium constant and the distribution coefficient is that we have defined the former in terms of molality and the latter in terms of mole fraction. Igneous geochemists tend to use the distribution coefficient, aqueous geochemists the apparent equilibrium constant.

Another special form of the equilibrium constant is the *solubility product*. Consider the dissolution of NaCl in water. The equilibrium constant is:

$$\mathbf{K} = \frac{a_{Na_{aq}^+} a_{Cl_{aq}^-}}{a_{NaCl_s}}$$

where *aq* denotes the dissolved ion and *s* denotes solid. Because the activity of NaCl in pure sodium chloride solid is 1, this reduces to:

$$\mathbf{K} = a_{Na_{aq}^+} a_{Cl_{aq}^-} = \mathbf{K}_{sp}$$
 3.93

CHAPTER 7: SOLUTIONS

where K_{sp} is called the *solubility product*. You should note that it is generally the case in dissolution reactions such as this that we take the denominator, i.e., the activity of the solid, to be 1.

3.9.4 HENRY'S LAW AND GAS SOLUBILITIES

Consider a liquid, water for example, in equilibrium with a gas, the atmosphere for example. Earlier in this chapter, we found that the partial pressure of component i in the gas could be related to the concentration of a component i in the liquid by Henry's Law:

$$P_i = h_i X_i \tag{3.10}$$

where h is Henry's Law constant. We can rearrange this as:

$$h_i = \frac{P_i}{X_i} \qquad 3.94$$

Notice that this equation is analogous in form to the equilibrium constant expression (3.88), except that we have used a partial pressure in place of one of the concentrations. A Henry's Law constant is thus a form of equilibrium constant used for gas solubility: it relates the equilibrium concentration of a substance in a liquid solution to that component's partial pressure in a gas.



Figure 3.17. Log of the solubility constant of barite plotted against the inverse of temperature. The slope of a tangent to the curve is equal to $-\Delta H/R$. The intercept of the tangent (which occurs at 1/T = 0 and is off the plot) is equal to $\Delta S/R$. After Blount (1977).

3.9.5 TEMPERATURE DEPENDENCE OF EQUILIBRIUM CONSTANT

Since $\Delta G^{\circ} = \Delta H^{\circ} - T\Delta S^{\circ}$ and $\Delta G_{r}^{\circ} = -RT \ln K$, it follows that in the standard state, the equilibrium constant is related to enthalpy and entropy change of reaction as:

$$\ln K = -\frac{\Delta H_r^o}{RT} + \frac{\Delta S_r^o}{R}$$
 3.95

Equation 3.95 allows us to calculate an equilibrium constant from fundamental thermodynamic data. Conversely, we can estimate values for ΔS° and ΔH° from the equilibrium constant, which is readily calculated if we know the activities of reactants and products. Equation 3.95 has the form:

Example 3.7: Using the Solubility Product. The apparent (molar) solubility product of fluorite (CaF₂) at 25° C is 3.9×10^{-11} . What is the concentration of Ca²⁺ ion in groundwater containing 0.1 mM of F⁻ in equilibrium with fluorite?

Answer: Expressing equation 3.93 for this case we have:

$$K_{sp-Fl} = \frac{[Ca^{2^+}][F^-]^2}{[CaF_2]} = [Ca^{2^+}][F^-]^2$$

We take the activity of CaF_2 as 1. Rearranging and substituting in values, we have:

$$[Ca^{2+}] = \frac{K_{sp-Fl}}{[F^{-}]^{2}} = \frac{3.9 \times 10^{-11}}{[0.1 \times 10^{-3}]^{2}} = \frac{3.9 \times 10^{-11}}{1 \times 10^{-8}} = 3.9 \times 10^{-3} M = 3.9 mM$$

Geochemistry

CHAPTER 7: SOLUTIONS

$$\ln \mathbf{K} = \frac{a}{T} + b$$

where *a* and *b* are $\Delta H^{\circ}/R$ and $\Delta S^{\circ}/R$ respectively. If we can assume that ΔH and ΔS are constant over some temperature range (this is likely to be the case provided the temperature interval is small), then a plot of ln K vs. 1/T will have a slope of $\Delta H^{\circ}/R$ and an intercept of $\Delta S^{\circ}/R$. Thus measurements of ln K made over a range of temperature and plotted vs. 1/T provide estimates of ΔH° and ΔS° . Even if ΔH and ΔS are not constant, they can be estimated from the instantaneous slope and intercept of a curve of ln K plotted against 1/T. This is illustrated in Figure 3.17, which shows measurements of the solubility constant for barite (BaSO₄) plotted in this fashion (though in this case the log₁₀ rather than natural logarithm is used). From changes of ΔH and ΔS with changing temperature and knowing the heat capacity of barite, we can also estimate heat capacities of the Ba²⁺ and SO²₄⁻ions, which would obviously be difficult to measure directly. We can, of course, also calculate ΔG directly from equation 3.86. Thus a series of measurements of the equilibrium constant for simple systems allows us deduce the fundamental thermodynamic data needed to predict equilibrium in more complex systems.

Taking the derivative with respect to temperature of both sides of equation 3.95 (while holding pressure constant), we have:

$$\left(\frac{\partial \ln K}{\partial T}\right)_{P} = \frac{\Delta H_{r}^{o}}{RT^{2}}$$
3.96

This equation is known as the *van't Hoff Equation*.

Example 3.8. Calculating Equilibrium Constants and Equilibrium Concentrations The hydration of olivine to form chrysotile (a serpentine mineral) may represented in a pure Mg system as:

$$H_2O + 2H^+ + 2Mg_2SiO_4 \rightleftharpoons Mg_3Si_2O_5(OH)_4 + Mg^{2+}$$

Species	∆H° kJ	S° J/K
$Mg_3Si_2O_5(OH)_4$	-4272.87	434.84
Mg^{2+}	-366.46	109.05
H^{+}	44.87	106.68
Mg_2SiO_4	-2132.75	186.02
H ₂ O	-232.19	211.50

If this reaction controlled the concentration of Mg^{2+} of the metamorphic fluid, what would be the activity of Mg^{2+} be in that fluid if it had a pH of 4.0 at 300° C?

Answer: Helgeson (1967) gives the thermodynamic data shown in the adjacent table for the reactants at 300° C. From these data, we use Hess's Law to calculate Δ H_r and Δ S_r as -231.38 kJ and -253.01 J/K respectively. The equilibrium constant for the reaction may be

calculated as:

$$\mathbf{K} = \exp\left(-\frac{\Delta H_r^o}{RT} + \frac{\Delta S_r^o}{R}\right) = \exp\left(-\frac{-231.38 \times 10^3}{8.134 \times 573} + \frac{253.01}{8.314}\right) = 7.53 \times 10^7$$

The equilibrium constant for this reaction can be written as:

 $\mathbf{K} = \frac{a_{Mg^{2+}} a_{Cry}}{a_{H^+}^2 a_{Fo}^2 a_{H_2O}}$ which reduces to $\mathbf{K} = \frac{a_{Mg^{2+}}}{a_{H^+}^2}$ if we take the activities of water, chrysotile, and forsterite as 1. Since pH = - log \mathbf{a}_{H^+} we may rearrange and obtain the activity of the magnesium ion as:

 $a_{Mg^{2+}} = K a_{H^+}^2 = 7.53 \times 10^7 \times 10^{-4 \times 2} = 7.53 \times 10^{-1}$

Geochemistry

CHAPTER 7: SOLUTIONS

3.9.6 Pressure Dependence of Equilibrium Constant

Si

nce
$$\left(\frac{\partial\Delta G_r}{\partial P}\right)_T = \Delta V$$
 and $\Delta G_r^o = -RT \ln K$
 $\left(\frac{\partial\ln K}{\partial P}\right)_T = -\frac{\Delta V_r^o}{RT}$ 3.97

then

If ΔV_r does not depend on pressure, this equation can be integrated to obtain:

$$\ln K_{P_2} = \ln K_{P_1} - \frac{\Delta V_r^o}{RT} (P_1 - P_2)$$

This assumption will be pretty good for solids because their compressibilities are very low, but slightly less satisfactory for reactions involving liquids (such as dissolution), because they are more compressible. This assumption will be essentially totally invalid for reactions involving gases, because their volumes are highly pressure dependent.

3.10 Practical Addroach to Electrolyte Equilibrium

With the equilibrium constant now in our geochemical toolbox, we have the tools necessary to roll up our sleeves and get to work on some real geochemical problems. Even setting aside non-ideal behavior, electrolyte solutions, geological ones in particular, often have many components and can be extremely complex. Predicting their equilibrium state can therefore be difficult. There are, however, a few rules for approaching problems of electrolyte solutions that, when properly employed, make the task much more tractable.

3.10.1 Choosing Components and Species

We emphasized at the beginning of the chapter the importance of choosing the components in a system. How well we chose components will make a difference in how easily we can solve a given problem. Morel and Hering (1993) suggest these rules for choosing components and species in aqueous systems:

- 1. All species should be expressible as stoichiometric functions of the components, the stoichiometry being defined by chemical reactions.
- 2. Each species has a unique stoichiometric expression as a function of the components.
- 3. *H*₂*O* should always be chosen as a component.
- 4. H^+ should always be chosen as a component.

H⁺ activity, or pH, is very often the critical variable, also called the "master variable", in problems in natural waters. In addition, recall that we define the free energy of formation of H⁺ as 0. For these reasons, it is both convenient and important that H⁺ be chosen as a component.

3.10.2 Mass Balance

This constraint, also sometimes called mole balance, is a very simple one, and as such it is easily overlooked. When a salt is dissolved in water, the anion and cation are added in stoichiometric proportions. If the dissolution of the salt is the only source of these ions in the solution, then for a salt of composition $C_{\nu+}A_{\nu-}$ we may write:

$$v^{-}[C] = v^{+}[A]$$
 3.98

Thus, for example, for a solution formed by dissolution of CaCl₂ in water the concentration of Cl⁻ ion will be twice that of the Ca^{2+} ion. Even if $CaCl_2$ is not the only source of these ions in solution, its congruent dissolution allows us to write the mass balance constraint in the form of a differential equation:

Geochemistry

Mg

Κ

Са

Cl

 NO_3

SO₄

CHAPTER 7: SOLUTIONS

Example 3.9. Determining the pH of Rain Water from its Composition

Determine the pH of the two samples of rain in the adjacent table. Assume that sulfuric and nitric acid are fully dissociated and that the ions in the table, along with H^+ and OH^- are the only ones present.

Answer: This problem is simpler than it might first appear. Analysis of Rain WATER Given the stated conditions, there are no reactions between these species that we need to concern ourselves with To solve the problem, we observe that this solution must be electrically ${
m N}\,{
m a}$ neutral: any difference in the sum of cations and anions must be due to the one or both of the two species not listed: OH^- and H^+ .

We startby making an initial guess that the rain is acidic and that the concentration of H⁺ will be much higher than that of OH⁻, and thatwe can therefore neglect the latter (we'll want to verify this assumption when we have obtained a solution). The rest is straightforward. We sum the product of charge times concentration (Equation 3.99) for both cations and anions and find

thatanions exceed cations in both cases: the difference is equal to the concentration of H^+ . Taking the log of the concentration (having first converted concentrations to M from μ M by multiplying by 10⁻⁶) we obtain a pH of 4.6 for the first

sample and 3.14 for the second. Now we need to check our simplifying assumption

thatwe could neglect OH⁻. The equilibrium between OH^{-} and H^{+} is given by:

$$K = [H^+][OH^-] = 10^{-14}$$

From this we compute $[OH^-]$ as 10^{-10} in the first case and 10⁻¹¹ in the second. Including these would not change the anion sum significantly, so our assumption was justified.

$$\frac{\partial Cl^{-}}{\partial Ca^{2+}} = 2$$

which just says that $CaCl_2$ dissolution adds two Cl^- ions to solution for every Ca^{2+} ion added.

By carefully choosing components and boundaries of our system, we can often write conservation equations for components. For example, suppose we have a liter of water containing dissolved CO_2 in equilibrium with calcite (for example, groundwater in limestone). In some circumstances, we may want to choose our system as the water plus the limestone, in which case we may consider Ca conserved and write:

$$\Sigma Ca = Ca_{aq}^{2+} + CaCO_{3s}$$

where $CaCO_{3s}$ is calcite (limestone) and Ca_{aq}^{2+} is aqueous calcium ion. We may want to avoid choosing carbonate as a component and chose carbon instead, since the carbonate ion is not conserved because of association and dissociation reactions such as:

$$CO_3^2 + H^+ \rightleftharpoons HCO_3^-$$

Choosing carbon as a component has the disadvantage that some carbon will be present as organic compounds, which we may not wish to consider. A wiser choice is to define CO_2 as a component. Total CO_2 would then include all carbonate species as well as CO_2 (very often, total CO_2 is expressed instead as total carbonate). The conservation equation for total CO_2 for our system would be:

$$\Sigma CO_2 = CaCO_{38} + CO_2 + H_2CO_3 + HCO_3 + CO_3^2 + CO_3^2$$

CHARGE BALANCE FOR RAINWATER
D' 1 D' 0

Rain 1

 (μM)

9

4

5

8

17

10

18

Rain 2

 (μM)

89

16

9

37

101

500

228

	Rain 1	Rain 2
Σ cations	38	204
Σ anions	63	1057
Δ	25	853
рН	4.60	3.07

Geochemistry

CHAPTER 3: SOLUTIONS

Here we see the importance of the distinction we made between components and species earlier in the chapter.

3.10.3 Electrical Neutrality

There is an additional condition that electrolyte solutions must meet: *electrical neutrality*. Thus the sum of the positive charges in solutions must equal the sum of the negative ones, or:

$$\sum_{i} m_i z_i = 0 \tag{3.99}$$

While this presents some experimental obstacles, for example we cannot add only Na⁺ ion to an aqueous solution while holding other compositional parameters constant, it also allows placement of an addition mathematical constraint on the solution. It is often convenient to rearrange equ. 3.99 so as to place anions and cations on different sides of the equation:

$$\sum_{i} m_{i}^{+} z_{i}^{+} = \sum_{n} \bar{m}_{n}^{-} \bar{z}_{n}^{-}$$
3.100

As an example, consider an natural water in equilibrium with atmospheric CO_2 and containing no other species. The charge balance equation in this case is:

$$[H^+] = [OH^-] + [HCO_3^-] + 2[CO_3^2^-]$$

As example 3.9 illustrates, the electrical neutrality constraint can prove extremely useful.

3.10.4 Equilibrium Constant Expressions

For each chemical reaction in our system, we can write one version of equation 3.85. This allows us to relate the equilibrium activities of the species undergoing reaction in our system to one another.

Solution of aqueous equilibria problems often hinge on the degree to which we can simplify the problem by minimizing the number of equilibrium constant expressions we must solve. For example, H_2SO_4 will be completely dissociated in all but the most acidic natural waters, so we need not deal with equilibrium between H⁺, SO_4^2 ⁻, HSO_4^- , and H_2SO_4 , and need not consider the two latter in our list of species. Similarly, though many natural waters contain Na⁺ and Cl⁻, NaCl will precipitate only from concentrated brines, so we generally need not consider equilibrium between NaCl, Na⁺, and Cl⁻.

Carbonate is a somewhat different matter. Over the range of compositions of natural waters, H₂CO₃,

Example 3.10: Soil Organic Acid

Consider soil water with a pH of 7 containing a weak organic acid, which we will designate HA, at a concentration of 1×10^{-4} M. If the apparent dissociation constant of the acid is $10^{-4.5}$, what fraction of the acid is dissociated?

Answer: We have two unknowns: the concentration of the dissociated and undissociated acid and we have two equations: the equilibrium constant expression for dissociation, and mass balance equation. We'll have to solve the two simultaneously to obtain the answer. Our two equations are:

$$K_{dis} = \frac{[H^+][A^-]}{[HA]} = 10^{-4.5} \qquad \Sigma HA = [HA] + [A^-]$$

Solving the dissociation constant expression for [A⁻] we have:

$$[A^{-}] = \frac{[HA]K_{dis}}{[H^{+}]}$$

Then solving the conservation equation for [HA] and substituting, we have

$$[A^{-}] = \frac{(\Sigma HA - [A^{-}])K_{dis}}{[H^{+}]}$$

Setting H⁺ to 10^{-7} and Σ HA to 10^{-4} , we calculate [A⁻] as 3.16×10^{-5} M, so 31.6% of the acid is dissociated.



CHAPTER 7: SOLUTIONS

 HCO_3^- , and $CO_3^2^-$ may all be present. In most cases, however, one of these forms will dominate and the concentrations of the remaining ones will be an order of magnitude or more lower than that of the dominant one. In some cases, two of the above species may have comparable concentrations and will be have to consider equilibrium between them, but it is rarely necessary to consider equilibrium between all three. Thus at most we will have to consider equilibrium between H_2CO_3 and HCO_3^- , or HCO_3^- and $CO_3^2^-$, and we can safely ignore the existence of the remaining species. A successful solution of problems involving carbonate equilibriu often requires correctly deciding which reactions to ignore. We will discuss carbonate equilibrium in greater detail in Chapter 6.

3.11 Oxidation and Reduction

An important geochemical variable that we have not yet considered is the *oxidation state* of a system. Many elements exist in nature in more than one valence state. Iron and carbon are the most important of these because of their abundance. Other elements, including transition metals such as Ti, Mn, Cr, Ce, Eu, and U, and non-metals such as N, S, and As, are found in more than one valence state in nature. The valence state of an element can significantly affect its geochemical behavior. For example, U is quite soluble in water in its oxidized state, U⁶⁺, but is much less soluble in its reduced state, U⁴⁺. Many uranium deposits have formed when an oxidized, U-bearing solution was reduced. Iron is reasonably soluble in reduced form, Fe²⁺, but much less soluble in oxidized form, Fe³⁺. The same is true of manganese. Thus iron is leached from rocks by reduced hydrothermal fluids and precipitated when these fluids mix with oxidized seawater. Eu²⁺ in magmas substitutes readily for Ca in plagioclase, whereas Eu³⁺ does not. The mobility of pollutants, particularly toxic metals, will depend strongly on the whether the environment is reducing or oxidizing. Thus the oxidation state of a system is an important geochemical variable.

The *valence number* of an element is defined as the electrical charge an atom would acquire if it formed ions in solution. For strongly electronegative and electropositive elements that form dominantly ionic bonds, valence number corresponds to the actual state of the element in ionic form. However, for elements that predominantly or exclusively form covalent bonds, valence state is a somewhat hypothetical concept. Carbon, for example, is never present in solution as a monatomic ion. Because of this, assignment of valence number can be a bit ambiguous. A few simple conventions guide assignment of valence number:

- The valence number of all elements in pure form is 0.
- The sum of valence numbers assigned to atoms in molecules or complex species must equal the actual charge on the species.
- The valence number of hydrogen is +1, except in metal hydrides, when it is -1.
- The valence number of oxygen is -2 except in peroxides, when it is -1.

The valence state in which an element will be present in a system is governed by the availability of electrons. Oxidation-reduction (*redox*) reactions involve the transfer of electrons and the resultant change in valence. Oxidation is the loss of electrons; reduction is the gain of electrons[‡]. An example is the oxidation of magnetite (which consists of 1 Fe²⁺ and 2 Fe³⁺) to hematite:

$$2\text{Fe}_{3}\text{O}_{4} + \frac{1}{2}\text{O}_{2} \rightleftharpoons 3\text{Fe}_{2}\text{O}_{3}$$

The Fe^{2+} in magnetite looses an electron in this reaction and thereby oxidized; conversely oxygen gains an electron and is thereby reduced.

We can divide the elements into *electron donors* and *electron acceptors*; this division is closely related to electronegativity, as you might expect. Electron acceptors are electronegative; electron donors are electropositive. Metals in 0 valence state are electron donors, non-metals in 0 valence state are usually electron acceptors. Some elements, such as carbon and sulfur, can be either electron donors or recep-

[‡] A useful mnemonic to remember this is **LEO** the lion says **GRR**! (Loss Equals Oxidation, Gain Refers to Reduction.) Silly, perhaps, but effective. Try it!

Geochemistry

CHAPTER 7: SOLUTIONS

tors. Oxygen is the most common electron acceptor, hence the term *oxidation*. It is nevertheless important to remember that oxidation and reduction may take place in the absence of oxygen.

A reduced system is one in which the availability of electrons is high, due to an excess of electron donors over electron acceptors. In such a system, metals will be in a low valence state, e.g., Fe²⁺. Conversely, when the availability of electrons is low, due to an abundance of electron acceptors, a system is said to be oxidized. Since it is the most common electron acceptor, the abundance of oxygen usually controls the oxidation state of a system, but this need not be the case.

To predict the equilibrium oxidation state of a system we need a means of characterizing the availability of electrons, and the valence state of elements as a function of that availability. Lowtemperature geochemists and high-temperature geochemists do this in different ways. The former use electrochemical potential while the latter use oxygen fugacity. We will consider both.

3.11.1 Redox in Aqueous Solutions

The simplest form of the chemical equation for the reduction of ferric iron would be:

$$\operatorname{Fe}_{aq}^{3+} + e^{-} \rightleftharpoons \operatorname{Fe}_{aq}^{2+} \qquad 3.101$$

where the subscript *aq* denotes the aqueous species. This form suggests that the energy involved might be most conveniently measured in an electrochemical cell.

The Daniell cell pictured in Figure 3.18 can be used to measure the energy involved in the exchange of electrons between elements, for example, zinc and copper:

$$\operatorname{Zn}_{s} + \operatorname{Cu}_{aq}^{2+} \rightleftharpoons \operatorname{Zn}_{aq}^{2+} + \operatorname{Cu}_{s}$$
 3.102

where the subscript *s* denotes the solid. Such a cell provides a measure of the *relative* preference of Zn and Cu for electrons. In practice, such measurements are made by applying a voltage to the system that is just sufficient to halt the flow of electrons from the zinc plate to the copper one. What is actually measured then is a potential energy, denoted E, and referred to as the *electrode potential*, or simply the *potential* of the reaction.

If we could measure the potential of two separate half-cell reactions:

$$Zn_{s} \rightleftharpoons Zn_{aq}^{2+} + 2e^{-}$$
$$Cu_{s} \rightleftharpoons Cu_{aq}^{2+} + 2e^{-}$$

we could determine the energy gain/loss in the transfer of an electron from an individual element. Unfortunately, such measurements are not possible (nor would these reactions occur in the natural environment: electrons are not given up except to another element or species*). This requires the establishment of an arbitrary reference value. Once such a reference value is established, the potential in-





^{*} Ionization reactions, where free electrons are formed, do occur in nature at very high temperatures. They occur, for example, in stars or other very energetic environments in the universe.

Geochemistry

CHAPTER 7: SOLUTIONS

volved in reactions such as 3.101 can be established.

3.11.1.1 Hydrogen scale potential, E_H

The established convention is to measure potentials in a standard hydrogen electrode cell (at standard temperature and pressure). The cell consists on one side of a platinum plate coated with fine Pt powder that is surrounded by H_2 gas maintained at a partial pressure of 1 atm and immersed in a solution of unit H^+ activity. The other side consists of the electrode and solution under investigation. A potential of 0 is assigned to the half-cell reaction:

$$\frac{1}{2} \mathbf{H}_{2(g)} \rightleftharpoons \mathbf{H}_{aq}^{+} + e^{-} \qquad 3.103$$

TADLE 7.7. $E_{\rm H}$ and pe for some	HALF-CEU	L REACTION
Half Cell Reaction	EH°	pε°
	(V)	
$Li^+ + e^- \rightleftharpoons Li$	-3.05	-51.58
$Ca^{2+} + 2e- \rightleftharpoons Ca$	-2.93	-49.55
$Th^{4+} + 4e- \rightleftharpoons Th$	-1.83	-30.95
U^{+4} +4e ⁻ \rightleftharpoons U	-1.38	-23.34
$Mn^{2+}+2e^{-} \rightleftharpoons Mn$	-1.18	-19.95
$Zn^{2+} + 2e^- \rightleftharpoons Zn$	-0.76	-12.85
$Cr^{3+} + 3e- \rightleftharpoons Cr$	-0.74	-12.51
$\mathrm{Fe}^{2+} + 2\mathrm{e}^{-} \rightleftharpoons \mathrm{Fe}$	-0.44	-7.44
$\mathrm{Eu}^{3+} + \mathrm{e}^{-} \rightleftharpoons \mathrm{Eu}^{2+}$	-0.36	-6.08
$Pb^{2+} + 2e^{-} \rightleftharpoons Pb$	-0.13	-2.2
$CO_{2(g)} + 4H^+ + 4e^- \rightleftharpoons CH_2O^* + 2H_2O$	-0.71	-1.2
$2H^+ + 2e^- \rightleftharpoons H_{2(g)}$	0	0
$N_{2(g)} + 6H^+ + 6e^- \rightleftharpoons 2NH_3$	0.093	1.58
$Cu^{2+} + 2e^{-} \rightleftharpoons Cu$	0.34	5.75
$UO_2^{2+} + 2e^- \rightleftharpoons UO_2$	0.41	6.85
$S + 2e^- \rightleftharpoons S^{2-}$	0.44	7.44
$Cu^+ + e^- \rightleftharpoons Cu$	0.52	8.79
$\mathrm{Fe}^{3+} + \mathrm{e}^{-} \rightleftharpoons \mathrm{Fe}^{2+}$	0.77	13.02
$NO_3^- + 2H^+ + e^- \rightleftharpoons NO_2 + H_2O$	0.80	13.53
$Ag^+ + e^- \rightleftharpoons Ag$	0.80	13.53
$Hg^{2+} + 2e^{-} \rightleftharpoons Hg$	0.85	14.37
$MnO_{2(s)} + 4H^+ + 2e^- \rightleftharpoons Mn^{2+} + 2H_2O$	1.23	20.8
$O_2 + 4H^+ + 4e^- \rightleftharpoons 2H_2O$	1.23	20.8
$MnO_4^- + 8H^+ + 5e^- \rightleftharpoons Mn^{2+} + 4H_2O$	1.51	25.53
$Au^+ + e^- \rightleftharpoons Au$	1.69	28.58
$Ce^{4+} + e^{-} \rightleftharpoons Ce^{3+}$	1.72	29.05
$Pt^+ +e^- \rightleftharpoons Pt$	2.64	44.64

Table 3.3. E_{H}° and $p\epsilon^{\circ}$ for some Half-cell Reactions

*CH₂O refers to carbohydrate, the basic product of photosynthesis.

where the subscript g denotes the gas phase. The potential measured for the entire reaction is then assigned to the half-cell reaction of interest. Thus for example, the potential of the reaction:

$$Zn_{aa}^{2+} + H_{2(g)} \rightleftharpoons Zn_s + 2H^4$$

is –0.763 V. This value is assigned to the reaction:

$$\operatorname{Zn}_{aa}^{2+} + 2e^{-} \rightleftharpoons \operatorname{Zn}_{s} = 3.104$$

and called the hydrogen scale potential, or $E_{H_{\ell}}$ of this reaction. Thus the E_{H} for the reduction of Zn^{+2} to Zn^{0} is -0.763 V. The hydrogen scale potentials of a few half-cell reactions are listed in Table 3.3. The sign convention for $E_{\rm H}$ is that the sign of the potential is positive when the reaction proceeds from left to right (i.e., from reactants to products). Thus if a reaction has positive E_{H} , the metal ion will be reduced by hydrogen gas to the metal. If a reaction has negative E_{H} , the metal will be oxidized to the ion and H⁺ reduced. The standard state potentials (298 K, 0.1 MPa) of more complex reactions can be predicted by algebraic combinations of the reactions and potentials in Table 3.3.

The half-cell reactions in Table 3.3 are arranged in order increasing E° . Thus a species on the product (right) side of a given reaction will reduce (give up electrons to) the species on the reactant side in all reactions listed below it. Thus in the Daniell Cell reaction in Figure 3.18, Zn metal will reduce Cu²⁺ in solution. Zn may thus be said to be a stronger reducing agent than Cu.

Geochemistry

CHAPTER 7: SOLUTIONS

Electrochemical energy is another form of free energy and can be related to the Gibbs Free Energy of reaction as:

$$\Delta G = -zFE \qquad 3.105$$

$$\Delta G^{\circ} = -z \mathcal{F} E^{\circ} \qquad 3.106$$

where *z* is the number of electrons per mole exchanged (e.g., 2 in the reduction of zinc) and is the Faraday constant (F = 96,485 coulombs; 1 joule = 1 volt-coulomb). The free energy of formation of a pure element is 0 (by convention). Thus, the ΔG in a reaction that is opposite one such as 3.104, i.e.:

$$Zn_{(s)} \rightleftharpoons Zn^{2+} + 2e^{-}$$

is the free energy of formation of the ion from the pure element. From equ. 3.105 we can calculate the ΔG for the reduction of zinc as 147.24 kJ/mol. The free energy of formation of Zn²⁺ would be –147.24 kJ/mol. Given the free energy of formation of an ion, we can also use 3.105 to calculate the hydrogen scale potential. Since

$$\Delta G = \Delta G^{o} + RT \ln \prod_{i} a_{i}^{v_{i}}$$
 3.107

we can substitute 3.105 and 3.106 into 3.108 and also write

$$E = E^{\circ} - \frac{RT}{z\mathcal{F}} \ln \prod_{i} a_{i}^{v_{i}}$$
3.108

Equation 3.108 is known as the *Nernst Equation*[‡]. At 298K and 0.1 MPa it reduces to:

$$E = E^{o} - \frac{0.0592}{z} \log \prod_{i} a_{i}^{v_{i}}$$
 3.109

We can deduce the meaning of this relationship from the relationship between ΔG and E in equ. 3.105. At equilibrium ΔG is zero. Thus in equation 3.108, *activities will adjust themselves such that E is 0*.

3.11.1.2 Alternative Representation of Redox State: parepsilon

Consider again the reaction:
$$\operatorname{Fe}_{aq}^{3+} + e^{-} \rightleftharpoons \operatorname{Fe}_{aq}^{2+}$$
 (3.101)

If we were to express the equilibrium constant for this reaction, we would write:

$$\mathbf{K} = \frac{a_{Fe^{2+}}}{a_{Fe^{3+}}a_{e^{-}}}$$

Thus we might find it convenient to define an activity for the electron. For this reason, chemists have defined an analogous parameter to pH, called $p\epsilon$. p ϵ is the negative log of the activity of electrons in solution:

$$p\varepsilon \equiv -\log a_{e^-}$$
 3.111

The log of the equilibrium constant for 3.101 may then be written as:

$$\log K = \log \frac{a_{Fe^{2+}}}{a_{Fe^{3+}}} + p\varepsilon$$

$$p\varepsilon = \log K - \log \frac{a_{Fe^{2+}}}{a_{Fe^{3+}}}$$
3.112

upon rearranging we have:

[‡] Named for Walther Nernst (1864-1941). Nernst was born in Briesau, Prussia (now in Poland) and completed a PhD at the University of Würzburg in 1887. Nernst made many contributions to thermodynamics and kinetics, including an early version of the third law. He was awarded the Nobel Prize in 1920.

Geochemistry

CHAPTER 7: SOLUTIONS

Example 3.11: Calculating the E_H of Net Reactions

We can calculate E_H values for reactions not listed in Table 3.3 by algebraic combinations of the reactions and potentials that are listed. There is, however, a "catch". Let's see how this works.

Calculate the E_H for the reaction:

$$Fe^{3+} + 3e^{-} \rightleftharpoons Fe$$

Answer: This reaction is the algebraic sum of two reactions listed in Table 3.2:

$$Fe^{3+} + e^- \rightleftharpoons Fe^{+2}$$

$$Fe^{2+} + 2e^{-} \rightleftharpoons Fe$$

Since the reactions sum, we might assume that we can simply sum the E_H values to obtain the E_H of the net reaction. Doing so, we obtain an E_H of 0.33 V. However, the true E_H of this reaction is -0.037 V. What have we done wrong?

We have neglected to take into consideration the number of electrons exchanged. In the algebraic combination of E_H values, we need to multiply the E_H for each component reaction by the number of electrons exchanged. We then divide the sum of these values by number of electrons exchanged in the net reaction to obtain the E_H of the net reaction, i.e.,

$$E_{H(net)} = \frac{1}{z_{net}} \sum_{i} z_{i} E_{Hi}$$
 3.110

where the sum is over the component reactions *i*. Looking at equation 3.105, we can see why this is the case. By Hess's Law, the ΔG of the net reaction must be the simple sum of the component reaction ΔG 's, but E_H values are obtained by multiplying ΔG by *z*. Equation 3.107 is derived by combining equ. 3.105 and Hess's Law. Using equ. 3.107, we obtain the correct E_H of -0.037 V.

When the activities of reactants and products are in their standard states, i.e., a = 1, then

$$p\varepsilon^{o} = \frac{1}{z}\log K \qquad 3.113$$

(where z again is the number of electrons exchanged: 1 in reaction 3.101). $p\epsilon^{\circ}$ values are empirically determined and may be found in various tables. Table 3.3 lists values for some of the more important reactions. For any state other than the standard state, $p\epsilon$ is related to the standard state $p\epsilon$ by:

$$p\varepsilon = p\varepsilon^{o} - \log \frac{a_{Fe^{2+}}}{a_{r,^{3+}}}$$
 3.114

 $p\epsilon$ and E_H are related by the following equation:

$$p\varepsilon = \frac{\mathcal{F}E_H}{2.303RT} = \frac{5039E_H}{T}$$
 3.115

(the factor 2.303 arises from the switch from natural log units to base 10 log units).

In defining electron activity and representing it in log units, there is a clear analogy between pɛ and pH. However, the analogy is purely mathematical, and not physical. Natural waters do not contain significant concentrations of free electrons. Also, though a system at equilibrium can have only one value for pɛ, just as it will have only one value of pH, redox equilibrium is often not achieved in natural waters. The pɛ of a natural system is therefore often difficult to determine. *Thus pɛ is a hypothetical unit*, defined for convenience of incorporating a representation of redox state that fits readily into established thermodynamic constructs such as the equilibrium constant. In this sense, equation 3.115 provides a more accurate definition of pɛ than does equ. 3.111.

The greater the p ϵ , the greater the tendency of species to lose their transferable, or valence, electrons. In a qualitative way we can think of the negative of p ϵ as a measure of the availability of electrons. p ϵ can be related in a general way to the relative abundance of electron acceptors. When an electron acceptor, such as oxygen, is abundant relative to the abundance of electron donors, the p ϵ is high and
Geochemistry

CHAPTER 7: SOLUTIONS

electron donors will be in electron-poor valence states, e.g., Mn^{4+} instead of Mn^{2+} . p ϵ , and E_H , are particularly useful concepts when combined with pH to produce diagrams representing the stability fields of various species. We will briefly consider how these are constructed.

3.11.1.3 ре – pH Diagrams

 $p\epsilon - pH$ and $E_H - pH$ diagrams are commonly used tools of aqueous geochemistry, and it is important to become familiar with them. An example, the $p\epsilon$ -pH diagram for iron, is shown in Figure 3.19. $p\epsilon$ pH diagrams look much like phase diagrams, and indeed there are many similarities. There are, however, some important differences. First, labeled regions do not represent conditions of stability for phases, rather they show which *species* will be *predominate* under the $p\epsilon$ -pH conditions within the regions. Indeed, in Fig. 3.19 we consider only a single phase: an aqueous solution. The bounded regions are called *predominance areas*. Second, species are stable beyond their region: boundaries represent the conditions under which the activities of species predominating in two adjoining fields are equal. However, since the plot is logarithmic, activities of species decrease rapidly beyond their predominance areas.

More generally, a pɛ-pH diagram is a type of *activity diagram*, or predominance diagram, in which the region of predominance of a species is represented as a function of activities of two or more species or ratios of species. We will meet variants of such diagrams in later chapters.

Let's now see how Figure 3.19 can be constructed from basic chemical and thermodynamic data. We will consider only a very simple Fe-bearing aqueous solution. Thus our solution contains only species of iron, the dissociation products of water and species formed by reactions between them. Thermodynamics allows us to calculate the predominance region for each species. To draw boundaries on this plot, we will want to obtain equations in the form of $p\epsilon = a + b \times pH$. With an equation in this form, *b* is a slope and *a* is an intercept on a $p\epsilon$ -pH diagram. Hence we will want to write all redox reactions so that they contain e⁻ and all acid-base reactions so that they contain H⁺.

In Fig. 3.18, we are only interested in the region where water is stable. So to begin construction of our diagram, we want to draw boundaries outlining the region of stability of water. The upper limit is the reduction of oxygen to water:

$$\frac{1}{2}O_{2(g)} + 2e^{-} + 2H_{aq}^{+} \rightleftharpoons H_{2}O$$

The equilibrium constant for this reaction is: $K = \frac{a_{H_2O}}{P_{O_2}^{1/2}a_{e^-}^2a_{H^+}^2}$

Expressed in log form: $\log K = \log a_{H_2O} - \frac{1}{2}\log P_{O_2} + 2p\varepsilon + 2pH$

The value of log K is 41.56 (at 25° C and 0.1 MPa). In the standard state, the activity of water and partial pressure of oxygen are 1 so that 3.116 becomes:

$$p\varepsilon = 20.78 - pH \qquad \qquad 3.117$$

Equation 3.117 plots on a $p\epsilon$ —pH diagram as a straight line with a slope of -1 intersecting the vertical axis at 20.78. This is labeled as line ① on Figure 3.19.

Similarly, the lower limit of the stability of water is the reduction of hydrogen:

$$\mathrm{H}_{aq}^{+} + e^{-} \rightleftharpoons \frac{1}{2} \mathrm{H}_{2(g)}$$

Because $\Delta G_r^{\circ} = 0$ and log K = 0 (by convention), we have $p\epsilon = -pH$ for this reaction, i.e., a slope of 1 and intercept of 0. This is labeled as line ② on Figure 3.19. Water is stable between these 2 lines (region shown in gray on Figure 3.19).

Now let's consider the stabilities of a few simple aqueous iron species. One of the more important reactions is the hydrolysis of Fe^{3+} :

3.116

Geochemistry

CHAPTER 7: SOLUTIONS

$$\operatorname{Fe}_{aq}^{3+} + \operatorname{H}_2\operatorname{O} \rightleftharpoons \operatorname{Fe}(\operatorname{OH})_{aq}^{2+} + \operatorname{H}^+$$

The equilibrium constant for this reaction is 0.00631. The equilibrium constant expression is then:

$$\log K = \log \frac{a_{Fe(OH)^{2+}}}{a_{Fe^{3+}}} - pH = -2.2$$

Region boundaries on $p\epsilon$ -pH diagrams represent the conditions under which the activities of two species are equal. When the activities of FeOH⁺² and Fe⁺³ are equal the equation reduces to:

$$-\log K = pH = 2.2$$

Thus this equation defines the boundary between regions of predominance of Fe³⁺ and Fe(OH)²⁺. The reaction is independent of $p\epsilon$ (no oxidation or reduction is involved), and it plots as a straight vertical line pH = 2.2 (line ③ on Figure 3.19). Boundaries between the successive hydrolysis products, e.g., $Fe(OH)^{\circ}_{3}$ and $Fe(OH)^{-}_{4}$ can be similarly drawn as vertical lines at the pH equal to their equilibrium constants, and occur at pH values of 3.5, 7.3, and 8.8. The boundary between Fe²⁺ and Fe(OH)⁻ can be similarly calculated and occurs at a pH of 9.5.

Now consider equilibrium between Fe⁺² and Fe⁺³. The p ϵ° for this reaction is 13.0 (Table 3.3), hence from equation 3.112 we have:

$$p\varepsilon = 13.0 - \log \frac{a_{Fe^{2+}}}{a_{Fe^{3+}}}$$

When the activities are equal, this equation reduces to:

and therefore plots as a horizontal line at $p\epsilon = 13$ that intersects the FeOH⁺² —Fe⁺³ line at an invariant point at pH = 2.2 (line ④ on Figure 3.19).

The equilibrium between Fe^{+2} , and $Fe(OH)^{+2}$ is defined by the reaction:

pε

$$\operatorname{Fe}(\operatorname{OH})_{aq}^{2+} + e^{-} + \operatorname{H}^{+} \rightleftharpoons \operatorname{Fe}_{aq}^{2+} + \operatorname{H}_{2}\operatorname{O}$$

Two things are occurring in this reaction: reduction of ferric to ferrous iron, and reaction of H^+ ions with the OH^- radical to form water. Thus we can treat it as the algebraic sum of the two reactions we just considered:

$$Fe_{aq}^{3+} + e^{-} \rightleftharpoons Fe_{aq}^{2+} \qquad p\epsilon = 13.0$$

$$Fe(OH)_{aq}^{2+} + H^{+} \rightleftharpoons Fe_{aq}^{3+} + H_{2}O \qquad pH = 2.2$$

$$Fe(OH)_{aq}^{2+} + e^{-} + H^{+} \rightleftharpoons Fe_{aq}^{2+} + H_{2}O \qquad p\epsilon + pH = 15.2$$
or:
$$p\epsilon = 15.2 - pH$$



Figure 3.19. $p\epsilon$ —pH diagram showing predominance regions for ferric and ferrous iron and their hydrolysis products in aqueous solution at 25°C and 1 bar.

Geochemistry

CHAPTER 7: SOLUTIONS

Thus this boundary has a slope of -1 and an intercept of 15.2 (line (5) on Figure 3.19). Slopes and intercepts of other reactions may be derived in a similar manner.

Now let's consider some solid phases of iron as well, specifically hematite (Fe_2O_3) and magnetite (Fe_3O_4). First, let's consider the oxidation of magnetite to hematite in the presence of an aqueous solution:

$$2Fe_{3}O_{4}+H_{2}O \rightleftharpoons 3Fe_{2}O_{3}+2H^{+}+2e^{-}$$
 3.119

Assuming unit activity of all phases, the equilibrium constant expression for this reaction is:

$$\log K = -2pH - 2p\varepsilon \qquad 3.120$$

From the free energy of formation of the phases ($\Delta G_f = -742.2 \text{ kJ/mol}$ for hematite, -1015.4 kJ/mol for magnetite, and -237.2 kJ/mol for water) we can calculate ΔG_r using Hess's Law and the equilibrium constant using Equ. 3.86. Doing so, we find log K = -5.77. Rearranging equation 3.120 we have:

$$p\varepsilon = 2.88 - pH$$

The boundary between hematite and magnetite will plot as a line with a slope of -1 and an intercept of 2.88. Above this line, i.e., at higher p ϵ , hematite will be stable, below that magnetite will be stable (Figure 3.20). Thus this line is equivalent to a phase boundary.

Next let's consider the dissolution of magnetite to form Fe²⁺ ions. The relevant reaction is:

 $Fe_3O_4 + 8H^+ + 2e^- \rightleftharpoons 3Fe^{2+} + 4H_2O$

The equilibrium constant for this reaction is 7×10^{29} . Written in log form:

$$\log \mathbf{K} = 3\log a_{Fe^{2+}} + 8pH - 2p\varepsilon = 29.85$$

or: $p\varepsilon = 14.92 - 4pH - \frac{3}{2} \log a_{Fe^{2+1}}$

We have assumed that the activity of water is 1 and that magnetite is pure and therefore that its activity is 1. If we again assume unit activity of Fe^{2+} , the predominance area of magnetite would plot as the line:

$$p\varepsilon = 14.92 - 4pH$$

i.e., a slope of -4 and intercept of 0.58. However, such a high activity of Fe²⁺ would be highly unusual in a natural solution. A more relevant activity for Fe²⁺ would be perhaps 10^{-6} . Adopting this value for the activity of Fe²⁺, we can draw a line corresponding to the equation:

$$p\varepsilon = 23.92 - 4pH$$

This line represents the conditions under which magnetite is in equilibrium with an activity of aqueous Fe^{2+} of 10^{-6} . For any other activity, the line will be shifted, as illustrated in Figure 3.20. For higher concentrations, the magnetite region will expand, for lower concentrations it will contract.

Now consider the equilibrium between hematite and Fe^{2+} . We can describe this with the reaction:

$$\text{Fe}_2\text{O}_3 + 6\text{H}^+ + 2\text{e}^- \rightleftharpoons 2\text{Fe}^{2+} + 3\text{H}_2\text{O}$$



Figure 3.20. Stability regions for magnetite and hematite in equilibrium with an iron-bearing aqueous solution. Thick lines are for a Fe_{aq} activity of 10^{-6} , finer lines for activities of 10^{-4} and 10^{-8} . The latter is dashed.

Geochemistry

CHAPTER 7: SOLUTIONS

The equilibrium constant (which may again be calculated from ΔG_r) for this reaction is 23.79.

Expressed in log form:

 $\log K = 2 \log a_{E_{e^{2+}}} + 6 pH + 2 p\varepsilon = 23.79$

Again using an activity of 10^{-6} for Fe²⁺, we can solve for pe as:

$$p\varepsilon = 11.9 - 3pH - \log a_{Fa^2}$$

For an activity of Fe²⁺ of 10⁻⁶, this is a line with a slope of 3 and an intercept of 17.9. This line represents the conditions under which hematite is in equilibrium with $a_{Fe^{2+}} = 10^{-6}$. Again, for any other activity, the line will be shifted as shown in Figure 3.20.

Finally, equilibrium between hematite and Fe³⁺ may be expressed as:

$$Fe_2O_3 + 6H^+ \rightleftharpoons 2Fe^{3+} + 3H_2O$$

The equilibrium constant expression is:

$$\log K = 2\log a_{Fa^{3+}} + 6\,pH = -3.93$$

For a Fe^{3+} activity of 10^{-6} , this reduces to:

$$pH = 1.34$$

Since the reaction does not involve transfer of electrons, this boundary depends only on pH.

The boundary between predominance of Fe^{3+} and Fe^{2+} is independent of the Fe concentration in solution, and is the same as equ. 3.118 and Fig. 3.18, namely $p\epsilon = 13$.



Figure 3.21. $p\epsilon$ and pH of various waters on and near the surface of the earth. After Garrels and Christ (1965).

Examining this diagram, we see that for realistic dissolved Fe concentrations magnetite can be in equilibrium only with a fairly reduced, neutral to alkaline solution. At pH of about 7 or less, it dissolves and would not be stable in equilibrium with acidic waters unless the Fe concentration were very high. Hematite is stable over a larger range of conditions, and becomes stable over a wider range of pH as pɛ increases. Significant concentrations of the Fe³⁺ ion (> 10⁻⁶ m) will be found only very acidic, oxidizing environments.

Figure 3.21 illustrates the pH and pɛ values that characterize a variety of environments on and near the surface of the Earth. Comparing this figure with pH–pɛ diagrams allows us to predict the species we might expect to find in various environments. For example, Fe^{3+} would be a significant dissolved species only in the acidic, oxidized waters that sometimes occur in mine drainages (the acidity of these waters results from high concentrations of sulfuric acid that is produced by oxidation of sulfides). We would expect to find magnetite precipitating only from reduced seawater or in organic-rich, highly saline waters.

3.11.2 Redox in Magmatic Systems

High temperature geochemists use *oxygen fugacity* to characterize the oxidation state of systems. Thus, equilibrium between magnetite and hematite might be written as:

$$4\mathrm{Fe}_{3}\mathrm{O}_{4} + \mathrm{O}_{2(\mathrm{g})} \rightleftharpoons 6\mathrm{Fe}_{2}\mathrm{O}_{3} \qquad \qquad 3.121$$

Geochemistry

CHAPTER 7: SOLUTIONS

rather than the way we expressed it in equ. 3.119. We note, however, there is negligible molecular oxygen in magmatic systems, and other species are often responsible for transfer of electrons and O²⁻. For example, the equilibrium between magnetite and hematite may be mediated by water:

$$2Fe_{3}O_{4} + H_{2}O_{(g)} \rightleftharpoons 3Fe_{2}O_{3} + H_{2} \qquad \qquad 3.122$$

The above two reactions are thermodynamically equivalent in terms of magnetite oxidation. The first reaction is simpler, of course, and hence preferred, but it may sometimes be necessary to consider the proportions of the actual gas species present.

If we can regard magnetite and hematite as pure phases, then their activities are equal to one and the equilibrium constant for reaction 3.121 is the inverse of the oxygen fugacity:

$$K_{MH} = \frac{1}{f_{O_2}}$$

$$K = e^{-\Delta G_f^o/RT}$$
3.123
3.124

We can rewrite equ. 3.86 as:

and taking the standard state as 1000 K and 1 bar, we can write:

$$-\log \mathbf{K} = \log f_{O_2} = \left(\frac{6\Delta G^o_{f(Fe_2O_3,1000)} - 4\Delta G^o_{f(Fe_3O_4,1000)}}{2.303RT}\right)$$

Thus oxygen fugacity can be calculated directly from the difference in the free energy of formation of magnetite and hematite at the appropriate T and P.

Substituting appropriate values into this equation yields a value for log f_{O_2} of -10.86.

It is important to understand that the oxygen fugacity is fixed at this level (though the exact level at which it is fixed is still disputed because of uncertainties in the thermodynamic data) simply by the equilibrium coexistence of magnetite and hematite. The oxygen fugacity does not depend on the proportion of these minerals. For this reason, it is appropriately called a buffer. To understand how this works, imagine some amount of magnetite, hematite and oxygen present in a magma. If the oxygen fugacity is increased by the addition of oxygen to the system, equilibrium in the reaction in equation 3.133 is driven to the right until the log of the oxygen fugacity returns to a value of -10.86. Only when all magnetite is converted to hematite can the oxygen fugacity rise. A drop in oxygen fugacity would be buffered in exactly the opposite way until all hematite were gone. A number of other buffers can be constructed based on reactions such as:

$$3Fe_2SiO_4 + O_2 \rightleftharpoons 2Fe_3O_4 + 3SiO_2$$
(fayalite) (magnetite) (quartz)

and

$$Fe + \frac{1}{2}O_2 \rightleftharpoons FeO$$

(iron) (wüstite)

These can be used to construct the oxygen buffer curves in Figure 3.21.



3.124

Figure 3.22. Oxygen buffer curves in the system Fe-Si-O at 1 bar. QIF, IW, WM, FMQ, and MH refer to the quartz-iron-fayalite, iron-wüstite, wüstite-magnetite, fayalite-magnetite-quartz and magnetite-hematite buffers respectively.

Geochemistry

CHAPTER 7: SOLUTIONS

References and Suggestions for Further Reading

- Anderson, G. M. and D. A. Crerar. 1993. *Thermodynamics in Geochemistry*. New York: Oxford Univ. Press.
- Broecker, W. S. and V. M. Oversby. 1971. Chemical Equilibria in the Earth. New York: McGraw-Hill.
- Blount, C. W. 1977. Barite solubilities and thermodynamic quantities up to 300° C and 1400 bars. *Am. Mineral.* 62: 942-957.
- Brookins, D. G. 1988. Eh-pH Diagrams for Geochemistry. New York: Springer Verlag.

Davies, C. W. 1962. Ion association, London: Butterworths.

Davies, C. W. 1938. The extent of dissociation of salts in water. VIII. An equation for the mean ionic activity coefficient in water, and a revision of the dissociation constant of some sulfates. *J. Chem. Soc.* 2093-2098.

Debye, P. and E. Hückel. 1923. On the Theory of Electrolytes, Phys. Z. 24: 185-208.

Fletcher, P. 1993. Chemical Thermodynamics for Earth Scientists, Essex: Logman Scientific and Technical.

Garrels, R. M. and C. L. Christ. 1982. Solutions, Minerals, and Equilibria. San Francisco: Freeman Cooper.

- Helgeson, H. C. and D. H. Kirkham. 1974. Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures, Debye-Hückel parameters for activity coefficients and relative partial molar quantities. *Am J. Sci.* 274: 1199-1261.
- Kerrick, D. M. and G. K. Jacobs. 1981. A modified Redlich-Kwong equation for H₂O, CO₂ and H₂O-CO₂ mixtures at elevated pressures and temperatures. *Am. J. Sci.* 281: 735-767.

Morel, F. M. M. and J. G. Hering. 1993. *Principles and Applications of Aquatic Chemistry*. 2nd ed., New York: John Wiley and Sons.

Navrotsky, A. 1994. *Physics and Chemistry of Earth Materials*. Cambridge: Cambridge University Press.

Nordstrom, D. K. and J. L. Munoz, 1986. *Geochemical Thermodynamics*, Palo Alto: Blackwell Scientific Publ.

Saxena, S. K., N. Chaterjee, Y. Fei and G. Shen. 1993. *Thermodynamic Data on Oxides and Silicates*. Berlin: Springer Verlag.

Problems

1. Consider the following minerals:

anhydrite:

bassanite: $CaSO_4 \cdot 1/2H_2O$ (the stuff of which plaster of paris is made)

gyspum: CaSO₄.2H₂O

CaSO₄

a.) If water vapor is the only phase of pure water in the system, how many phases are there in this system and how many components are there?

b.) How many phases are present at invariant points in such a system? How many univariant reactions are possible? Write all univariant reactions, labeling each according the phase that does not participate in the reaction.

2. Consider a system consisting of olivine of variable composition $((Mg,Fe)_2SiO_4)$ and orthopyroxene of variable composition $((Mg,Fe)SiO_3)$. What is the *minimum* number of components needed to describe this system?

3. In section 3.2.1.3, we showed that a system containing H_2O , H_2CO_3 , HCO_3^- , CO_3^{2+} , H^+ , and OH^- could be described in terms of components CO_3^{2+} , H^+ , and OH^- . Find a different set of components that describe the system equally well. Show that each of the species in the system is an algebraic sum of your chosen components.

4. Use the data in Table 2.2 to construct a temperature-pressure phase diagram that showing the stability fields of calcite and aragonite.

Geochemistry

CHAPTER 7: SOLUTIONS

5. Consider the following hypothetical gaseous solution: gases 1 and 2 form an ideal binary solution; at 1000° K, the free energies of formation from the elements are -50kJ/mol for species 1 and -60kJ/mol for species 2.

- a.) Calculate ΔG_{mixing} for the solution at 0.1 increments of X₂. Plot your results.
- b.) Calculate \overline{G} for *ideal* solution at 0.1 increments of X₂. Plot your results.
- c.) Using the method of intercepts, find μ^1 and μ^2 in the solution at $X_2 = 0.2$

6. Using the thermodynamic data in Table 2.2, determine which side of the reaction:

$$2Al(OH)_3 \rightleftharpoons Al_2O_3 + 3H_2O$$

is stable at 600° C and 400 MPa.

7. The following analysis of water is from the Rhine River as it leaves the Swiss Alps:

HCO_3^-	113.5 ppm	SO_4^{2-}	36.0 ppm
Cl-	1.1 ppm	NO_3^-	1.9 ppm
Ca ²⁺	40.7 ppm	Mg^{2+}	7.2 ppm
Na ⁺ 1.4 ppm	K^+	1.2 ppm	

a. Calculate the ionic strength of this water. (Recall that concentrations in ppm are equal to concentrations in mmol kg^{-1} multiplied by formula weight.)

b. Using the Debye-Hückel equation and the data in Table 3.2, calculate the practical activity coefficients for each of these species at 25°C.

8. Seawater has the following composition:

Na ⁺	0.481 M	Cl-	0.560 M
Mg^{2+}	0.0544 M	SO_4^{2-}	0.0283 M
Ca ²⁺	0.0105 M	HCO_3^-	0.00238 M
K^+	0.0105 M		

a. Calculate the ionic strength.

b. Using the Davies equation and the data in Table 3.2, calculate the practical activity coefficients for each of these species at 25°C.

9. Water from Thonon, France has the following composition:

Anions	mg/L	Cations	mg / L
HCO_{3}^{-}	332	Ca ²⁺	103.2
SO_4^{2-5}	14	Mg^{2+}	16.1
NO_3^{-}	14	K^+	1.4
Cl	8.2	Na^+	5.1

a. What is the ionic strength of this water?

b. What are the activity coefficients for HCO_3^- and CO_3^{2-} in this water?

c. Assuming an equilibrium constant for the dissociation of bicarbonate:

$$HCO_3^- \rightleftharpoons H^+ + CO_3^{2-}$$

of 4.68×10^{-11} and a pH of 7.3, what is the equilibrium concentration of CO_3^{2-1} in this water?

10. Given the following analysis of biotite and assuming a *mixing-on-site model* for all sites, calculate the activities of the following components:

a.) $KMg_3Si_3AlO_{10}(OH)_2$ (phlogopite)

Geochemistry

CHAPTER 3: SOLUTIONS

b.)

 $KFe_3^{+2}Si_3AlO_{10}(OH)_2$ (annite)

Site	Ion	Ions per site
Tetrahedral	Si	2.773
	Al	1.228
Octahedral	Al	0.414
	Ti	0.136
	Fe ⁺³	0.085
	Fe ⁺²	1.390
	Mn	0.009
	Mg	0.850
Interlayer	Ca	0.013
•	Na	0.063
	K	0.894
Anion	OH	1.687
	F	0.037

Hint: Check your result by making sure the activity of phlogopite in pure phlogopite is 1.

11. Write the equilibrium constant expression for the reaction:

$$CaCO_{3(s)} + 2H_{(aq)}^{+} + SO_{4}^{2-} + H_2O_{(liq)} \Leftrightarrow CaSO_4 \cdot 2H_2O + CO_{2(g)}$$

assuming the solids are pure crystalline phases and that the gas is ideal.

12. The equilibrium constant for the dissolution of galena:

 $PbS_{solid} + 2H^+ \rightleftharpoons Pb_{aq}^{2+} + H_2S^{aq}$

is 9.12×10^{-7} at 80° C. Using the $\gamma_{Pb^{2+}} = 0.11$ and $\gamma_{H_2S} = 1.77$, calculate the equilibrium concentration of Pb^{2+} in aqueous solution at this temperature and at pH's of 6, 5 and 4. Assume the dissolution of galena is the only source of Pb and H₂S in the solution and that there is no significant dissociation of H₂S. *Hint:* mass balance requires that $[H_2S] = [Pb^{2+}]$.

13. The first and second dissociation constants for phosphoric acid (H₃PO₄) are:

$$K_{1} = \frac{\left[H^{+}\right]\left[H_{2}PO_{4}^{-}\right]}{\left[H_{3}PO_{4}\right]} = 7.52 \times 10^{-3}$$
$$K_{2} = \frac{\left[H^{+}\right]\left[HPO_{4}^{2^{-}}\right]}{\left[H_{2}PO_{4}^{-}\right]} = 6.23 \times 10^{-8}$$

Assuming ideal behavior ($\gamma = 1$), a pH of 3.5, and a third dissociation constant is 0. If you added 1 mole of phosphoric acid to 1 liter water, what will the concentration of H₂PO₄⁻ be?

14. The first dissociation constant for H_2S is $K_1 = 9.1 \times 10^{-3}$. Neglecting the second dissociation and assuming ideality (i.e. activity equals concentration), what is the pH of 1 liter of pure water if you dissolve 0.01 moles of H_2S in it? What fraction of H_2S has dissociated? (*HINT: assume that the concentration of OH is negligible (in other words, no autodissociation of water and use the quadratic equation for your final solution).*

15. Assuming ideal solution behavior for the following:

a.) Show that the boiling point of a substance is increased when another substance is dissolved in it assuming the concentration of the solute in the vapor is small.

Geochemistry

CHAPTER 3: SOLUTIONS

b.) By how much will the boiling point of water be elevated when 10% salt is dissolved in it?

15. Find $\Delta \overline{G}$ for the reaction:

$$Pb^{2+} + Mn \rightleftharpoons Pb + Mn^{2-}$$

Which side of the reaction is favored? (HINT: use the data in Table 3.3)

16. What is the $\Delta \overline{G}$ for the reaction:

$$Cu^{2+} + e^{-} \rightleftharpoons Cu^{+}$$

What is the $p\epsilon^{\circ}$ for this reaction?

17. Consider a stream with a pH of 6.7 and a total dissolved Fe concentration of 1 mg/L. Assume ideal behavior for this problem.

a. If the stream water is in equilibrium with the atmospheric O_2 (partial pressure of 0.2 MPa), what is the p ϵ of the water?

b. Assuming they are the only species of Fe in the water, what are the concentrations of Fe^{3+} and Fe^{2+} . Use the pe you determined in part a.

18. Construct a p ϵ -pH diagram for the following species of sulfur: HSO₄⁻, SO₄²⁻, H₂S, HS⁻, and S²⁻ at 25°C and 1 MPa. The following free energies of formation should provide sufficient information to complete this task.

species	ΔG_{f}^{O}	species	ΔG_{f}^{O}	
S ²⁻ (aq)	+85.81	H ₂ O	-237.19	
HS ⁻ (aq)	+12.09	H^{+}	0	
H_2S_2 (aq)	-27.82	H ₂ (g)	0	
SO_4^{2-} (aq)	-744.54	$O_2(g)$	0	
HSO_4^- (aq)	-755.92	-		

Values are in kJ/mole, standard state is 25°C and 1 MPa. R = 8.314 J/mole-K.

19. Construct a p ϵ -pH diagram for dissolve species of uranium: UO₂²⁺ and U(OH)₅⁻, and the two solid phases UO₂ and U₃O₈ at 25°C and 0.1 MPa. Assume the activity of dissolved uranium is fixed at 10⁻⁶. The following free energies of formation should provide sufficient information to complete this task.

species	ΔG_{f}^{0}
U(OH) ⁻ ₅ (aq)	-1630.80
UO_{2}^{2+} (aq)	-952.53
$UO_{2}(s)$	-1031.86
$U_{3}O_{8}(s)$	-3369.58
H ₂ O	-237.19

Values are in kJ/mole, standard state is 25°C and 0.1 MPa. R = 8.314 J/mole-K.

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

Chapter 4: Applications of Thermodynamics to the Earth

4.1 INTRODUCTION

In the previous 2 chapters, we developed the fundamental thermodynamic relationships and saw how they are applied to geochemical problems. The tools now in our thermodynamic toolbox are sufficient to deal with most of the phenomena we will encounter in the second half of this book. They are not sufficient, however, to deal with all geochemical problems. In this chapter, will add a final few thermodynamic tools. These allow us to deal with non-ideal behavior and exsolution phenomena in solids and silicate liquids. With that, we can use thermodynamics to determine the pressure and temperature at which rock assemblages formed, certainly one of the most useful applications of thermodynamics to geology. Along the way, we will see how thermodynamics is related to one of the most useful tools in petrology: phase diagrams. Finally, we return to the question of non-ideal behavior in electrolyte solutions and examine in more depth the problems of ion association and solvation and how this affects ion activities. Deviations from ideal behavior tend to be greater in solutions of high ionic strength, which includes such geologically important solutions as hydrothermal and ore-forming fluids, saline lake waters, metamorphic fluids, and formation and oil field brines. We briefly examine methods of computing activity coefficients at ionic strengths relevant to such fluids.

4.2 Activities in Non-Ideal Solid Solutions

4.2.1 MATHEMATICAL MODELS OF REAL SOLUTIONS: MARGULES EQUATIONS

Ideal solution models often fail to describe the behavior of real solutions; a good example is water and alcohol, as we saw in Chapter 3. Ideal solutions fail spectacularly when exsolution occurs, such as between oil and vinegar, or between orthoclase and albite, a phenomenon we will discuss in more detail shortly. In non-ideal solutions, even when exsolution does not occur, more complex models are necessary.

Power, or Maclaurin, series are often a convenient means of expressing complex mathematical functions, particularly if the true form of the function is not known, as is often the case. This approach is the basis of Margules[†] equations, a common method of calculating excess free energy. For example, we could express the excess volume as a power series:

$$\overline{V}_{ex} = A + BX_2 + CX_2^2 + DX_2^3 + \dots$$
4.1

where X_2 is the mole fraction of component 2.

Following the work of Thompson (1967), Margules equations are used extensively in geochemistry and mineralogy as models for the behavior of non-ideal solid solutions. It should be emphasized that this approach is completely empirical –true thermodynamic functions are not generally power series. The approach is successful, however, because nearly any function can be *approximated* as a power series. Thus Margules equations are attempts to approximate thermodynamic properties from empirical observations when the true mathematical representation is not known. We will consider two variants of them: the symmetric and asymmetric solution models.

4.2.1.1 The Symmetric Solution Model

In some solutions, a sufficient approximation of thermodynamic functions can often be obtained by using only a second order power series, i.e., in equ. 4.1, D = E = ... = 0. Now in a binary solution, the excess of any thermodynamic function should be entirely a function of mole fraction X_2 (or X_1 , however

⁺ Named for M. Margules, who first used this approach in 1895.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

we wish to express it). Put another way, where $X_2 = 0$, we expect $\overline{V}_{ex} = 0$. From this we can see that the first term in Equ. 4.1, A, must also be 0. Thus equation 4.1 simplifies to:

$$\overline{V}_{ex} = BX_2 + CX_2^2 \tag{4.2}$$

The simplest solution of this type would be one that is symmetric about the midpoint, $X_2 = 0.5$; this is called a *Symmetric Solution*. In essence, symmetry requires that:

$$BX_2 + CX_2^2 = BX_1 + CX_1^2 4.3$$

Substituting $(1 - X_2)$ for X_1 and expanding the right hand side of 4.3, we have:

$$BX_2 + CX_2^2 = B - BX_2 + C - 2CX_2 + CX_2^2$$
4.4

Collecting terms and rearranging:

$$B(2X_2 - 1) = C(1 - 2X_2)$$

$$4.5$$

which reduces to B = -C. Letting $W_V = B$ in equation 4.2, we have: $\overline{V} = W V W V^2 = W V (1 - V) = V V W$

$$V_{ex} = W_V X_2 - W_V X_2^- = W_V X_2 (1 - X_2) = X_1 X_2 W_V$$
4.6

W is known as an *interaction parameter* (recall that non-ideal behavior arises from *interactions* between molecules or atoms), and depends on temperature, pressure, and the nature of the solution, but not on X. Expressions similar to 4.2–4.6 may be written for enthalpy, entropy, and free energy; for example:

$$\overline{G}_{ex} = X_1 X_2 W_G \tag{4.7}$$

$$W_G = W_U + PW_V - TW_S$$

The W_G term may be expressed as:

Since the W_H term can be written as: $W_H = W_U + PW_V$ then 4.8 may also be written: $W_G = W_H - TW_S$

$$W_G = W_H - TW_S \tag{4.8a}$$

4.8

The temperature and pressure dependence of W_G are then

$$\left(\frac{\partial W_G}{\partial T}\right)_P = -W_s \qquad 4.9 \qquad \left(\frac{\partial W_G}{\partial P}\right)_T = W_V \qquad 4.10$$

Regular solutions[‡] are a special case of symmetric solutions where:

$$V_s = 0$$
 and therefore $W_G = W_H$

Regular solutions correspond to the case where $\Delta S_{ex} = 0$, i.e., where $\Delta S_{mixing} = \Delta S_{ideal'}$ and therefore where $W_S = 0$. From equation 4.9, we see that W_G is independent of temperature for regular solutions. Examples of such solutions include electrolytes with a single, uncoupled, anionic or cationic substitution, e.g., CaCl₂—CaBr₂, or solid solutions where there is a single substitution in just one site (e.g., Mg₂SiO₄—Fe₂SiO₄).

Setting equation 4.7 equal to equation 3.57, we have for binary solutions:

$$\overline{G}_{ex} = X_1 X_2 W_G = RT [X_1 \ln \lambda_1 + X_2 \ln \lambda_2]$$

$$4.11$$

For a symmetric solution we have the additional constraint that at $X_2 = X_1$, $\lambda_1 = \lambda_2$. From this relationship it follows that:

$$RT\ln\lambda_i = X_i^2 W_G \tag{4.12}$$

This leads to the relationships:

W

$$\mu_1 = \mu_1^\circ + RT \ln X_1 + X_2^2 W_G$$
4.13

$$\mu_2 = \mu_2^\circ + RT \ln X_2 + X_1^2 W_G$$
4.13a

The symmetric solution model should reduce to Raoult's and Henry's Laws in the pure substance and infinitely dilute solution respectively. We see that as $X_1 \rightarrow 1$ equations 4.13 and 4.13a reduce respectively to:

[‡] The term regular solution is often used to refer to symmetric solutions. In that case, what we termed a regular solution is called a *strictly regular solution*.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

$$\mu_1 = \mu_1^{\circ} + RT \ln X_1 \qquad 4.14$$

$$\mu_2 = \mu_2^\circ + RT \ln X_2 + W_G \quad 4.15$$

Equation 4.14 is Raoult's Law; letting:

$$\mu^* = \mu^\circ + W_c$$

$$W_G = RT \ln h$$

or

then 4.15 is Henry's Law. Thus the interaction parameter can be related to the parameters of Henry's Law, and activity coefficient. In the Margules representation, a solution that is ideal throughout is simply the special case where A = B = C = D = ... = 0.

4.2.1.2 The Asymmetric Solution Model

Many real solutions, for example mineral solutions with asymmetric solvi, are not symmetric. This corresponds to the case where D in equation 4.01 is nonzero; i.e., we must carry the expansion to the third order. It can be shown that in this case the excess free energy in binary solutions is given by: $\overline{G}_{ex} = (W_{G_1}X_2 + W_{G_2}X_1)X_1X_2$ 4.16

(You can satisfy yourself that this may be written as a power-series to the third order of either X_1 or X_2 .) The two coefficients are related to the Henry's Law constants:

$$W_{G_i} = \mu_i^* - \mu_i^\circ = RT \ln h_i$$
 4.17

Activity coefficients are given by:

$$vRT \ln \lambda_i = (2W_{G_j} - W_{G_i})X_j^2 + 2(W_{G_i} - W_{G_i})X_j^3 + 3.18$$

where j=2 when i=1 and visa versa and v is the stoichiometric coefficient. As for the symmetric solution model, the interaction parameters of the asymmetric model can be expressed as the sum of the W_U , W_V , and W_S interaction parameters to account for temperature and pressure dependencies.

The alkali feldspars (NaAlSi₃O₈ –



Figure 4.1. Alkali feldspar solid solution computed at 600° C and 200 MPa (2 kb) using the data of Thompson and Waldbaum (1969). $\Delta G_{real} = \Delta G_{ideal} + \Delta G_{excess}$.



Figure 4.2. ΔG_{real} of alkali feldspar solution computed for a series of temperatures and 200 MPa.

KAlSi₃O₈) are an example of a solid solution exhibiting asymmetric exsolution. Figure 4.1 shows the ΔG_{real} , ΔG_{ideal} , and ΔG_{excess} for the alkali feldspar solid solution computed for 600° C and 200 MPa using the asymmetric solution model of Thompson and Waldbaum (1969). ΔG_{excess} is computed using equation 4.16, ΔG_{ideal} is computed using equation 3.30. Figure 4.2 shows ΔG_{real} computed for a series of tem-

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

peratures. Perhaps a clearer picture of how ΔG will vary as a function of both composition and temperature can be obtained by plotting all 3 variables simultaneously, as in Figure 4.3.

4.3 Exsolution PHENOMENA

Now consider a binary syssuch as NaAlSi₃O₈ tem, KAlSi₃O₈ in the example above, of components 1 and 2, each of which can form a pure phase, but also together form a solution phase, which we will call *c*. The condition for spontaneous exsolution of components 1 and 2 to form two phases *a* and *b* is simply that $G_a + G_b < G_c$.

As we saw in Chapter 3, the free energy of a real solution can be expressed as the sum of an ideal solution a non-ideal or excess free energy term:

$$G_{real} = G_{ideal} + G_{ex}$$

 $W_{GAb} = 10.344$ kJ and $W_{GOr} = 18.938$ kJ.

We can then calculate the activity coef-

ficient as a function of X_{Ab} and X_{Or} from

equ. 4.19. The activity is then computed



Figure 4.3. ΔG surface for the alkali-feldspar solid solution as a function of the mole fraction albite and temperature.





0.8

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

The free energy of the ideal part is given by:

$$\overline{G}_{ideal} = \sum_{i} X_{i} \mu_{i}^{o} + RT \sum_{i} X_{i} \ln X_{1}$$
(3.31)

Further, the ideal part itself consists of two terms, the first term in 3.31 corresponding to the free energy of a mechanical mixture (G_{mixture}), the second part being the free energy of ideal mixing ($\Delta G_{ideal mixing}$). Figure 4.5a illustrates the variation of G_{excess} , $G_{mixture}$, and G_{ideal} in this hypothetical system. G_{mixture} is simply the free energy of a mechanical mixture of pure components 1 and 2 (e.g., orthoclase and albite). Figure 4.5b illustrates the variation of G_{real} in this system So long as G_{real} is less than G_{mixture}, a solution is stable relative to pure phases 1 and 2. You can see that G_{ideal} is always less than G_{mix} $_{ture}$, so as long as the G_{ex} term is not too great. In the hypothetical case illustrated in Fig. 4.5, a solution is always stable relative to a mechanical mixture of the pure end member phases. However, if we look carefully at Fig. 4.5b, we see there is yet another possibility, namely that two phases *a* and *b*, each of which is a *limited* solid solution of components 1 and 2, are stable relative to a single solid solution. Thus at equilibrium, two phases will exsolve from the single solution; this is just what occurs at lower temperatures in the alkali feldspar system. It would be useful if thermodynamics could predict when such exsolution will occur. Let's see if our thermodynamics tools are up to the task.

Looking at Figure 4.2, we see that at 800° C, ΔG_{real} defines a continuously concave upward path, while at lower temperatures, such as 600° C (Figure 4.1), inflections occur and there is a region where ΔG_{real} is concave downward. All this suggests we can use calculus to predict exsolution. For a binary solution of components 1 and 2, the $G_{mixturer}$ and $\Delta G_{ideal mixing}$ terms are:

$$G_{mixture} = X_1 \mu_1^o + X_2 \mu_2^o$$
$$\Delta G_{ideal mixing} = RT(X_1 \ln X_1 + X_2 \ln X_2)$$

Equation 3.31 can thus be written as:

$$G = X_1 \mu_1^o + X_2 \mu_2^o + RT(X_1 \ln X_1 + X_2 \ln X_2) + G_{ex}$$
4.19

Differentiating with respect to X_2 (and recalling that $X_1 = 1 - X_2$), we obtain:

$$\left(\frac{\partial \overline{G}}{\partial X_2}\right) = \mu_2^o - \mu_1^o + RT \ln \frac{X_2}{X_1} + \left(\frac{\partial \overline{G}_{ex}}{\partial X_2}\right) \qquad 4.20$$

This is the equation for the slope of the curve of G vs. X_2 . The second derivative is:



Figure 4.5. (a) Schematic isothermal, isobaric G-X plot for a real solution showing Δ G of mechanical mixing, ideal mixing and excess mixing. (b) Sum of ideal and excess mixing free energies shown in (a). Tangents to the minima give the chemical potentials in immiscible phases *a* and *b*. (c). T-X plot for same system as in (b). Solid line is the *solvus*, dashed line is the *spinodal*. Exsolution may not occur between the spinodal and solvus because the free energy can locally increase during exsolution. After Nordstrom and Munoz (1986).

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

$$\left(\frac{\partial^2 \overline{G}}{\partial X_2^2}\right) = \frac{RT}{X_1 X_2} + \left(\frac{\partial^2 \overline{G}_{ex}}{\partial X_2^2}\right)$$

$$4.21$$

This tells us how the slope of the curve changes with composition. For an ideal solution, G_{excess} is 0, the second derivative is always positive, and the free energy curve is concave upward. But for real solutions G_{excess} can be positive or negative. If for some combination of T and X (and P), the second derivative of G_{excess} becomes negative and its absolute value is greater than the RT/X_1X_2 term, inflection points appear in the G-X curve. Thus exsolution is thermodynamically favored if for some composition:

$$\frac{RT}{X_1X_2} + \left(\frac{\partial^2 \overline{G}_{ex}}{\partial X_2^2}\right) \le 0$$

The inflection points occur where the second derivative is 0, however, as may be seen in Fig. 4.5b, the inflection points do not correspond with the thermodynamic limits of solubility, which in this diagram are between X_2^a and X_2^b .

We can draw a straight line that is tangent to the free energy curve at X_2^a and X_2^b . This line is the free energy of a mechanical mixture of the two limited solutions *a* and *b*. Phase *a* is mostly component 1, but contains X_2^a of component 2. Similarly, phase *b* is mostly component 2 but contains $1 - X_2^b$ of component 1. The mechanical mixture of *a* and *b* has less free energy than a single solution phase everywhere between X_2^a and X_2^b . It is therefore thermodynamically more stable, so exsolution can occur in this region.

In Figure 4.2, we can see inflection points developing at about 650° C in the alkali feldspar solution. The inflection points become more marked and occur at increasingly different values of X_{Ab} as tem-

perature decreases. The alkali feldspar system illustrates a common situation where there is complete solid solution at higher temperature, but decreasing *miscibility* at lower temperature. This occurs because free energy of ideal mixing becomes less negative with decreasing temperature (Figure 3.6).

Figure 4.5c shows a schematic drawing of a temperature–composition plot in which there is complete solution at higher temperature with a widening two-phase region at lower temperatures. The boundary between the two-phase and one-phase regions is shown as a solid line and is known as the *solvus*.

The analysis of exsolution above is relevant to immiscible liquids (e.g., oil and vinegar, silicate and sulfide melts) as well as solids. There is a difference, however. In solids, exsolution must occur through diffusion of atoms through crystal lattices, while in liquids both diffusion and advection serve to redistribute components in the exsolving phases. As exsolution begins, the exsolving phases begin with the composition of the single solution and must rid themselves of unwanted components. In a solid, this only



Figure 4.6. A small portion of a G-X plot illustrating the origin of the spinodal. The process of exsolution of two phases from a single solid solution must overcome an energy barrier. As exsolution from a solution of composition C begins, the two exsolving phases have compositions that move away from C, e.g., A' and B'. But the free energy of a mechanical mixture of A' and B' has greater free energy, by ΔG_{unmix} than the original single solution phase. Exsolution will therefore be inhibited in this region. This problem does not occur if the original solution has composition C'.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

occurs through diffusion, which is very slow. This leads to a kinetic barrier that often prevents exsolution even though 2 exsolved phases are more stable than 1 solution. This is illustrated in Figure 4.6. For example, consider a solution of composition C. It begins to exsolve protophases of A and B, which initially have compositions A' and B'. Even though a mechanical mixture of A and B will have lower free energy than solution phase C, A' and B', the initial products of exsolution, have higher free energy than C. Furthermore, as exsolution proceeds and these phases move toward compositions A and B, this free energy excess becomes larger. Thus exsolution causes a local increase in free energy and therefore cannot occur. This problem is not encountered at composition C' though, because a mixture of the exsolving protophases A" and B" has lower free energy than original solution at C'. Thus the actual limit for exsolution is not tangent points such as B but at inflection points (where $\partial^2 G / \partial X^2 = 0$) such as S. The locus of such points is plotted in Figure 4.5c as the dashed red line and is known as the spinodal.



Figure 4.7. P-T phase diagram for SiO₂. This system has 1 component but 7 phases. L designates liquid, dashed lines indicate where phase boundaries are uncertain. The α — β quartz transition is thought to be partially second order, that is, it involves only stretching and rotation of bonds rather than a complete reformation of bonds as occurs in first order phase transitions.

4.4 Thermodynamics and Phase Diagrams

A *phase diagram* is a representation of the regions of stability of one or more phases as a function of two or more thermodynamic variables such as temperature, pressure, or composition. In other words, if we plot 2 thermodynamic variables such as temperature and pressure or temperature and composition, we can define an area on this plot where a phase of interest is thermodynamically stable. Figure 4.7 is an example of a T-P phase diagram for a 1-component system: SiO₂. The diagram shows the SiO₂ phase stable for a given combination of pressure and temperature. Figure 4.8 is an example of a simple T-X diagram for the two-component system diopside-anorthite (CaMgSi₂O₆ or clinopyroxene and Caplagioclase, CaAl₂Si₂O₈; two of the more common igneous rock forming minerals). In multicomponent systems we must always be concerned with at least 3 thermodynamic variables: P, T, and X. Thus any T-X phase diagram will be valid for only one pressure, 0.1 MPa (1 bar \approx 1 atm) in this case. Of course, with a three dimensional drawing it is possible to represent both temperature and pressure as well as composition in a binary system.

It should not surprise you at this point to hear that the phase relationships in a chemical system are a function of the thermodynamic properties of that system. Thus phase diagrams, such as Figures 4.7 and 4.8, can be constructed from thermodynamic data. Conversely, thermodynamic information can be deduced from phase diagrams.

Let's now see how we can construct phase diagrams, specifically T-X diagrams, from thermodynamic data. Our most important tool in doing so will be the \overline{G} – X diagrams that we have already encountered. The guiding rule in constructing phase diagrams from \overline{G} – X diagrams is that the stable phases are those that combine to give the lowest \overline{G} . Since a \overline{G} – X diagram is valid for only one particular temperature, we will need a number of \overline{G} – X diagrams at different temperatures to construct a single T – X diagram (we could also construct P – X diagrams from a number of \overline{G} – X diagrams at different pressures). Before we begin, we will briefly consider the thermodynamics of melting in simple systems.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

4.4.1 The Thermodynamics of Melting

One of the more common uses of phase diagrams is the illustration of melting relationships in igneous petrology. Let's consider how our thermodynamic tools can be applied to understanding melting relationships. We begin with melting in a simple one component system, for example quartz. At the melting point, this system will consist of two phases: a solid and a melt. At the melting point, the liquid and solid are in chemical equilibrium. Therefore, according to equation 3.17: $\mu_1 = \mu_s$.

The Gibbs Free Energy of melting, $\Delta G_{m\nu}$ must be 0 at the melting point (and only at the melting point). Since

$$\Delta G_m = \Delta H_m - T_m \Delta S_m \qquad 4.22$$

and $\Delta G_m = 0$ at $T_{m\nu}$ then:

$$\Delta H_m = T_m \Delta S_m$$

where ΔH_m is the heat (enthalpy) of melting or fusion^{*}, T_m is the melting temperature, and ΔS_m is the entropy change of melting. Thus the melting temperature of a pure substance is simply:

$$T_m = \frac{\Delta H_m}{\Delta S_m}$$
 4.23

This is a very simple, but very important, relationship. This equation tells us that temperature of melting of a substance is the ratio of the enthaply change to entropy change of melting. Also, if we can measure temperature and enthalpy change of the melting reaction, we can calculate the entropy change.

The pressure dependence of the melting point is given by the Clapeyron Equation:

$$\frac{dT}{dP} = \frac{\Delta V_m}{\Delta S_m} \qquad 4.24$$

Precisely similar relationships hold for vaporization (boiling). Indeed, the temperature and pressure boundaries between any two phases, such as quartz and tridymite, calcite and aragonite, etc., depend on thermodynamic properties in an exactly analogous manner.



Figure 4.8. Phase diagram (T-X) for the two-component system diopside-anorthite at 1 atm. Four combinations of phases are possible as equilibrium assemblages: liquid (L), liquid plus diopside (L + Di), liquid plus anorthite (L + An), and diopside plus anorthite.



Figure 4.9. Computed phase diagram for the system Anorthite-Diopside (CaAl₂Si₂O₈–CaMgSi₂O₆). The eutectic occurs at $X_{Di} = 0.7$ and 1334°C. The dashed lines beyond the eutectic give the apparent melting points of the components in the mixture.

^{*} The heat of fusion is often designated by ΔH_{f} . I have chosen to use the subscript *m* to avoid confusion with heat of formation, for which we have already been using the subscript *f*.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

In equation 3.66 we found that addition of a second component to a pure substance depresses the melting point. Assuming ΔS_m and ΔH_m are independent of temperature, we can express this effect as:

$$\frac{T_{i,m}}{T} = 1 - \frac{R \ln a_{i,m}^{\ell}}{\Delta S_{i,m}}$$

$$4.25$$

Since enthalpies of fusion, rather than entropies, are the quantities measured, equation 4.25 may be more conveniently expressed as:

$$\frac{T_{i,m}}{T} = 1 - \frac{T_{i,m}R\ln a_{i,m}^{\ell}}{\Delta H_{i,m}}$$

$$4.26$$

Example 4.2 shows how the approximate phase diagram for the diopside-anorthite system (Figure 4.9) may be constructed using this equation.

It must be emphasized that in deriving equation 3.66, and hence the equations 4.25 and 4.26, we made the assumption that the solid was a pure phase. This assumption is a reasonably good one for ice, and for anorthite-diopside binary system, but it is not generally valid. Should the solid or solids involved exhibit significant solid solution, this assumption breaks down and these equations are invalid. In that case, melting phase diagrams can still be constructed from thermodynamic equations, but we need to model solid solution as well as the liquid one. Section 4.4.2.1 below illustrates an example (Anorthite-Albite) where the two solutions can be modeled as ideal.

4.4.2 THERMODYNAMICS OF PHASE DIAGRAMS FOR BINARY SYSTEMS

In a one component system, a phase boundary, such as the melting point, is univariant since at that point two phases coexist and f = c-p + 2 = 1 - 2 + 2 = 1. Thus specifying either temperature or pressure fixes the other. A three-phase point, e.g., the triple point of water, is invariant. Hence simply from knowing that three phases of water coexist (i.e., knowing we are at the triple point), we know the tem-

Example 4.2: Calculating Melting Curves

Using the data given below and assuming (1) that the melt is an ideal solution and (2) diopside and anorthite solids are pure phases, calculate a T-X phase diagram for melting of an anorthite-diopside mixture.

Answer: Solving equation 4.26 for T, and replacing activity with mole fraction (since we may assume ideality), we have:

	T _m	ΔH_{m}
	°C	joules/mole
Diopside	1391	138100
Anorthite	1553	136400

(Data from Stebbins et al., 1983)

$$T = \frac{\Delta H_{i,m}}{\Delta H_{i,m}} - R \ln X_i^{\ell}$$

$$4.27$$

We then calculate T for every value of X_{An} and X_{Di} . This produces two curves on a T-X plot, as shown in Figure 4.09. The curves intersect at the *eutectic*, or lowest point at which melt may exist in the system.

Comparing our result with the actual phase relationships determined experimentally (Figure 4.08), we see that while the computed phase diagram is similar to the actual one, our computed eutectic occurs at $X_{Di} = 0.70$ and 1335° C and the a ctual eutectic occurs at $X_{Di} \approx 0.56$ and 1274° C. The difference reflects the failure of the several assumptions we made. First, and most importantly, silicate liquids are not ideal solutions. Second, the entropies and enthalpies of fu sion tend to decrease somewhat with decreasing temperature, violating the assumption we made in deriving equation 4.26. Third, the diopside crystallizing from anorthite-diopside mixtures is not pure, but contains some Al and an excess of Mg.

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

perature and pressure.

In binary systems, the following phase assemblages are possible according to the Gibbs Phase Rule (ignoring for the moment gas phases):

	Phases	Free compositional variables
Univariant	2 solids + liquid, 2 liquids + solid, 3 solids or liquids	0
Divariant	1 solid + 1 liquid, 2 solids, 2 liquids	0
Trivariant	1 solid or 1 liquid	1

When a \overline{G} -X diagram is drawn, it is drawn for a specific temperature and pressure, i.e., \overline{G} -X are isobaric and isothermal. Thus we have already fixed two variables, and the compositions of all phases in univariant and divariant assemblages are fixed by virtue of our having fixed T and P. Only in trivariant systems are we free to choose the composition of a phase on a \overline{G} -X diagram. Figure 4.10 is schematic diagram of a two component, one phase (trivariant) assemblage, in which there is complete solution between component 1 and component 2. This phase might be either a liquid, or a solid such as plagioclase. The composition of the phase may fall anywhere on the curve. Of course, since this diagram applies only to one temperature, we cannot say from this diagram alone that there will be complete solution at all temperatures.

Figure 4.11 illustrates four possible divariant systems. The first case (Figure 4.11a) is that of a liquid solution of composition L' in equilibrium with a solid of fixed composition S_2 (pure component 2). Because the system is divariant, there can be only one possible liquid composition since we have implicitly specified P and T. As usual, the equilibrium condition is described by $\mu_i = \mu_i^s$ (equation 3.17). For i = 2, this means the tangent to the free energy curve for the melt must intersect the $X_2 = 1$ line at μ_2^s as is shown. In other words, the chemical potential of component 2 in the melt must be equal to the chemical potential of component 2 in the solid. Again, this diagram is valid for only one temperature; at any other temperature, the free energy curve for the liquid would be different, but the composition of this new liquid in equilibrium with solid S_2 would still be found by drawing a tangent from S_2 to the free energy curve of the liquid. At sufficiently high temperature, the tangent would always intersect below S_2 . The temperature at which this first occurs is the melting temperature of S_2 (because it is the point at which the free energy of a liquid of pure 2 is less than the solid). The shaded region shows the compositions of systems that will have a combination of solid S_2 and liquid L' as their equilibrium phases as this temperature.

We can also think of the tangent line as defining the free energy of a mechanical mixture of S_2 and L'. In the range of compositions denoted by the shaded region, this mixture has a lower free energy than the liquid solution, hence at equilibrium we expect to find this mixture rather than the liquid solution.

Figure 4.11b illustrates a system with a liquid plus a solid solution, each of which has its own G-X curve. Again, the equilibrium condition is $\mu_i^1 = \mu_i^s$ so the compositions of the coexisting liquid and solid are given by a tangent to both curves. Since the system is divariant and we have fixed P and T the compositions of the solutions are fixed. All system compositions in the shaded region can be accommodated by a mixture of liquid and solid. Compositions lying to the left of the region would have only a liquid; compositions to the right of the shaded region would be accommodated by a solid solution.

Figure 4.11c illustrates the case of two immiscible solids (pure components 1 and 2). The molar free energy of the system is simply that of a mechanical mixture of S_1 and S_2 : a straight line drawn between the free energy points of the two phases.



Figure 4.10. Molar free energy vs. composition $(\overline{G} - X_2)$ for a one -phase assemblage that exhibits complete solution of either a liquid or solid.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS



Figure 4.11. Plot of molar free energy vs. composition $(\overline{G} - X_2)$ for two phase divariant systems. (a) shows a liquid solution (*L*) in equilibrium with a solid (S_2) of pure X_2 . The shaded area shows the range of composition of systems for which *L'* and S_2 . (b) is the case of where both solid and liquid have variable composition. Equilibrium compositions are determined by finding a tangent to both free energy curves. *L'* and *S'* will be the equilibrium phases for systems having compositions in the shaded area. (c) is the case of 2 immiscible solids. (d) shows two limited solid solutions of composition S_1 and S_2 . In the case, the compositions of the solids are given by the point where a straight line is tangent to the curve in two places.

Figure 4.11d illustrates the case of a limited solution. We have chosen to illustrate a solid solution, but the diagram would apply equally well to the case of two liquids of limited solubility.

Figure 4.12a shows the case of two solid solutions plus one liquid. The chemical potential of each component in each phase must be equal to the chemical potential of that component in every other phase, so chemical potentials are given by tangents to all three phases. This is an univariant system, specifying either temperature, pressure, or the composition of a phase fixes other variables in the system. Because of this, if we move to a slightly higher or low temperature at fixed pressure one of the phases must be eliminated in a phase elimination reaction. If the liquid is the liquid is between the two solids in composition, the reaction is known as a eutectic, which is the lowest temperature at which the liquid can exist. Moving to a higher temperature would result in elimination of one of the solids. If, alternatively, the liquid is not between two solids (for example, if the curves L and S2 in Fig. 4.12 were switched), the reaction would be known as a peritectic, and moving to lower temperature eliminates one of the solids. Thus, it is possible for a liquid to persist below a peritectic if the composition is right, but a liquid will never persist at equilibrium below a euctectic. Figure 4.12b is a eutectic in a system where the two solids are the phases of pure components 1 and 2. A line drawn between the free energies of the pure components is also tangent to the liquid curve.

4.4.2.1 An Example of a Simple Binary System with Complete Solution: Albite-

ANORTHITE

Phase diagrams in T-X space can be constructed by analyzing G-X diagrams at a series of temperatures. Let's examine how this can be done in the case of a relatively simple system of two components albite (NaAlSi₃O₈) and anorthite (CaAl₂Si₂O₈) whose solid (plagioclase) and liquid exhibit complete solid solution. Figure 4.13 shows G-X diagrams for various temperatures as well as a T-X phase diagram for this system. Since both the solid and liquid exhibit complete solution, we need to consider G-X curves for both.

We start at the highest point at which liquid and solid coexist, $T_m(T_1)$ for anorthite. Here the solid and liquid curves both have the same value at $X_{An} = 1$; i.e., they are at equilibrium. A G-X plot above

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS



Figure 4.12. Two univariant systems: a liquid plus two solid solutions, and two pure solids and a liquid. Since these systems are univariant, they occur only at one fixed T if P is fixed.

phase. As we move toward Ab (left) in composition, tangents to the solid curve eventually touch the curve for the liquid. The point where the tangent touches each curve gives composition of the liquid and the solid stable at this temperature. In the compositional range between the points where the tangent touches the two curves, the tangent is below both curves, thus a mechanical mixture of solid and liquid is stable over this compositional range at this temperature. For compositions to the left of the point where the tangent touches the liquid curve, the liquid curve is lower than both the solid curve and a tangent to both, so it is stable relative to both the solid and any mixture of solid and liquid.

Going to progressively lower temperatures (e.g., T_3), the points where a tangent intersects the two curves move toward Ab (to the left). Eventually, at a sufficiently low temperature (T_4), the curve for the solid is everywhere below that of the liquid and only solid solution is stable. By extracting information from G-X curves at a number of temperatures, it is possible to reconstruct the phase diagram shown at the bottom of Figure 4.13.

Since both the solid and liquid show complete miscibility in this system, we will make the simplifying assumption that both solutions are ideal and do an approximate mathematical treatment. We recall that the condition for equilibrium was:

$$\mu_i^{\alpha} = \mu_i^{\beta}$$

We can express the chemical potential of each component in each phase as:

this temperature would show the curve for the liquid to be everywhere below that of the solid, indicating the liquid to be the stable phase for all compositions.

At a somewhat lower temperature (T_2) , we see that the curves for the solid and liquid intersect at some intermediate composition. To the right, the curve for the solid is lower than that of the liquid, and tangents to the solid curve extrapolated to both $X_{Ab}=1$ and $X_{An}=1$ are always below the curve for the liquid, indicating the solid is the stable



Figure 4.13. G-X diagrams and a T-X phase diagram for the plagioclase-liquid system. After Richardson and McSween (1987).

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

$$\mu_i^{\alpha} = \mu_i^{o\alpha} + RT \ln X_i^{\alpha}$$

$$4.28$$

Combining these relationships, we have:

$$\mu_{Ab}^{os} - \mu_{Ab}^{o\ell} = RT \ln\left(\frac{X_{Ab}^{\ell}}{X_{Ab}^{s}}\right) 4.29 \quad \text{and} \quad \mu_{An}^{os} - \mu_{An}^{o\ell} = RT \ln\left(\frac{X_{An}^{\ell}}{X_{An}^{s}}\right) 4.30$$

Here our standard states are the pure end members of the melt and solid. The left side of both of these equations corresponds to the standard free energy change of crystallization, thus:

$$\Delta \overline{G}_{m}^{Ab} = RT \ln\left(\frac{X_{Ab}^{\ell}}{X_{Ab}^{s}}\right) \qquad 4.31 \qquad \text{and} \qquad \Delta \overline{G}_{m}^{An} = RT \ln\left(\frac{X_{An}^{\ell}}{X_{An}^{s}}\right) \qquad 4.32$$

Both sides of these equations reduce to 0 if and only if $X_i^{\ell} = X_i^{\ell} = 1$ and $T = T_m$. Rearranging:

$$X_{Ab}^{s} = X_{Ab}^{\ell} e^{\Delta \overline{G}_{m}^{Ab}/RT}$$

$$4.33$$

$$X_{An}^s = X_{An}^{\ell} e^{\Delta \overline{G}_m^{An}/RT}$$

$$4.34$$

Thus the fraction of each component in the melt can be predicted from the composition of the solid and thermodynamic properties of the end members. Since $X_{An}^{\ell} = 1 - X_{Ab}^{\ell}$ and $X_{An}^{s} = 1 - X_{Ab}^{s}$, we can combine equations 4.33 and 4.34 to obtain:

$$\left(1 - X_{Ab}^{\ell}\right)e^{\Delta \bar{G}_{m}^{An}/RT} = 1 - X_{Ab}^{\ell}e^{\Delta \bar{G}_{m}^{Ab}/RT}$$

$$4.35$$

$$X_{Ab}^{\ell} = \frac{1 - e^{\Delta G_m / RT}}{e^{\Delta \bar{G}_m^{Ab} / RT} - e^{\Delta \bar{G}_m^{An} / RT}}$$
 4.36

The point is that the mole fraction of any component of any phase in this system can be predicted from the thermodynamic properties of the end-members. We must bear in mind that we have treated this as an ideal system; i.e., we have ignored any G_{excess} term. Nevertheless, the ideal treatment is relatively successful for the plagioclase system. For non-ideal systems, we merely replace mole fraction in the above equations with activity. Provided they are known, interaction parameters can be used to calculate activity coefficients (e.g., equations 4.18 or 4.12 as the case may be). Beyond that, non-ideal systems can be treated in a manner exactly analogous to the treatment above.

4.5 GEOTHERMOMETRY AND GEODAROMETRY

An important task in geochemistry is estimating the temperature and pressure at which mineral assemblages equilibrate. The importance extends beyond petrology to tectonics and all of geology because it reveals the conditions under which geological processes occur. Here we take a brief look at the thermodynamics underlying geothermometry and geobarometry.

Geothermometry and geobarometry involve two nearly contradictory assumptions. The first is that the mineral assemblage of interest is an equilibrium one, the second is that the system did not reequilibrate during the passage through lower P and T conditions that brought the rock to the surface where it could be collected. As we will see in the next chapter, reaction rates depend exponentially on temperature, hence these assumptions are not quite as contradictory as they might seem. In this section, we will focus only on "chemical" thermobarometers. In Chapter 9, we will see that temperatures can also be deduced from the distribution of isotopes between phases.

4.5.1 Theoretical considerations

In general, geobarometers and geothermometers make use of the pressure and temperature dependence of the equilibrium constant, K. In Section 3.9 we found that $\Delta G^{\circ} = -RT \ln K$. Assuming that ΔC_p and ΔV of the reaction are independent of temperature and pressure, we can write:

$$\Delta G^{o} = \Delta H^{o}_{T,P_{ref}} - T \Delta S^{o}_{T,P_{ref}} + \Delta V^{o}_{T,P_{ref}} (P - P_{ref}) = -RT \ln \mathbf{K}$$

$$4.37$$

Geochemistry

4.39

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

where the standard state of all components is taken as the pure phase at the temperature and pressure of interest, and the enthalpy, entropy and volume changes are for the temperature of interest and a reference pressure (generally 0.1 MPa).

Solving 4.37 for ln K and differentiating the resulting equation with respect to temperature and pressure leads to the following relations:

 $\left(\frac{\partial \ln \mathbf{K}}{\partial P}\right)_{T} = \frac{\Delta V_{T, P_{ref}}^{o}}{RT}$

$$\left(\frac{\partial \ln K}{\partial T}\right)_{P} = \frac{\Delta H^{o}_{T,P_{ref}} + \Delta V^{o}_{T,P_{ref}} \left(P - P_{ref}\right)}{RT^{2}}$$

$$4.38$$

and

These equations provide us with the criteria for reactions that will make good geothermometers and geobarometers. For a good geothermometer, we want the equilibrium constant to depend heavily on *T*, but be approximately independent of *P*. Looking at equation 4.38, we see this means the ΔH term should be as large as possible and the ΔV term as small as possible. A fair amount of effort was devoted to development of a geothermometer based on the exchange of Fe and Mg between olivine and pyroxenes in the late 1960's. The effort was abandoned when it was shown that the ΔH for this reaction was very small. As a rule, a reaction should have a ΔH° of at least 1 kJ to be a useful geothermometer. For a good geobarometer, we want the ΔV term to be as large as possible. Even though the rhodonite ([Mn,Fe,Ca]SiO₃) and pyroxmangite ([Mn,Fe]SiO₃) pairs commonly occur in metamorphic rocks, the reaction rhodonite \rightarrow pyroxmangite does not make a useful geobarometer because the ΔV of reaction is only 0.2 cc/mol. In general, a reaction should have a ΔV of greater than 2 cc/mol if it is to be used for geobarometry.

The following discussion presents a few examples of useful chemical geothermometers and geobarometers (since most reactions are both temperature and pressure dependent, it is perhaps more accurate to use the term "thermobarometer"). It is not an exhaustive treatment, nor should it be inferred

that those examples discussed are in any way superior to other geothermometers and geobarometers. Reviews by Essene (1982, 1989) and Bohlen and Lindsley (1987) summarize a wide range of igneous and metamorphic thermobarometers.

4.5.2 Practical Thermobarometers

4.5.2.1 Univariant Reactions and Displaced Equilibria

We can broadly distinguish 3 main types of thermobarometers. The first is the *univariant reaction*, in which the phases have fixed compositions. They are by far the simplest, and often make good geobarometers as the ΔV of such reactions is often large. Examples include the graphite-diamond transition, any of the SiO₂ transitions (Figure 4.7), and the transformations of Al₂SiO₅, shown in Figure 4.14. While such thermobarometers are simple, their utility for estimating temperature and pressure is limited. This is because exact temperatures and pressures can



Figure 4.14. Phase diagram for Al_2SiO_5 (kyanite-sillimanite-andalusite) as determined by Holdaway (1971).

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

be obtained only if two or more phases coexist, for example, kyanite and andalusite in Figure 4.14. If kyanite and andalusite are both found in a rock, we can determine either temperature or pressure if we can independently determine the other. Where 3 phases, kyanite, sillimanite, and andalusite coexist the system is invariant and P and T are fixed. If only one phase occurs, for example sillimanite, we can only set a range of values for temperature and pressure. Unfortunately, the latter case, where only 1 phase is present, is the most likely situation. It is extremely rare that kyanite, sillimanite, and andalusite occur together.

The term *displaced equilibria* refers to variations in the temperature and pressure of a reaction that results from appreciable solution in one or more



Figure 4.15. Isopleths of Al in orthopyroxene (thin red lines; weight percent) coexisting with forsterite plus and aluminous phase in the CMAS (Ca-Mg-Al-Si) system. After Gasparik (1984).

phases. Thermobarometers based on this phenomenon are more useful than univariant reactions because the assemblage can coexist over a wide range of P and T conditions. In the example shown in Figure 4.15, the boundaries between garnet-bearing, spinel-bearing, and plagioclase-bearing assemblages are curved, or "displaced" as a result of the solubility of Al in enstatite. In addition to the experimental calibration, determination of P and T from displaced equilibria requires (1) careful determination of phase composition and (2) an accurate solution model.

Geobarometers based on the solubility of Al in pyroxenes have been the subject of extensive experimental investigations for the past 25 years. The general principal is illustrated in Figure 4.15, which shows the concentration of Al in orthopyroxene (opx) coexisting with olivine (forsterite) and an aluminous phase, anorthite, spinel, or garnet. The Al content of opx depends almost exclusively on pressure in the presence of anorthite, is essentially independent of pressure in the presence of spinel, and depends on both temperature and pressure in the presence of garnet. Orthopyroxene-garnet equilibrium has proved to be a particularly useful geobarometer.

Garnet is an extremely dense phase. So we might guess that the ΔV of reactions that form it will be comparatively large, and therefore that it is potentially a good geobarometer. The concentration of Al in opx in equilibrium with garnet may be used as a geobarometer if temperature can be independently determined. Although there has been a good deal of subsequent work and refinement of this geobarometer, the underlying thermodynamic principles are perhaps best illustrated by considering the original work of Wood and Banno (1973).

Wood and Banno (1973) considered the following reaction:

$$Mg_{2}Si_{2}O_{6} + MgAl_{2}SiO_{6} \rightleftharpoons Mg_{3}Al_{2}Si_{3}O_{12}$$

$$Opx Solid Solution \rightleftharpoons Pyrope Garnet$$

$$4.40$$

In developing a geobarometer based on this reaction, they had to overcome a number of problems. First, the substitution of Al in orthopyroxene is a coupled substitution. For each atom of Al substituting in the M1 octahedral site, there must be another Al atom substituting for SiO_2 in the tetrahedral site. Second, there was a total lack of thermodynamic data on the MgAl₂SiO₆ phase component. Data was lacking for a good reason: the phase does not exist and cannot be synthesized as a pure phase. Another

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

problem was the apparent non-ideal behavior of the system, which was indicated by orthopyroxenes in Fe- and Ca-bearing systems containing less alumina than in pure MgO systems at the same pressure.

The equilibrium constant for reaction 4.40 is:

$$\mathbf{K} = \frac{a_{Mg_3Al_2Si_3O_{12}}}{a_{Mg_2Si_2O_6}a_{MgAl_2SiO_6}}$$
 4.41

where the activities in the denominator represent the activities of the enstatite and the hypothetical aluminous enstatite phase components in the enstatite solid solution. In the pure MgO system (i.e., no CaO, FeO, MnO, etc.), the numerator, the activity of pyrope, is 1, of course, and we may write:

$$\Delta G^{\circ} = RT \ln(a_{Mg_2Si_2O_6}a_{MgAl_2SiO_6}) = \Delta H^{\circ} - T\Delta S^{\circ} + (P - P_{ref})\Delta V^{\circ}$$

$$4.42$$

(compare equation 4.37). For an ideal case, this may be rewritten as:

$$RT\ln(X_{Mg_2Si_2O_6}X_{MgAl_2SiO_6}) = \Delta H^o - T\Delta S^o + (P - P_{ref})\Delta V^o$$

$$4.43$$

Wood and Banno first estimated thermodynamic parameters (Δ H, Δ S, and Δ V for aluminous pyroxene) from experimental data. They dealt with the non-ideality in two ways. First, they assumed ideal solution behavior at 1 bar and assumed all non-ideality associated with substitution of Al in orthopyroxene at higher pressure could be accounted for in the volume term in 4.42, which they rewrote as:

$$\Delta \bar{V}^{o} = \bar{V}^{o}_{Mg_{3}Al_{2}Si_{3}O_{12}} - \bar{V}^{o\,opx}_{Mg_{2}Si_{2}O_{6}} - \bar{V}^{o\,opx}_{MgAl_{2}SiO_{6}}$$

$$4.44$$

As for non-ideality related to substitution of Ca and Fe in the system, they noted that non-idealities of most silicate systems were of similar size and magnitude and hence the activity coefficients for garnet tend to cancel those for orthopyroxene. Furthermore, the ΔV and ΔH terms are both large and tend to reduce the errors due to non-ideal behavior.

Since equation 4.42 contains temperature as well as pressure terms, it is obvious that the temperature must be known to calculate pressure of equilibration. In the same paper, Wood and Banno (1973) provided the theoretical basis for estimating temperature from the orthopyroxene–clinopyroxene miscibility gap. Thus in a system containing garnet, orthopyroxene and clinopyroxene, both temperature and pressure of equilibration may be estimated from the composition of these phases.

This geobarometer-geothermometer is commonly used to estimate the temperature and pressure (depth) of equilibration of mantle-derived garnet

lherzolite xenoliths. One of the first applications was by Boyd (1973), who calculated P and T for a number of xenoliths in South African kimberlites, and hence reconstructed the geotherm in the mantle under South Africa.

4.5.2.2 Solvus Equilibria

Solvus Equilibria provides a second kind of T, °C thermobarometer. Generally, these make better geothermometers than geobarometers. A good example is the ortho- and clinopyroxene system, illustrated in Figure 4.16. The two-pyroxene solvus has been the subject of particularly intensive experimental and theoretical work because ortho- and clinopyroxene coexist over a wide range of conditions in Mg, Fe-rich rocks of the crust and upper mantle.

One of the inherent thermodynamic difficulties with this type of geothermometer is that since



Figure 4.16. Phase relationships in the system $Mg_2Si_2O_6$ (enstatite) — CaMgSi_2O_6 (diopside) system (after Lindsley, 1983).

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

it involves exsolution, ideal solution models will clearly be very poor approximations. Thus considerable effort has been made to develop solution models for the pyroxenes. Several factors further complicate efforts to use the pyroxene solvus as a thermobarometer. The first is the existence of a third phase, pigeonite (a low-Ca clinopyroxene), at high temperatures and low pressures; the second is that the system is not strictly binary: natural pyroxenes in igneous rocks are solutions of Mg, Ca, and Fe components. The presence of iron is problematic because of the experimental difficulties encountered with Fecontaining systems. These difficulties include the tendency both for iron to dissolve in the walls of commonly used platinum containers and for Fe^{2+} either to oxidize to Fe^{3+} or to reduce to metallic iron, depending on the oxygen fugacity. In addition, other components, particularly Na and Al are often present in the pyroxenes, as we have just seen.

Despite its complexities, the system has been modeled with some success using a symmetric solution model developed by Wood (1987). There are two octahedral sites in both ortho- and clinopyroxenes, generally called M1 and M2. Ca²⁺ occurs only in the M2 site, while Fe and Mg can occupy either site. Ignoring pigeonite and components other than Ca, Mg and Fe, we can treat mixing in the M2 and M1 sites separately. Mixing in the M2 site can be treated as a ternary Mg, Fe, and Ca solution. In a symmetric ternary solution consisting of components A, B, and C, the activities of the components may be calculated from:

$$RT \ln \gamma_A = X_B^2 W_G^{AB} + X_C^2 W_G^{AC} + X_B X_C \left(W_G^{AB} + W_G^{AC} - W_G^{BC} \right)$$

$$4.45$$

where W_{G}^{AB} is the A-B binary interaction parameter, etc. Mixing of Fe and Mg between the M1 and M2 sites was treated as a simple exchange reaction:

$$Fe_{M2} + Mg_{M1} \rightleftharpoons Fe_{M1} + Mg_{M2}$$

with Δ H of 29.27 kJ/mol and Δ S of 12.61 j/mol. Using this approach, Wood calculated the temperature dependence of the solvus in shown in Figure 4.17. The model fits experimental observation reasonably well for the Mg-rich pyroxenes, but significant deviations occur for the Fe-rich pyroxenes.

4.5.2.3 Exchange Reactions

Exchange reaction thermobarometers depend on the exchange of two species between phases. We will consider two examples of these.

The Roeder and Emslie olivine-liquid geothermometer is a rather simple one based on the equilibrium between magma and olivine crystallizing from it. Consider the exchange reaction:

$$MgO_{Ol} + FeO_{liq} \rightleftharpoons MgO_{liq} + FeO_{Ol}$$



Figure 4.17. Comparison of calculated (solid lines) and experimentally observed (red dashed lines) phase relationships between clino- and orthopyroxene shown in the 'pyroxene quadrilateral', a part of the $CaSiO_3$ -MgSiO_3-FeSiO_3 system. Di: diopside, En: enstatite, Hd: hedenbergite, Fs: ferrosilite. Lines show the limit of solid solution at the corresponding temperatures (° C).

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

where *Ol* denotes olivine and *liq* denotes liquid. We can write the equilibrium constant for this reaction as:

$$K_{D} = \frac{X_{FeO}^{Ol} X_{MgO}^{liq}}{X_{FeO}^{liq} X_{MgO}^{Ol}}$$
4.46

Recalling our criteria for a good geothermometer, we can guess that this reaction will meet at least several of these criteria. First, olivine exhibits complete solid solution, so we might guess we can treat it as an ideal solution, which turns out to be a reasonably good assumption. We might also guess that the molar volumes of forsterite and fayalite and of their melts will be similar, meaning the ΔV term, and hence pressure dependence, will be small, which is also true. As it turns out, however, the ΔH term, which is related to the difference in heats of fusion of forsterite and fayalite, is also relatively small, so the exchange reaction itself is a poor geothermometer. However, we can consider two separate reactions here:

$$MgO_{liq} \rightarrow MgO_{Ol}$$
 and $FeO_{liq} \rightarrow FeO_{Ol}$

and we can write two expressions for K_D . This was the approach of Roeder and Emslie (1970), who deduced the following relations from empirical (i.e., experimental) results:

$$\log \frac{X_{MgO}^{Ol}}{X_{MgO}^{liq}} = \frac{3740}{T} - 1.87$$
4.47



Fig. 4.18. Olivine saturation surface constructed by Roeder and Emslie (1970).

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

$$\log \frac{X_{FeO}^{Ol}}{X_{FeO}^{liq}} = \frac{3911}{T} - 2.50$$
 4.48

Example 4.3. Calculating Magma Temperatures Using the Olivine Geothermometer

From the electron microprobe analysis of glass of a mid-ocean ridge basalt and its coexisting olivine microphenocryst, calculate the temperature at which the olivine and liquid equilibrated:

-	
SiQ	50.3
$A l_2 O_3$	14.3
ΣFeO	11.1
MgO	7.2
CaO	11.5
N a ₂ O	2.6
K ₂ O	0.23
MnO	0.20
TiO ₂	1.71
Total	99.02
Mol % Fo in Ol	82

Answer: We will answer this assuming the glass composition represents that of the liquid and using equations 4.47 and 4.48. To use the equations, we will have to convert the analysis of the glass from weight percent to mole fraction.

Let's setup a spreadsheetto do these calculations. First we must deal with the Fe analysis. The analysis reports only iron as FeO. Generally, about 10% of the iron in a basaltic magma will be present as ferric iron (Fe₂O₃), so we will have to assign 10% of the totaliron to Fe₂O₃. To do this, we get the weight percent FeO simply by multiplying the total FeO by 0.9. To get weight percent Fe₂O₃, we multiply total FeO (11.1%) by 0.1, then multiply by the ratio of the molecular weight of Fe₂O₃ to FeO and divide by 2 (since there are 2

	wt%	w/10% ferric	Mol. wt	moles	mol frac.
SiO2	50.3	50.3	60.09	0.8371	0.5265
Al203	14.3	14.3	102	0.1402	0.0882
total FeO	11.1	11.1			
FeO		9.99	71.85	0.1390	0.0875
Fe2O3		1.22	157.7	0.0077	0.0049
MgO	7.8	7.8	40.6	0.1921	0.1208
CaO	11.5	11.5	56.08	0.2051	0.1290
Na2O	2.6	2.6	61.98	0.0419	0.0264
K20	0.23	0.23	94.2	0.0024	0.0015
MnO	0.2	0.2	70.94	0.0028	0.0018
TiO2	1.71	1.71	79.9	0.0214	0.0135
Total	99.74	99.85		1.590	1.000
XMgO-OI					0.82
XFeO-OI					0.18
	TMgO	1384	kelvin	1111	°C
	TFeO	1390	kelvin	1117	°C

Fe atoms per 'molecule').

Now we are ready to cal culate the mole fractions. We'll set up a column with molecular weights and divide each weight percent by the molecular weight to get the num ber of moles per 100 grams. To convert to mole fraction, we divide the number of moles by the sum of the num ber of moles.

Since the mole fraction of Mg in olivine is equal to the mole fraction of forsterite, we need only convert percent to fraction (i.e., divide by 100) The mole fraction of FeO in olivine is simply $1 - X_{MgO}$. Thus $X_{MgO(ol)} = 0.82$ and X_{FeC} calculate 2 temperatures: one

 $_{(o1)} = 0.18$. Now we are ready to calculate temperatures. We can calculate 2 temperatures: one from MgO, and the other from FeO. The temperature based on the FeO exchange is:

$$T_{FeO} = \frac{3911}{\log\left(\frac{X_{FeO}^{Ol}}{X_{FeO}^{liq}}\right) + 2.50} \quad \text{and that based on MgO is:} \quad T_{MgO} = \frac{3740}{\log\left(\frac{X_{MgO}^{Ol}}{X_{MgO}^{liq}}\right) + 1.87}$$

We find that the temperatures of the two methods agree within 6°, which is fairlygood. This in dicates the analyzed olivine probably was in equilibrium with the liquid.

and:

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

These K_D 's are much more temperature dependent than for the combined exchange reaction. Subtracting equation 4.47 from 4.48 yields:

$$\log K_D = \frac{171}{T} - 0.63 \qquad 4.49$$

where K_D is defined as in equation 4.46. Note that these equations have the form of equation 3.95. Roeder and Emslie (1970) used these equations to construct the diagram in Figure 4.18.

The iron-titanium oxide system evaluated by Buddington and Lindsley (1964) was one of the first means of obtaining quantitative estimates of crystallization temperatures of igneous rocks. It is important not only because it is useful over a wide range of temperatures and rock types, but also because it yields oxygen fugacity as well. Figure 4.19 shows the TiO_2 –FeO–Fe₂O₃ (rutile–wüstite–hematite) ternary system. The geothermometer is based on the reaction:

$$yFe_{2}TiO_{4} + (1-y)Fe_{3}O_{4} + \frac{1}{4}O_{2}$$

$$\approx yFeTiO_{3} + \left(\frac{3}{2} - y\right)Fe_{2}O_{3}$$

4.50



Figure 4.19. The TiO_2 -FeO-Fe₂O₃ ternary system. Phases are: FeO: wüstite; Fe₂O₃: hematite; TiO₂: rutile; Fe₂TiO₄: ulvospinel; Fe₂O₄: magnetite; FeTiO₃: ilmenite. The system also includes the FeTi₂O₅—Fe₂TiO₅ solution, which is not shown.

which describes equilibrium between the ulvospinel–magnetite (titanomagnetite) and ilmenite– hematite solid solution series. The equilibrium constant expression may be written as:

$$\mathbf{K} = \frac{a_{FeTiO_3}^{y} a_{Fe_2O_3}^{3/2 - y}}{a_{Fe_2TiO_4}^{y} a_{Fe_3O_4}^{1 - y} f_{O_2}^{1/4}}$$

$$4.51$$

The original Buddington and Lindsley geothermometer was based on empirica observations of compositional dependence on oxygen fugacity and temperature, as shown in Figure 4.20. Having values for the compositions of the titanomagnetite and ilmenite phases, one simply read T and f_{O_2} from the graph. To understand the system from a thermodynamic perspective, it is better to consider the two fundamental reactions occurring separately in this system:

$$\operatorname{Fe}_{3}\operatorname{O}_{4} + \operatorname{Fe}_{1}\operatorname{TiO}_{3} \rightleftharpoons \operatorname{Fe}_{2}\operatorname{TiO}_{4} + \operatorname{Fe}_{2}\operatorname{O}_{3}$$
 4.52

$$4Fe_{3}O_{4} + O_{2} \rightleftharpoons 6Fe_{2}O_{3} \qquad 4.53$$

The first reaction represents a temperature dependent exchange between the titanomagnetite and ulvospinel solutions; the second reaction is the oxidation of magnetite to hematite.

Several investigators have studied the iron-titanium oxides attempting to improve upon the work of Buddington and Lindsley (1964). The approach of Spencer and Lindsley (1981) was to consider two reactions 4.52 and 4.53. They modeled the ilmenite as a binary asymmetric Margules solution and titanomagnetite as a binary asymmetric Margules solution below 800° C and as an ideal binary solution above 800° C. They modeled configurational entropy based ordering of Fe²⁺, Fe³⁺, and Ti⁴⁺ in the ilmenite lattice structure (they assumed Fe³⁺ mixed randomly with Ti⁴⁺ in 'A' sites and Fe³⁺ and Fe²⁺ randomly in 'B' sites). The ΔG of reactions above were written as:

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

$$-\frac{\Delta G^{\circ}}{RT} = \ln\left[\frac{X_{Usp}^{\alpha}(1-X_{Ilm})^{\alpha}}{(1-X_{Usp})^{\alpha}X_{Ilm}^{\alpha}}\right] + \ln\left[\frac{\lambda_{Usp}^{\alpha}\lambda_{Hem}^{\alpha}}{\lambda_{Mt}^{\alpha}\lambda_{Ilm}^{\alpha}}\right]$$

$$4.54$$

and:

$$-\frac{\Delta G}{RT} = \ln\left[\frac{X_{Hem}^{6\alpha}}{X_{Mt}^{4\alpha}}\right] + \ln\left[\frac{\gamma_{Hem}^{6\alpha}}{\gamma_{Mt}^{6\alpha}}\right] - \ln f_{O_2}$$

$$4.55$$

The α parameter is related to the number of sites involved in the exchange; Spencer and Lindsley assumed α was 2 for ilmenite and 1 for titanomagnetite. The excess free energy was expressed in the usual way for an asymmetric solution (equation 4.16):

$$\overline{G}_{ex} = \left(W_{G_1} X_2 + W_{G_2} X_1 \right) X_1 X_2$$

for each solution series. When the effect of pressure is neglected, the free energy interaction parameter expression (equation 4.08) simplifies to:

$$W_G = W_H - TW_S \tag{4.56}$$

Values for W_H and W_S were obtained from least-squares fits to experimental data. The parameters obtained are listed in Table 4.1.

Substituting equations 4.56 and 4.16 into the free energy of solution expression ($\Delta G_{excess} = \Delta G_{ideal} - \Delta G_{real}$), the following equation can be obtained:

$$T(K) = \frac{AW_{H}^{U_{Sp}} - BW_{H}^{M_{t}} - CW_{H}^{II} + DW_{H}^{H_{em}} + \Delta H^{o}}{AW_{S}^{U_{Sp}} - BW_{S}^{M_{t}} - CW_{S}^{II} + DW_{S}^{H_{em}} + \Delta S^{o} - R\ln K^{exch}}$$

$$4.57$$

Oxygen fugacity is determined as:

$$\log f_{O_2} = \log MH + \left(\frac{12\ln(1 - X_{ilm}) - 4\ln(1 - X_{Usp}) +}{\frac{1}{RT} \left[\frac{8X_{Usp}^2(\alpha_{Usp} - 1)W_G^{Usp} + 4X_{Usp}^2(1 - 2X_{Usp})W_G^{Mt} +}{12X_{llm}^2(1 - X_{llm})W_G^{llm} - 6X_{llm}^2(1 - 2X_{llm})W_G^{Hem}} \right] \right) / 2.303 \quad 4.58$$

where:

$$A = 3X_{Usp}^{2} - 4X_{Usp} + 1, B = 3X_{Usp}^{2} - 2X_{Usp}, C = 3X_{Ilm}^{2} - 4X_{Ilm} + 1, D = 3X_{Ilm}^{2} - 2X_{Ilm}$$
$$K^{exch} = \frac{(X_{Usp}X_{Hem}^{2})}{(X_{Mt}X_{Ilm}^{2})}, \Delta H^{o} = 27.799 \text{ kJ/mol}, \Delta S^{o} = 4.1920 \text{ J/K-mol}$$

and MH is the magnetite-hematite buffer: $\log MH = 13.966 - 24634/T$.

We have reviewed just a few of the available thermobarometers in use. These were selected to illustrate the underlying principles. There are, however, many thermobarometers in use by geochemists

Table 4.1. Margules Parameters for Ilmenite and Titanomagnetite Solid Solutions

	Usp (<800 ° C)	Mag (<800° C)	llm	Hem		
W _H (joules)	64835	20798	102374	36818		
W _S (joules)	60.296	19.652	71.095	7.7714		
W _G (>800°C) (joules)	0	0				
ΔS_{Usp}^{o} (joules)	4.192					
ΔH_{Usp}^{o} (joules)	27799					

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

and petrologists. Some of these are listed in Table 4.2.

Example 4.4: Using the Iron-Titanium Oxide Geothermometer

An electron microprobe analysis of oxide phases in an andesite reveals that there is 68 mole percent of ulvospinel in an ulvospinel–magnetite phase and 93.3% of ilmenite in an illmenite-hematite phase. Calculate the temperature and fO_2 at which these phases equilibrated.

Answer: We can use equations 4.57 and 4.58 to answer this question. The data in Table 4.2 are relevant to the binary asymmetric solution model for the system below 800° C. Above 800° C, an ideal solution is assumed for the ulvospinel-magnetite phase, so the interaction parameters for this phase go to 0. But if we don't know the temperature, how do we know which equation to use? We begin by temperature using computing the parameters for less than 800° C. If the temperature computed in this way is greater than 800° C (1073 K), we set the \tilde{W}_{H} and W_{s} for ulvospinel and magnetite to 0 and recompute.

Once we have temperature, we can compute the W_G terms using the relationship $W_G = W_H - TW_s$, bearing in mind that $W_{G_{usp}} = W_{G_{Mt}} = 0$ if the temperature is greater than 800° C. With these values in hand, we can use equation 4.58 to calculate the fO_2 . Our spreadsheet is shown on the right. These data we taken from one of Lindsley's Spencer and (1981)experiments, performed at 938° C and $\log fO_2 = -12.76$. Our calculations are in good agreement with the experimental observation.

		1 1			
	XUsp	XIIm			
	0.68	0.933			
	<u>Λ</u> μ	27799			
		4 1 9 2			
	R R	8 3 1 4			
	Interaction	Parameters			
	WHII	64835	WSU	60 296	
	WHM	20798	WSM	19 652	
	WHI	102374	WSI	71.095	
	WHH	36818	WSH	7,7714	
		00010	Worr		
	A	-0.3328			
	В	0.0272			
	C	-0.12053			
	D	0.745467			
	K	0.010958			
Γ=	(A*WHU-	∟ B*WHM-C*W	HI+D	└ *WHH+∧H)	
	(A*WSU-	B*WSM-C*V	$VSI+D^{3}$	*WSH+ΔS-F	{*ln(K)
	(
Γ ((<800)	1281	К	1008	°C
Γ	(>800)	1205	Κ	932	°C
	WG=WH-1	T*WS			
	WGU	-7829.52	WGI	16695.29	
	WGM	-2885.21	WGH	27452.45	
			МН	-6.47	
		Log <i>f</i> 02 (<800)	-12.58	
		LogfO2 (2)	800)	-12.69	
	1	$=$ $=$ \sim $=$ \sim \sim \sim \sim \sim \sim \sim \sim			1

4.6 Thermodynamic Models of Magmas

Silicate liquids have played an extremely important role in the development of the Earth, as well as other bodies in the solar system. As we shall see, the Earth's crust formed as melts from the mantle rose to the surface and cooled. Thus an understanding of igneous processes is an essential part of earth science. Until the a few decades ago, the primary approaches to igneous petrology were observational and experimental. Results of melting experiments in the laboratory were used to interpret observations on igneous rocks. This approach proved highly successful and is responsible for most of our understanding igneous processes. However, such an approach has inherent limitations: virtually every magma is unique in its composition and crystallization history. Yet the experimental database is limited: it is not practical to subject each igneous rock to melting experiments in the laboratory. Realizing this, igneous petrologists and geochemists turned to thermodynamic models of silicate melts as a tool to interpret their evolution. With a proper 'model' of the interaction of various components in silicate melts and adequate thermodynamic data, it should be possible to predict the equilibrium state of any

Geochemistry

CHAPTER 4 : APPLICATIONS OF THERMODYNAMICS

magma under any given set of conditions. The obstacles in develproper thermodynamic oping models of silicate liquids, however, have been formidable. Because they are stable only at high temperatures, obtaining basic thermodynamic data on silicate liquids is difficult. Furthermore, silicate liquids are very complex solutions, with 8 or more elements present in high enough concentrations to affect the properties of the solution. Nevertheless, sufficient progress has been made on these problems that thermodynamics is now an important tool of igneous petrology.

4.6.1 STRUCTURE OF SILICATE MELTS

As was the case for silicate solids and electrolyte solutions, application of thermodynamics to silicate liquids requires some understanding of the interactions that occur on the atomic level. Thus we will once again have to consider the microscopic viewpoint before developing a useful thermodynamic approach. In this section,





we briefly consider the nature of silicate melts on the atomic level.

 Table 4.2.
 Commonly Used Thermobarometers

Reaction	Туре	Reference					
Garnet=Biotite Fe-Mg	exchange	(Ferry and Spear, 1978)					
$(Fe,Mg)_{3}Al_{2}Si_{3}O_{12} \rightleftharpoons K(Mg,Fe)AlSi_{3}O_{10}(OH)_{2}$							
Plagioclase = Garnet + Kyanite + Quartz	displaced equilibria	(Ghent, 1976; Koziol					
$3(Ca,Na)Al_2Si_2O_8 \rightleftharpoons (Fe,Ca)_3Al_2Si_3O_{12} + 2Al_2SiO_5 + 3A_2SiO_5 + 3A_$	SiO ₂	and Newton, 1988)					
Garnet + Quartz = Plagioclase + Wollastonite	displaced equilibria	(Gasparik, 1984b)					
$(Fe,Ca)_3Al_2Si_3O_{12} + SiO_2 \rightleftharpoons (Ca,Na)Al_2Si_2O_8 + 2CaSiO_3$							
Calcite = Dolomite	solvus equilibria	Goldsmith and Newton (1978)					
$CaCO_3 \rightleftharpoons (Ca,Mg)CO_3$							
Calcite = Aragonite	univariant	(Johannes and Puhan, 1971)					
$CaCO_3 \rightleftharpoons CaCO_3$							
Ilmenite + Al_2SiO_5 = Garnet + Rutile + Quartz	displaced equilibria	(Bohlen et al., 1983)					
3 FeTiO ₃ + Al ₂ SiO ₅ \rightleftharpoons 3 TiO ₂ + (Fe,Ca) ₃ Al ₂ Si ₃ O ₁₂ + Si	O_2						
Hercynite + Quartz = Garnet + Sillimanite	displaced equilibria	(Bohlen et al., 1986)					
$\text{FeAl}_2\text{O}_4 + 5\text{SiO}_2 \rightleftharpoons \text{Fe}_3\text{Al}_2\text{Si}_3\text{O}_{12} + \text{Al}_2\text{SiO}_5$							

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

Most, though not all, of our knowledge of the structure has come from studies of glasses rather than melts. While the thermodynamic properties of silicate liquids and their respective glasses differ, other studies have confirmed the general structural similarities of glasses and liquids. Spectral studies of glasses, which in some respects can be viewed as supercooled liquids, have revealed that silicate liquids have structures rather similar to those of silicate solids. In fact, the principal difference between silicate liquids and solids is the absence of long-range ordering in the former; short range ordering is similar. As in silicate minerals, the primary structural element of silicate liquids is the silicon tetrahedron (see Fig. 1.11), consisting of a silicon atom surrounded by four oxygens. As in silicate minerals, tetrahedra may be linked by a shared oxygen, called a bridging oxygen; not surprisingly, unshared oxygens are termed non-bridging (Figure 4.21a). Unlinked silica tetrahedra, that is, those with no bridging oxygens, are termed monomers, SiO_4^{4-} (Figure 4.21b). Two tetrahedra linked by a single oxygen are termed *dimers* and have the formula $Si_2O_7^{6-}$. Tetrahedra may also be linked by two oxygens to form infinite chains; these have a chemical formula of SiO_3^{2-} . In silicates such as quartz and feldspar, the tetrahedra are all linked into a framework, and all oxygens are shared. All these structural elements can be present in silicate glasses.

The degree to which the silica tetrahedra are linked, or polymerized, in silicate liquids affects chemical and physical properties. The degree of polymerization in turn depends on other cations present. These may be divided into two groups, network formers and network modifiers. Relatively small, highly charged cations such Al³⁺ and Fe³⁺ (more rarely, Ti⁴⁺, P³⁺, B³⁺ as well) often substitute for silicon in tetrahedral sites and, along with Si, are termed network formers. The other common cations of natural silicate liquids, Ca^{2+} , Mg^{2+} , K^+ , Na^+ , and H⁺, are network modifiers. These ions cannot substitute for silicon in tetrahedra and their positive charges can only be balanced by non-bridging oxygens. Addition of these ions disrupts the linkages between silica tetrahedra. Thus as silicate melts become richer in these network modifiers they become progressively depolymerized. This is illustrated in Figure 4.22, which compares the structure of pure silica glass (liquid) and



Figure 4.21. Silicate structures. a: Short range silicate structures in melts resemble those in solids. Individual tetrahedra may be linked by bridging oxygens and linked to 2 silicon atoms. b. Unit in silicate melts may include monomers, with no bridging oxygens, and dimers, where only 1 of 4 oxygens in each tetrahedra are 'bridging'.



Figure 4.22. (a) structure of pure silica glass and (b) a silica-rich glass with additional component ions.

a silica-rich glass (liquid). Melt structure in turn affects the physiochemical properties of the melt. For example, SiO_2 -rich melts tend to have low densities and high viscosities. As ions such as MgO or CaO are added to the melt, viscosity decreases and density increases as the polymer structure is disrupted.

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

4.6.2 MAGMA Solution Models

Advances on several fronts have moved thermodynamic modeling of magmas from an academic curiosity to useful petrological tool. First, spectroscopic (mainly Raman and infrared spectroscopy, both of which are sensitive to atomic and molecular vibrations) studies are revealing the structure of silicate melts, which provides the theoretical basis for thermodynamic models. Second, more sophisticated thermodynamic models more accurately reflect interactions in silicate melts. Third, the thermodynamic database has become more complete and more accurate. Finally, the wide accessibility and power of computers and appropriate programs have made the extensive matrix calculations involved in these models possible. Several factors complicate the task of thermodynamic modeling of magmas. First, magmas are solutions of many components (typically 8 or more). Second, the solids crystallizing from magmas are themselves solutions. Third, magmas crystallize over a substantial temperature range (as much as 400-500° C, more in exceptional cases). Furthermore, crystallization may occur over a range of pressures as a magma ascends through the Earth, and crystallization may be accompanied by melting and assimilation of the surrounding 'country' rock. Despite these complications several models that are sufficiently accurate to be useful to petrologists have been published, most notably those of Ghiorso (Ghiorso et al., 1983; Ghiorso and Sack, 1995) and Nielsen and Dungan (1983). The goal of these models is to describe the phases and their proportions in equilibrium with a magma, and the resulting evolution of liquid composition. In particular the models of Ghiorso and colleagues are applicable to both melting and crystallization. In the section below, we briefly consider the model of Ghiorso.

4.6.2.1 The Regular Solution Model of Ghiorso and Others: "MELTS"

Ghiorso (Ghiorso et al., 1983; Ghiorso, 1987; Ghiorso and Sack, 1995; Ghiroso et al. (2002) noted that silicate liquids have substantial compositional regions in which immiscibility occurs and therefore argued that the simplest model that might be able to describe them is the regular solution model. As we saw earlier in the chapter, regular solution models attempt to describe excess functions with interaction, or Margules, parameters. The Margules equation for excess Gibbs Free Energy for many components is:

$$\bar{G}_{ex} = \frac{1}{2} \sum_{i} \sum_{j, j \neq i} X_i X_j W_G^{i,j}$$
4.59*

and the Gibbs Free Energy is:

$$\overline{G} = \sum X_i \mu_i^o + RT \sum_i X_i \ln X_i + \frac{1}{2} \sum_i \sum_j X_i X_j W_G^{i,j}$$

$$4.60^{\dagger}$$

The chemical potentials of individual components are:

$$\mu_{i} = \mu_{i}^{o} = +RT \ln X_{i} + \sum_{j,j \neq i} X_{j} W_{G}^{i,j} - \frac{1}{2} \sum_{j,j \neq k} \sum_{k,k \neq j} X_{j} X_{k} W_{G}^{j,k}$$

$$4.61$$

and the activity coefficients are:

$$RT\ln\lambda_i = \sum_j X_j W_G^{i,j} - \frac{1}{2} \sum_j \sum_k X_j X_k W_G^{j,k}$$

$$4.62$$

Having chosen a general form for the solution model, the next step is to select the components. For practical reasons, Ghiorso et al. (1983) placed all components on an 8-oxygen basis. Ghiorso and Sack (1995) chose liquid components that were "mineral-like": SiO₂, TiO₂, Al₂O₃, Fe₂O₃, MgCr₂O₄, Fe₂SiO₄, Mg₂SiO₄, CaSiO₃, KAlSiO₄, etc. and H₂O. For components of solid phases, they chose pure end-member

^{*} The $\frac{1}{2}$ term arises because the sum contains both $X_i X_j W_g^{ij}$ and $X_j X_i W_g^{ij}$ terms and $W_g^{ij} = W_g^{ij}$.

⁺ For clarity, we have simplified Ghiorso's equation by neglecting H₂O, which they treated separately.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

phase components (e.g., MgSiO₃ in orthopyroxene). The problem with this approach is that the concentrations of these components varied greatly; for example, the mole fraction of SiO₂ is typically 0.4 in basaltic magmas where as that of Mg₂SiO₄ is typically less than 0.1 and that of KAlSiO₄ is typically less than 0.05. We can see from equation 4.60 that when X_i is small, the contribution of the interaction parameters for this component, $W_G^{i,i}$ to the free energy will also be small. Consequently, in the most recent version of this model, called *pMELTS*, Ghiroso et al. (2002) redefined the liquid components so that their mole fractions were more similar, e.g., SiO₂ \rightarrow Si₄O₈, Na₂SiO₃ \rightarrow NaSi_{0.5}O_{1.5}, etc.

The next task is to find values for the interaction parameters. These can be calculated from solid-liquid equilibria experiments. The principle involved is an extension of that which we used in constructing phase diagrams: when a solid and liquid are in equilibrium, the chemical potential of each component in each phase must be equal. Since thermodynamic properties of the solids involved are available (determined using standard thermodynamics techniques), the thermodynamic properties of the coexisting liquid may be calculated.

The reaction of a solid phase, φ , with the melt can be described with a set of *p* reactions of the form:

$$\varphi_{p} \rightleftharpoons \sum_{i} v_{p,i} c_{i}$$

$$4.63$$

where φ_p is the p^{th} end member component of phase φ , c_i refers to the formula for the i^{th} component in the liquid and $v_{p,i}$ refers to the stoichiometric coefficient of this component. Thus for reaction of olivine with the liquid, we have two versions of 4.63:

$$(Mg_2SiO_4)_{O1} \rightleftharpoons 2MgO_\ell + SiO_{2-\ell}$$
 4.63a

and

$$(Fe_2SiO_4)_{Ol} \rightleftharpoons 2FeO_\ell + SiO_{2-\ell}$$
 4.63b

We can express the Gibbs Free Energy change for each of these reaction as:

$$\Delta \overline{G}_r = \Delta \overline{G}_{\varphi_p}^o + RT \sum_i v_{p,i} \ln a_i^\ell - RT \ln a_{\varphi_p}$$

$$4.64$$

where a_i^{ℓ} is the activity of the oxide component in the liquid and φ_p refers to phase component p in phase φ . $\Delta \overline{G}_r$ is, of course, 0 at equilibrium. For example, for reaction 4.63a above, we have:

$$\Delta \overline{G}_r = 0 = \Delta \overline{G}_{Fo}^o + RT \left[2 \ln a_{MgO}^\ell + \ln a_{SiO_2}^\ell \right] + RT \ln a_{Fo}$$

where the subscript *Fo* refers to the forsterite (Mg_2SiO_4) component in olivine and the superscript ℓ refers to the liquid phase. Expanding the liquid activity term, we have:

$$0 = \Delta \overline{G}_{\varphi_p}^o + RT \sum_i v_{p,i} \ln X_i^\ell + RT \sum_i v_{p,i} \ln \lambda_i^\ell - RT \ln a_{\varphi_p}$$

$$4.65$$

Substituting 4.62 for the activity coefficient term in 4.65 and rearranging to place the "knowns" on the left-hand side, we have:

$$-\Delta \bar{G}_{\varphi_{p}}^{o} + RT \ln a_{\varphi_{p}} - RT \sum_{i} v_{p,i} \ln X_{i}^{\ell} = \sum_{i} v_{p,i} \sum_{j} X_{j} W_{G}^{i,j} - \frac{1}{2} \sum_{j} \sum_{k,k \neq j} X_{k} X_{j} W_{G}^{k,j}$$

$$4.66$$

The quantities on the left-hand side of the equation are terms that can be calculated from the compositions of coexisting solids and liquids and solution models of the solids. The right hand side contains the unknowns. One statement of equation 4.66 can be written for each component in each solid phase at a given temperature and pressure. With enough experiments, values for the interaction parameters can be extracted from the phase relations. Ghiorso et al. (1983) and Ghiorso and Sack (1995) used a statistical technique called least squares[†] to determine the interaction parameters from a large number of

[†] 'Least squares' is a numerical technique that attempts to minimize the square of the difference between calculated and observed value of some parameter. The square is taken to give greater weight to large deviations. Thus least squares techniques yield results where there are relatively few large deviations between the calculated and observed value of the parameter of interest. We discuss this technique further in Chapter 8.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

	SiO ₂	TiO ₂	Al_2O_3	Fe ₂ O ₃	$MgCr_2O_4$	Fe ₂ SiO ₄	Mg_2SiO_4	CaSiO ₃	Na ₂ SiO ₃	$KAlSiO_4$	Ca ₃ (PO ₄)
TiO ₂	26267										
Al_2O_3	39120	-29450									
Fe ₂ O ₃	8110	-84757	-17089								
$MgCr_2O_4$	27886	-72303	-31770	21606							
Fe ₂ SiO ₄	23661	5209	-30509	-179065	-82972						
Mg_2SiO_4	3421	-4178	-32880	-71519	46049	-37257					
CaSiO ₃	-864	-35373	-57918	12077	30705	-12971	-31732				
Na ₂ SiO ₃	-99039	-15416	-130785	-149662	113646	-90534	-41877	-13247			
KAlSiO ₄	-33922	-48095	-25859	57556	75709	23649	22323	-17111	6523		
$Ca_3(PO_4)_2$	613892	25939	52221	-4214	5342	87410	-23209	37070	15572		
H ₂ O	30967	81879	-16098	31406		28874	35634	20375	96938	10374	43451

Table 4.3. Interaction Parameters for the Ghiorso Regular Solution Model

Values are in kJ/mol. From Ghiorso and Sack (1995).

published experimental data. Ghiorso and Sack (1995) also noted that the absence of a phase in an experiment provides thermodynamic information about that phase, i.e., that its free energy must be higher than that of the phases that are present. Their approach made use of this information as well, though discussion of that aspect of their method would take us too far afield. In all, Ghiroso and Sack (1995) used data from 1593 published laboratory experiments. The interaction parameters they determined are listed in Table 4.3. In constructing the *pMELTS* model, Ghiroso et al. (2002) used mineral-liquid equilibrium constraints derived from published results of 2439 different laboratory experiments.

One of the goals of the *pMELTs* was to improve the thermodynamic predictions at higher pressures. Since many melting reactions involve significant volume changes, this required an improved equation of state for the liquid, i.e., an improved description of the relationship between volume and pressure. Ghiroso et al. (2002) chose a third order Birch-Murnaghan equation:

$$P = \frac{3}{2} K \left[\left(\frac{V^{\circ}}{V} \right)^{\frac{1}{3}} - \left(\frac{V^{\circ}}{V} \right)^{\frac{5}{3}} \right] \left\{ 1 - \frac{3}{4} (4 - K') \left[\left(\frac{V^{\circ}}{V} \right)^{\frac{3}{2}} - 1 \right] \right\}$$

$$4.67$$

where *K* is the bulk modulus. New experimental data on density derived from new shock wave experiments and new experimental determinations of silicate liquid density (by suspending olivine crystals in the liquid and observing if they sink or float) were used to constraint the *K*' parameter. A new equation of state for water was also incorporated into pMELTS.

With values for the interaction parameters, the model can then be used to predict the assemblage of solids, their compositions, and the liquid composition that will be present in the system as a function of temperature and pressure. The equilibrium condition for a magma, as for any other system, is the condition where Gibbs Free Energy is at a minimum. Thus the problem becomes finding a composition for the liquid and coexisting solids that minimizes *G* at a particular temperature and pressure. In other words, we want to find values of G_{ℓ} and $G_{\phi l}$, $G_{\phi 2}$, ... $G_{\phi n}$ such that G_{sys} is minimal where:

$$G_{sys} = G_{\ell} + \sum_{\varphi} G_{\varphi}$$
 4.68
Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS



Figure 4.23. SiO_2 concentrations in a melt produced by melting of peridotite at 1 GPa as a function of *F*, the percent fraction of melt in the system. Figure compares the predictions of the earlier version of the MELTS model, the newer version, pMELTS, and experimentally determined composition. From Ghiroso et al. (2002).



Figure 4.24. Compositions of pyroxenes found in lavas from two the Cameroon Line. Diamonds, circles, and triangles are *megacrysts* and likely to have crystallized from these lavas. Stars are pyroxenes in *xenoliths* accidentally incorporated in the lavas. Lines show the compositions of the pyroxenes predicted by *pMELTS* to crystallize from these magmas as they cool and evolve. From Rankenburg et al. (2004).

Inherent in the problem is finding which solids will be in equilibrium with the liquid for a given bulk system composition at specified temperature and pressure. In Ghiorso's approach, an initial guess is made of the state of the system. This is done by taking the liquid composition as equal to the system composition and estimating what phases are likely to be in equilibrium with this liquid. Then G is expanded as a 3term Taylor Series[‡] about that initial point, N', where N' is the composite vector containing the mole fractions describing the compositions of all phases in the system. The second term in the expansion is the matrix of first derivatives of G with respect to n_i , the moles of component *i*, which is simply the matrix of the chemical potentials. A minimum of G occurs where the first derivative

is 0. Thus the second term is set to 0 and solution sought by successive iterations. After each iteration N' is reset to the composition found in the most recent iteration. This approach clearly involves repetitive matrix calculations and would not be practical without a computer, but they can easily be performed on the current generation of computers.

The goal of a thermodynamic model such as MELTS is to predict the both the composition of the melt and composition of coexisting solid phases if temperature, pressure and the composition of the system

can be specified. Thus such a program should be able to predict the composition of the melt generated in a region undergoing melting and how the composition of that melt evolves as it rises and cools nearer the surface. Figure 4.23 compares the predictions of the 1995 and 2002 versions of the model

[‡] A Taylor series expansion of a function f(z) in the vicinity of some point z = a has the form:

$$f(z) = f(a) + \frac{(z-a)}{1!} f'(a) + \frac{(z-a)^2}{2!} f''(a) + \dots$$

where f' and f'' are the first and second derivatives of f with respect to z.

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

with experimentally determined compositions of the liquid produced by melting peridotite at 1 GPa. The agreement between the model and experimental observation is clearly improved in *pMELTS*, but it is also clear the predictions still do not agree perfectly with observations.

Figure 4.24 compares the compositions of clinopyroxene crystals found in basaltic lavas of Cameroon Line volcanoes with the compositions predicted by *pMELTS* to crystallize from these magmas. Diamonds and circles are *megacrysts*, and are likely to have crystallized from the magmas. Stars are pyroxenes in ultamafic xenoliths, which are more likely pieces of mantle accidentally incorporated in the magma. The kink in the lower link reflects the onset of garnet crystallization. The agreement is not perfect but this diagram nevertheless shows the enormous value of the this thermodynamic approach in igneous petrology. In this example, Rankenburg et al. (2004) were able to estimate the pressue and temperature of crystallization as 1400°C and 1.7-2.3 GPa. These pressures correspond to depths greater than the thickness of the crust in this area, hence the authors concluded the pyroxene megacrysts must have crystallized in the mantle. Future refinements of the MELTS will undoubtedly close the gap between predictions and observations and enhance the value of this tool.

The latest version of the model, *pMELTS* runs on UNIX-based computers (including Mac OS X), and is available on World Wide Web at "<u>http://melts.ess.washington.edu/</u>"). This web site also has an on-line *Java* version available.

4.7 REPRISE: THERMODYNAMICS OF ELECTROLYTE SOLUTIONS

We discussed the nature of electrolyte solutions and introduced one approach to dealing with their non-ideality, namely the Debye-Hückel activity coefficients, in Section 3.7. We also noted a number of theoretical weaknesses in the Debye-Hückel approach and that this approach is restricted to fairly dilute solutions (ionic strengths less than 0.1 M). In this section we will return to the problem of electrolyte solutions and examine the causes of non-ideal behavior in high ionic strength solutions in more detail. Before doing so, however, we need to introduce a new variation on our now-familiar thermodynamic parameters, namely mean ionic quantities.

4.7.1 MEAN IONIC QUANTITIES

Consider an aqueous NaCl solution. In Chapter 3 we saw that the thermodynamic properties of a salt are related to those of its component ions by:

$$\Psi_{AB} \equiv \nu_A \Psi_A + \nu_B \Psi_B \tag{3.73}$$

So, for example, the chemical potential of NaCl in solution is:

$$\mu_{NaCl} = \mu_{Na^+} + \mu_{Cl^-}$$

which we can express as:

$$\mu_{NaCl} = \mu_{Na^{+}}^{o} + \mu_{Cl^{-}}^{o} + RT \left(\ln a_{Na^{+}} + \ln a_{Cl^{-}} \right)$$

$$\mu_{NaCl} = \mu_{Na^{+}}^{o} + \mu_{Cl^{-}}^{o} + RT \left(\ln m_{Na^{+}} + \ln m_{Cl^{-}} \right) + RT \left(\ln \gamma_{Na^{+}} + \ln \gamma_{Cl^{-}} \right)$$

$$\mu_{NaCl} = \mu_{Na^{+}}^{o} + \mu_{Cl^{-}}^{o} + RT \left(\ln m_{Na^{+}} + \ln m_{Cl^{-}} \right) + RT \left(\ln \gamma_{Na^{+}} + \ln \gamma_{Cl^{-}} \right)$$

$$\mu_{NaCl} = \mu_{Na^{+}}^{o} + \mu_{Cl^{-}}^{o} + RT \left(\ln m_{Na^{+}} + \ln m_{Cl^{-}} \right) + RT \left(\ln \gamma_{Na^{+}} + \ln \gamma_{Cl^{-}} \right)$$

$$\mu_{NaCl} = \mu_{Na^{+}}^{o} + \mu_{Cl^{-}}^{o} + RT \left(\ln m_{Na^{+}} + \ln m_{Cl^{-}} \right) + RT \left(\ln \gamma_{Na^{+}} + \ln \gamma_{Cl^{-}} \right)$$

or:

Though we can certainly determine the concentrations of Na and Cl in solution, how do we independently determine their activity coefficients? Since we cannot create a pure Na⁺ solution or a pure Cl⁻ one, we cannot say what part of the non-ideality of NaCl solution is due to Na⁺ and what part is due to Cl⁻. The practical solution then is to assign all non-ideality equally to both ions. This leads to the concept of the *mean ion activity coefficient*:

$$\gamma_{\pm} = (\gamma_{Na^{+}} \gamma_{Cl^{-}})^{1/2}$$
 4.70

Thus the mean activity coefficient of a salt is the geometric mean of the activity coefficients of its component ions. Equation 4.69 then becomes:

$$\mu_{NaCl} = \mu_{Na^{+}}^{o} + \mu_{Cl^{-}}^{o} + RT \left(\ln m_{Na^{+}} + \ln m_{Cl^{-}} + \ln \gamma_{\pm}^{2} \right)$$

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

Equation 4.70 is valid for 1:1 salts (i.e., 1 cation for each anion). A general expression for the mean activity coefficient of a salt of composition $A_{\nu+}B_{\nu-}$ is:

$$\gamma_{\pm} = (\gamma_{+}^{\nu^{+}} \gamma_{-}^{\nu^{-}})^{1/\nu}$$
 4.71

where v is the sum of the component positive and negative ions:

$$\mathbf{v} = \mathbf{v}^+ + \mathbf{v}^- \tag{4.72}$$

Example 4.5: Calculating Single Ion Activity Coefficients from Mean Ionic Activity Coeffi-

The measured mean ionic activity coefficient of KCl in a

solution of 1.0 m ionic strength is 0.604; that of $CaCl_2$ in a solution of the same ionic strength is 0.449. What is the

the mean ion activity coefficient for CaCl₂ is related to

Answer: We begin by noting that $\gamma_{CI-} = \gamma_{K^+} = \gamma_{\pm KCI}$ and therefore that $\gamma_{CI-} = 0.604$. According to equ. 4.71,

activity coefficient of Ca²⁺? Assume $\gamma_{CF} = \gamma_{K^+}$.

 $\gamma_{\pm CaCl_2} = \left(\gamma_{Ca^+} \gamma_{Cl^-}^2\right)^{1/3}$

 $\gamma_{Ca^+} = \frac{\gamma_{\pm CaC1_2}^3}{\gamma_{C1^-}^2} = \frac{0.449^3}{0.604^2} = 0.248$

the single ion activity coefficients as:

Solving this for $\gamma_{Ca^{2+}}$ we have:

Mean activity coefficients have the advantage that they are readily measurable (through electrochemical means or solubility, for example). Given a well-behaved salt, such as KCl, where the relationship $\gamma_{-} = \gamma_{+}$ appears to hold, it is then possible to determine single ion activity coefficients. For example, we can obtain $\gamma_{Na^{+}}$ in our NaCl solu-

CIENTS

tion by first determining γ_{CI-} in KCI^{*}:

$$\gamma_{Cl^-} = \gamma_{K^-} = \gamma_{\pm KCl}$$

then determining the mean ion activity coefficient of NaCl experimentally *in a solution of the same ionic strength* and calculating γ_{Na^+} as:

$$\gamma_{Na^{+}} = \frac{\gamma_{\pm NaCl}^{2}}{\gamma_{Cl^{-}}}$$

We can extend the concept of mean ionic quantities to other thermodynamic variables as well. The *mean ionic potential*, μ_{+} , is defined as:

$$\mu_{\pm} = \frac{v^+ \mu_+ v^- \mu^-}{v} \qquad 4.73$$

Thus the mean ionic potential is simply the arithmetic mean of the potential of the individual ions weighted by their stoichiometric coefficients. We could also express the mean ionic potential as:

$$\mu_{\pm} = \mu_{\pm}^{o} + \frac{RT\left(\ln a_{\pm}^{v^{+}} + \ln a_{-}^{v^{-}}\right)}{v}$$

$$4.74$$

Rearranging once more, we obtain:

$$\mu_{\pm} = \mu_{\pm}^{o} + RT \ln \left(a_{+}^{v^{+}} a_{-}^{v^{-}} \right)^{1/v}$$

$$4.75$$

Comparing this relationship with equation 4.69, we define a *mean ionic activity* such that:

$$a_{\pm} = \left(a_{+}^{v^{+}}a_{-}^{v^{-}}\right)^{1/v}$$
 4.76

We can also define mean ionic molalities such that $a_{\pm} = \gamma_{\pm} m_{\pm}$. Substituting $a_{-} = \gamma_{-} m_{-}$, and $a_{+} = \gamma_{+} m_{+}$, we find the *mean ionic molality* is then:

$$m_{\pm} = \left(m_{+}^{\nu^{+}} m_{-}^{\nu^{-}}\right)^{1/\nu}$$
 4.77

Mass balance requires that:

^{*} The use of KCl as a reference for determining mean ion activity coefficients is based on the observation that K⁺ and Cl⁻ have about the same effective radius and ion mobility and is known as the MacInnes Convention. Like that of Debye-Hückel, however, this approach breaks down at high ionic strength.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

 $m^+ = v^+ m$ and $m^- = v^- m$ 4.78 Substituting this into equation 4.77, we see that:

> $m_{\pm} = m \left(v_{+}^{v^{+}} v_{-}^{v^{-}} \right)^{1/v}$ 4.79

Let's return to our NaCl example. Dissociation is essentially complete and v^+ and v^- are unity, so that:

and

$$m_{Na^+} = m_{NaCl}$$

 $m_{Cl^-} = m_{NaCl}$

Since v = 2:

 $m_{\pm NaCl} = \sqrt{m_{NaCl}^2} = m_{NaCl}$ Mean ionic molality is simply equal to molality for a completely dissociated salt consisting of monovalent ions such as NaCl.

The mean ionic activity coefficient, or the stoichiometric activity coefficient as it's sometimes referred to, of NaCl would be the square root of the product of the component activity coefficients according to equation 4.76, as would the mean ionic activity. The individual ion activities can be measured in a number of ways. Therefore, the above relationships allow calculation of the mean ionic activity coefficient from measurable quantities.

For strong electrolytes, i.e., salts that completely dissociate, it can also be shown that mean activity coefficient and mean activity of the salt are related to its activity coefficient and activity by:

$$\begin{array}{l} \gamma = \gamma_{\pm}^{\nu} & 4.80 \\ a = a_{\pm}^{\nu} & 4.81 \end{array}$$

and

We can modify the Debye-Hückel equations to obtain mean ion activity coefficients as follows:

Debye-Hückel Extended Law:

$$\log_{10} \gamma_{\pm} = \frac{-Az_{\pm}|z_{-}|\sqrt{I}}{1 + Ba\sqrt{I}}$$
 4.82

Limiting Law:

$$\log_{10} \gamma_{\pm} = -A z_{+} |z_{-}| \sqrt{I}$$
 4.83

where å is taken as the sum of the radii of the anion and cation, i.e., $a^{\dagger} = a^{\dagger} + a^{\dagger}$.

4.7.1.1 Relationship between Activity and Molality of A SALT

Let's consider the relationship between activity and molality of a salt in an electrolyte solution such as a NaCl solution. Figure 4.25a illustrates this relationship. What we immediately notice is that the slope in the Henry's Law region is essentially

Example 4.6: Mean Ionic Parameters for A fully dissociated electrolyte

If the molality of a CaCl₂ solution is 0.3 M and the activity coefficients of Ca²⁺ and Cl⁻ are 0.5 and 0.7 respectively, calculate the activity and mean ionic molality of CaCl₂ in the solution. Assume that CaCl₂ fully dissociates.

Answer: For CaCl₂, $v^+ = 1$, $v^- = 2$, and v = 3. So we can use equation 4.79 to calculate mean ionic molality:

$$m_{\pm CaCl_2} = m_{CaCl_2} (1^{1}2^2)^{1/3} = 4^{1/4} m_{CaCl_2}$$

Substituting 0.3 for m, we find that $m_{\pm} = 0.4762$ M.

We then use equation 4.71 to calculate the mean ionic activity coefficient:

$$\gamma_{\pm} = (\gamma_{+}^{\nu^{+}} \gamma_{-}^{\nu^{-}})^{1/\nu} = (0.5^{1} 0.7^{2})^{1/3} = 0.625$$

The mean ionic activity is then:

 $a_{\pm} = \gamma_{\pm} m_{\pm} = 0.625 \times 0.4762 = 0.298$ and the activity of CaCl₂ is:

 $a_{CaCl_2} = a_{\pm}^{v} = \gamma_{\pm}m_{\pm} = 0.298^3 = 0.0263M$





Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

zero, which is not at all what we expect for Henry's Law behavior.

It can easily be shown that the relationship in Fig. 4.25a is a simple consequence of the dissociation of the NaCl into Na⁺ and Cl⁻ ions. From 3.75 we have:

$$\mu_{NaCl} = \mu_{Na_{ee}^+} + \mu_{Cl_{ee}^-}$$
 4.84

Substituting this into equation 3.46, we obtain:

 $\mu_{NaCl} = \mu_{Na^+}^o + \mu_{Cl^-}^o + RT \ln a_{Na^+} + RT \ln a_{Cl^-}$

In the reference state of infinitely dilute solution, $m_i = a_i$, so that:

$$\mu_{NaCl} = \mu_{Na^+}^o + \mu_{Cl^-}^o + RT \ln m_{Na^+} + RT \ln m_{Cl^-}$$

$$4.85$$

Furthermore, charge balance requires that:

$$m_{Na^+} = m_{Cl^-} = m_{NaCl} 4.86$$

Substituting 4.86 into 4.85 and rearranging:

$$\mu_{NaCl} = \mu_{Na^+}^o + \mu_{Cl^-}^o + 2RT \ln m_{NaCl} = \mu_{Na^+}^o + \mu_{Cl^-}^o + RT \ln m_{NaCl}^2$$

$$4.87$$

Comparing this equation with equation 3.46, we see that

$$a_{NaCl} \propto m_{NaCl}^2$$

When we plot activity versus the square of molality, we obtain a linear relationship (Fig. 4.25b).

Generalizing this result for dissociation of a substance into a positive ion A and negative ion B, such as:

$$A_{v^+}B_{v^-}v^+A^{v^+}+v^+B^{v}$$

the relationship between activity of a salt and its molality is:

$$a_{AB} \propto m_{AB}^{v}$$
 4.88

For example, v is 3 for $CaCl_2$, 4 for $FeCl_3$, etc.

Now let's see what happens if we substitute the mean ion activity for activity. Since:

$$a_{\pm}^{v} = a_{AB}$$

We have:

 $a_{\pm}^{v} = m_{AB}^{v}$ or $a_{\pm} \propto m_{AB}$ This is the relationship that we observed in Figure 4.25, so we see that the mean ionic activity accounts for the effects of dissociation.

4.7.2 Activities in High Ionic Strength Solutions

The Debye-Hückel equation becomes inaccurate at ionic strengths above about 0.1 m. This is illustrated in Figure 4.26, which shows the experimentally determined mean ion activity coefficient for NaCl as a function of ionic strength and temperature. At low temperatures, the activity begins to increase about ionic strengths of 1 m, whereas Debye-Hückel predicts continual decrease. The activities of many electrolytes eventually exceed 1 at high concentrations. The difference between the observed



Figure 4.26. Observed mean ion activity coefficient, γ_{+} , of NaCl as a function of ionic strength and temperature (solid lines; data from Helgeson, 1981) compared with value predicted by the Debye-Hückel Law (computed as $(\gamma_{Na^+}\gamma_{Cl^-})^{1/2}$).

Geochemistry

CHAPTER 4 : APPLICATIONS OF THERMODYNAMICS

activity coefficients and those predicted by the Debye-Hückel equation are due to the effects of ion association and solvation. Debye and Hückel explicitly assumed complete dissociation, i.e., no ion associations, and while their treatment included in a general way the dielectric properties of water, it neglected the effects of solvation. As we noted in Chapter 3, the effects of both ion association and solvation become increasingly important with increasing ionic strength. It should be no surprise then that the Debye-Hückel treatment breaks down at high ionic strength. Here we will consider these effects in greater detail.

4.7.2.1 Correction for the Concentration of Water

At low and moderate ionic strength, we can assume that the mole fraction of water in solution is 1. For example in seawater, with an ionic strength of 0.7, the mole fraction of water about 0.99. Generally, activity coefficients and equilibrium constants are not known within 1%, so the error introduced by this assumption is still small compared to other errors. In higher ionic strengths, however, this assumption is increasingly invalid (for example, at a molality of 3, the mole fraction of water has decreased below 0.95), and this must be taken into account. A convenient way to do this is to incorporate it into the activity coefficient. The corrected activity coefficient is:

$$\gamma_{corr} = \frac{\gamma}{\left(1 + 0.018 \sum_{i} m_{i}\right)}$$

$$4.89$$

4.7.2.2 Effects of Solvation

Water molecules bound to ions in solvation shells have lost their independent translational motion and move with the ion as a single entity. These water molecules are effectively unavailable for reaction, hence solvation has the effect of reducing the activity of water, which increases the apparent concentration, or activity, of the solutes. In addition to solvation, i.e., the direct association of some water molecules with the ion, the charge of the ion causes collapse of the water structure beyond the solvation shell.

For a solution consisting of a single salt, Robinson and Stokes (1959) proposed the contribution of solvation to the mean ion activity coefficient could be expressed as:

$$\log \gamma_{\pm}^{solv} = -\frac{h}{v} \log a_{w} - \log(1 - 0.018hm)$$
4.90

where $\gamma_{\pm}^{\text{solv}}$ is the solvation contribution to the mean ion activity coefficient, *h* is the number of moles of water molecules bound to each mole of salt, a_w is the activity of water, *m* is the concentration of the salt in solution, and v is a defined in equation 4.72 (i.e., total moles of ions produced upon dissolution of a mole of salt). Table 4.3 listed estimated values for the solvation number, i.e., number of water molecules in the solvation shell of each ion. From these, the value of *h* for equ. 4.90 can be calculated. The activity of water can be adequately estimated as:

$$a_w = 1 - 0.04m$$

Figure 4.27 illustrates the effect of solvation on the activity coefficient. As may be seen, solvation substantially affects the activity coefficient at ionic strengths above about 0.5 m.

4.7.2.2 Effects of Ion Association

An ion pair can be considered to have formed when ions approach closer than some critical distance, r_{cr} where the electrostatic energy, which tends to bind them, exceeds twice the thermal energy, which tends to

Species	h	Species	h
Li ⁺ 2.3	OH⁻	7.6	
Na ⁺	3.3	F^-	6.7
K^+	2.3	Cl⁻	2.7
Rb^+	2.3	Br-	1.7
Mg^{2+}	8.9	CO_{3}^{2-}	14.4
Ca ²⁺	8.9	SO_4^{2-}	10.4
Cd^{2+}	6.3	7	
Ba ²⁺	9.2		

 Table 4.3. Ion Solvation Numbers

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

move them apart. When this happens, the ions are electrostatically bound and their motions are linked. They are said to form an ion pair. The thermal energy of an ion is kT and electrostatic interaction energy is:

$$U_{electro.} = \frac{q_1 q_2}{4\pi\varepsilon r} \qquad 4.91$$

The ratio of these two energies when the distance is less than the critical one is then: γ_{\pm}

$$\frac{U_{electro}}{U_{therm}} = \frac{z_1 z_2 e^2}{4\pi\varepsilon_0 \varepsilon_r rT}$$
 4.92

We can use this equation to solve for the critical distance r_c :

$$r_c = \frac{z_1 z_2 e^2}{8\pi\varepsilon_0 \varepsilon_r T}$$
 4.93

For two singly charged ions, the critical distance is 3.57 Å. In a 1 molar solution, the average separation between ions is about 12 Å, so even in such a relatively concentrated solution, ion pairs will not form between singly charged ions. Indeed, the critical distance is smaller than the combined Debye-Hückel radii of all pairs of singly charged ions. Thus we do not expect ion



Figure 4.27. Comparison of the electrostatic contribution to the mean ion activity coefficient of NaCl (calculated by the Debye-Hückel Extended Law), the solvation contribution (calculated from equation 4.92 assuming h = 4) and the sum of the two.

associations to form from pairs of singly charged ions under most circumstances. In contrast, the critical distance for ion association between a singly and a doubly charged ion is 70 Å, considerably greater than the sum of their Debye-Hückel radii. It also exceeds the average separation of ions in a 0.01 m solution (about 55 Å), so that even in dilution solutions, we would expect significant ion pair formation for multiply charged ions.

As we saw earlier, all ions in solution are surrounded by a solvation shell of water molecules. This solvation shell may or may not be disrupted when ion pair formation occurs (Fig. 4.28). If it is not disrupted, and the two solvation shells remain intact, an *outer sphere ion pair* (also called an outer sphere complex) is said to have formed. If water molecules are excluded from the space between the ions, an *inner sphere ion pair* (or complex) is said to have formed.

For some purposes, ion pairs can be treated as distinct species having charge equal to the algebraic sum of the charge of the ions involved. These can be included, for example, in calculation of ionic strength to obtain a somewhat more accurate estimate of activities. On the other hand, ion pairs, including neutral ones, can be highly dipolar and may behave as charge-separated ions.

Ion associations affect activities in two ways. First, associated ions are less likely to participate in reactions, thus reducing the activity of the ions involved. Second, ion association reduces the ionic strength of the solution, and hence reduces the extent of electrostatic interactions among ions. This has the effect of increasing activity. To understand the first effect, consider the case where a certain fraction of the free ions re-associates to form ion pairs, e.g.:

$$v^+ A^{z+} + v^- B^{z-} \rightleftharpoons (A_{v^+} B_{v^-})^0_{aq}$$

where the superscript 0 indicates neutrality and the subscript aq a dissolved aqueous species. A salt that only partially dissociates in solution is called a weak electrolyte. Let α be the fraction of the ions that associate to form ion pairs or complexes. The associate of this fraction of ions as ion pairs will be thermodynamically equivalent to that fraction of the substance not dissociating to begin with. The fraction of free ions is then $1 - \alpha$. Equation 4.78 becomes:

$$m^+ = (1 - \alpha)v^+m$$
 and $m^- = (1 - \alpha)v^-m$ 4.94

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS



Solvation shells intact

Partial disruption of solvation shells

Disruption of solvation shells

Figure 4.28. In formation of ion pairs, the solvation shells may remain intact or be partially or totally disrupted. The former results in an outer sphere ion pair, the latter results in an inner sphere ion pair.

where m is the molality of the solute. We can rewrite equation 4.76 as:

$$a_{\pm} = \left[\left(\gamma_{+} m_{+} \right)^{\nu^{+}} \left(\gamma_{-} m_{-} \right)^{\nu^{-}} \right]^{1/\nu}$$

$$4.95$$

Substituting 4.94 into 4.95 and rearranging, we obtain:

$$a_{\pm} = \left(\gamma_{+}^{v^{+}} \gamma_{-}^{v^{-}}\right)^{1/v} \left\{ \left[(1-\alpha)v^{+}m \right]^{v^{+}} \left[(1-\alpha)v^{-}m \right]^{v^{-}} \right\}^{1/v}$$

A little more rearranging and we have:

$$a_{\pm} = \left(\gamma_{+}^{v^{+}} \gamma_{-}^{v^{-}}\right)^{1/\nu} \left\{ \left[\left(1 - \alpha\right) m \right]^{(v^{+} + v^{-})} \left(v^{+}\right)^{v^{+}} \left(v^{-}\right)^{v^{-}} \right\}^{1/\nu}$$

Finally, since $v = v^+ + v^-$, we obtain:

$$a_{\pm} = \left(\gamma_{+}^{\nu^{+}} \gamma_{-}^{\nu^{-}}\right)^{1/\nu} (1 - \alpha) m \left\{ \left(\nu^{+}\right)^{\nu^{+}} \left(\nu^{-}\right)^{\nu^{-}} \right\}^{1/\nu}$$

$$4.96$$

We can recognize the last term as m_{\pm} . Since $a_{\pm} = \gamma_{\pm}m_{\pm}$, we see that the mean ionic activity coefficient will be

$$\gamma_{\pm} = (1 - \alpha) \left(\gamma_{+}^{\nu^{+}} \gamma_{-}^{\nu^{-}} \right)^{1/\nu}$$
4.97

for an incompletely dissociated electrolyte. Thus the mean ion activity coefficients are reduced by a factor of $1 - \alpha$. Provided we have appropriate stability constants for the ion pairs or complexes, α can be calculated and an appropriate correction applied.

Now consider a CaSO₄ solution of which some fraction of the Ca²⁺ and SO₄²⁻ ions, α , associate to form CaSO₄^o. The ionic strength of the this solution would be

$$I = \frac{(1-\alpha)}{2} \Big(4m_{Ca^{2+}} + 4m_{SO_4^{2-}} \Big)$$

Thus the ionic strength is reduced by a factor of 1 - α as well.

Ion pairs and complexes need not be neutral species (AlCl²⁺, for example). When they are not, they will contribute to ionic strength. A general expression for ionic strength taking account of ion associations must include charged ion pairs and complexes:

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

$$I = \frac{1}{2} \left[\sum_{i} (1 - \alpha) m_i z_i^2 + \sum_{n} c_n z_n^2 \right] \quad 4.98$$

where α_i is the fraction of each ion involved in ion associations, and c_n is the concentration of each ion pair or complex and z_n is its charge. We could use this result directly in the Debye-Hückel equation to make an improved estimate of ionic strength, and hence of the single ion activity coefficient.

Figure 4.29 illustrates the effect of ion pair formation for a hypothetical $CaCl_2$ solution in which some fraction of the ions combine to form ion pairs. The fraction of Ca^{2+} ions forming $CaCl^-$ was assumed to increase linearly with ionic strength up to the maximum value shown.

If the formation of ion pairs depends on the ratio of thermal to electrostatic energy, we might expect that ion pair formation will decrease with temperature. However, the relative permittivity of water decreases with temperature, allowing increased electrostatic interaction between ions, and this effect dominates over the increased thermal energy of ions. As a



Figure 4.29. Effects of ion association on the activity coefficient. Mean ion activity coefficient of CaCl₂ for varying extents of ion association. Fraction of Ca²⁺ ions forming CaCl⁻ was assumed to increase linearly with ionic strength up to a maximum value (α_{max}) at I = 5 m. Solid line shows electrostatic term (Debye-Hückel) after correction for ion association, dashed line shows the combined electrostatic and solvation term.

Example 4.7. Activity Coefficients in a Brine Specie							vecies Conc		
The following concentrations were measured in a shield brine from Sudbury, Canada at 22° C. Calc ulate the activity coefficients of these species using the Truesdell-Jones equation. Na K K Answer: Our first task is to convert g/kg to molal concentrations. We do this by dividing by molecular weight. Next, we need to calculate ionic strength (equation 3.75) which we find to be 5.9 m. Calculation of activity coefficients is then straightforward using the parameters in Tables 3.2 and 4.6. Finally, we apply a HCO3 correction for the decreased concentration of water (equation 4.89). Our final SO4							$\begin{array}{r} g/kg \\ 18.9 \\ 0.43 \\ 63.8 \\ 0.078 \\ 0.223 \\ 0.042 \\ 162.7 \end{array}$		
spreadsheet is	snown be	now.							
	%00	m	Z	å_TJ	b_TJ	log (gamma)	ganma	gamma o	corr
Na	18.9	0.822	1	5	0.165	0.728	5.341	4.741	
K	0.43	0.017	1	3.5	0.015	-0.238	0.579	0.514	

Κ	0.43	0.017	1	3.5	0.015	-0.238	0.579	0.514
Ca	63.8	1.595	2	5	0.165	-0.017	0.963	0.855
Mg	0.078	0.003	2	5.5	0.2	0.264	1.836	1.630
SO4	0.223	0.002	2	5	-0.04	-1.229	0.059	0.052
HCO3	0.058	0.001	1	5.4	0	-0.233	0.585	0.519
Cl	162.7	4.590	1	3.5	0.015	-0.238	0.579	0.514
	m	7.030	A	0.5092				
	Ι	5.913	В	0.3283				

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

result, the extent of ion association increases with temperature. Increasing pressure, on the other hand, favors dissociation of ions.

4.7.2.2 Alternative Expressions for Activity Coefficients

There have been a number of attempts to develop working equations that account for all the effects on activity coefficients at high ionic strength. Many of these are ultimately based on the specific ion interaction theory of Brønsted (1922). Brønsted proposed an equation of the form:

TERS			
Ion	å	b	
Na^+	4.0	0.075	
K^+	3.5	0.015	
Mg^{2+}	5.5	0.20	
Ca ²⁺	5.0	0.165	
Cl⁻	3.5	0.015	
SO_4^{2-}	5.0	-0.04	
CO_{3}^{2-}	5.4	0	
HCO ₃	5.4	0	

$$\log \gamma_i = \alpha m^{1/2} + \beta_i m \tag{4.99}$$

where α is a constant that is independent of the solute ions and β is the "specific ion interaction parameter" and is different for each ionic species. Guggenheim (1935) replaced the first term on the right with a simplified form of the Debye-Hückel equation and the second term with the summation of ion-ion interaction parameters:

$$\log \gamma_i = \frac{-z_i^2 A \sqrt{I}}{1 + \sqrt{I}} + 2\sum_k \beta_{i,k} m_k$$

$$4.100$$

where $\beta_{i\kappa}$ is parameter describing the interactions between ions *i* and *k*. For natural waters with many species, the Guggenheim equation becomes complex. Also starting from Debye-Hückel, Truesdell and Jones (1974) proposed the following simpler equation:

$$\log \gamma_i = \frac{-z_i^2 A \sqrt{I}}{1 + B \mathring{a}_i \sqrt{I}} + b_i I \log \gamma_i = \frac{-z_i^2 A \sqrt{I}}{1 + B \mathring{a}_i \sqrt{I}} + b_i I$$

$$4.101$$

The first term on the right is identical in form to Debye-Hückel; the second term is similar to the Brønsted specific ion interaction term. Truesdell and Jones determined parameters å and b empirically. Table 4.5 lists these parameters for some common ions. Figure 4.30 compares mean activity coefficient of calculated with the Debye-Hückel, Davies, and Truesdell-Jones equations with the actual measured values. The Truesdell-Jones equations fit these observations very well. This is not always the case, however. The fit for Na₂CO₃, for example is little better than for Debye-Hückel.

Other equations include those developed by Pitzer (1979) and the National Bureau of Standards. While these equations are generally more accurate than the above, their complexity places them beyond the scope of this book. The interested reader is referred to any of several texts on geochemical thermodynamics that treat them (Nordstrom and Munoz, 1986; Fletcher, 1993; Anderson and Crerar, 1993) as well as the original literature.

References and Suggestions For Further Reading

Anderson, G. M. and D. A. Crerar. 1993. *Thermodynamics in Geochemistry*. New York: Oxford Univ. Press.

Bohlen, S. R., W. A. Dollase and V. J. Wall. 1986. Calibration and application of spinel equilibria in the system FeO-Al₂O₃-SiO₂. *J. Petrol.* 27: 1143-1156.

- Bohlen, S. R. and D. H. Lindsley, 1987. Thermometry and barometry of igneous and metamorphic rocks. *Ann. Rev. Earth Planet. Sci.* 15:397-420.
- Bohlen, S. R., V. J. Wall and A. L. Boettcher. 1983. Experimental investigations and geologica applications of equilibria in the system FeO-TiO₂-Al₂O₃-SiO₂-H₂O. *Am. Mineral.* 68: 1049-1058.

 Table 4.5.
 Truesdell-Jones
 Parame

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

- Bottinga, Y. and D. Weill. 1972. The viscosity of magmatic silicate liquids: a model for calculation. *Am. J. Sci.* 272: 438-475.
- Boyd, F.R. 1973. A pyroxene geotherm. Geochim. Cosmochim. Acta, 37, 2533-2546.
- Bradley, R. S. 1962. Thermodynamic calcuations on phase equilibria involving fused salts. Part II. Solid solutions and applications to the olivines. *Am. J. Sci.*, 260, 550-554.
- Brønsted, J. N. 1922. Studies on solubility, IV. The principle of specific interaction of ions. J. Am. Chem. Soc. 44: 877-898.
- Buddington, A. F. and D. H. Lindsley. 1964. Iron-titanium oxide minerals and synthetic equivalents, *J. Petrol.* 5: 310-357.
- Carmichael, I. S. E. and H. P. Eugster (eds). 1987. Thermodynamic Modeling of Geological Materials: Minerals, Fluids and Melts (Reviews in Mineralogy, vol 17). Washington: Mineral. Soc. Am.
- Essene, E. 1982. Geologic thermometry and barometry. In *Characterization of metamorphism through mineral equilibria, Reviews in Mineralogy Vol. 10,* J. M. Ferry. ed., pp. 153-206. Washington: Mineral. Soc. Am.
- Essene, E. 1989. The current status of thermobarometry in metamorphic rocks. In *The Evolution of Metamorphic Belts, Vol.* J. S. Daly, R. A. Cliff and B. W. D. Yardley. ed., pp. 1-44. London: Blackwell Sci. Pub.
- Ferry, J. M. and F. S. Spear. 1978. Experimental calibration of the partitioning of Fe and Mg between biotite and garnet. *Contrib. Mineral. Petrol.* 25: 871-893.
- Fletcher, P. 1993. Chemical Thermodynamics for Earth Scientists, Essex: Logman Scientific and Technical.
- Garrels, R. M., and Christ, C. L. 1965. Solutions, Minerals, and Equilibria. New York: Harper and Row.
- Gasparik, T. 1984a. Two-pyroxene thermobarometry with new experimental data in the system CaO-MgO-Al₂O₃-SiO₂. *Contrib. Mineral. Petrol.* 87: 87-97.
- Gasparik, T. 1984b. Experimental study of subsolidus phase relations and mixing properties of pyroxene and plagioclase in the system CaO-Al₂O₃-SiO₂. *Geochim. Cosmochim. Acta.* 48: 2537-2446.
- Ghent, E. D. 1976. Plagioclase-garnet-Al₂O₅-quartz: a potential geobarometer-geothermometer. *Am. Mineral.* 61: 710-714.
- Ghiorso, M. S. 1987. Modelling magmatic systems: Thermodynamic relations. in *Thermodynamic Modelling of Geological Materials: Minerals, Fluids,*
- and Melts. eds. I. S. E. Carmichael and H. P. Eugster. 443-465. Washington: Mineral. Soc. Am.
- Ghiorso, M. S. and I. S. E. Carmichael. 1985. Chemical mass transfer in magmatic processes: II. Applications in equilibrium crystallization, fractionation crystallization, and assimilation, *Contrib. Mineral. Petrol.* 90: 121-141.
- Ghiorso, M. S., I. S. E. Carmichael, M. L. Rivers and R. O. Sack. 1983. The Gibbs free energy of mixing of natural silicate liquids; an expanded regular solution approximation for the calculation of magmatic intensive variables. *Contrib. Mineral. Petrol.* 84: 107-145.
- Ghiorso, M. S., M. M. Hirshmann and R. O. Sack. 1994. New software models thermodynamics of magmatic systems. *EOS*. 75: 571-576.
- Ghiorso, M. S., M. M. Hirschmann, P. W. Reiners and V. C. Kress, 2002. The pMELTS: a re-



Figure 4.30. Measured mean ionic activity coefficients in $MgCl_2$ solution as a function of ionic strength compared with values calculated from the Debye-Hückel, Davies and Truesdell-Jones equations.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

vision of MELTS for improved calculation of phase relations and major element partitioning related to partial melting of the mantle to 3 GPa, *Geochem. Geophys. Geosyst.*, 3:2001GC000217.

- Ghiorso, M. S. and R. O. Sack. 1995. Chemical mass transfer in magmatic processes IV. A revised and internally consistent thermodynamic model for the interpolation and extrapolation of liquid-solid equilibria in magmatic systems at elevated temperatures and pressures. *Contrib. Mineral. Petrol.* 119: 197-212.
- Goldsmith, J. R. and R. C. Newton. 1969. P-T-X relations in the system CaCO₃-MgCO₃ at high temperatures and pressures. *Am. J. Sci.* 267A: 160-190.
- Grover, J., 1977. Chemical mixing in multicomponent solutions: An introduction to the use of margules and other thermodynamic excess functions to represent non-ideal behavior, in *Thermodynamics in Geology*. ed. D. G. Fraser. 67-97. Dordrecht: D. Reidel.
- Guggenheim. 1935. The specific thermodynamic properties of aqueous solutions of strong electrolytes. *Phil. Mag.* 19: 588.
- Helgeson, H. C. 1969. Thermodynamics of hydrothermal systems at elevated temperatures and pressures. *Am. J. Sci.* 267: 729-804.
- Helgeson, H. C. and D. H. Kirkham. 1974. Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures, Debye-Hückel parameters for activity coefficients and relative partial molar quantities. *Am J. Sci.* 274: 1199-1261.
- Helgeson, H. C., Kirkham, D. H., and G. C. Flowers. 1981. Theoretical predicitons of the thermodynmaic behavior of electrolytes at high pressures and temperatures: IV. Calculation of activity coefficients, osmotic coefficients, and apparent molar and standard and relative partial molar properties to 600°C and 5 kbar. *Am. J. Sci.* 281: 1249-1316.
- Holdaway, M. J. 1971. Stability of andalusite and the aluminum silicate phase diagram. *Am. J. Sci.* 271: 91-131.
- Holland, T. J. B., A. Navrotsky, and R. C. Newton, 1979. Thermodynamic parameters of CaMgSi₂O₆ Mg₂Si₂O₆ pyroxenes based on regular solution and cooperative disordering models. *Contrib. Mineral. Petrol.* 69: 337-344.
- Johannes, W. and D. Puhan. 1971. The calcite-aragonite equilibrium reinvestigated. *Contrib. Mineral. Petrol.* 31: 28-38.
- Koziol, A. M. and R. C. Newton. 1988. Redetermineation of the garnet breakdown reaction and improvement of the plagioclase-garnet-Al₂O₅-quartz geobarometer. *Am. Mineral.* 73: 216-223.

Navrotsky, A. 1994. Physics and Chemistry of Earth Materials. Cambridge: Cambridge University Press.

- Nicholls, J. and J. K. Russell. 1990. Modern Methods in Igneous Petrology, Reviews in Mineralology, vol. 24, Washington: Min. Soc. Am.
- Nielsen, R. L. and M. A. Dungan. 1983. Low pressure mineral-melt equilibria in natural anhydrous mafic systems, *Contrib. Mineral. Petrol.*, 84: 310-326.
- Nordstrom, D. K. and J. L. Munoz, 1986. *Geochemical Thermodynamics*, Palo Alto: Blackwell Scientific Publ.
- Pitzer, K. S. 1979. Theory: ion interaction approach. in *Activity Coefficients in Electolytes*, ed. R. M. Pytkowicz. 157-208. Boca Raton: CRC Press.
- Rankenburg, K., J. C. Lassiter and G. Brey, 2004. Origin of megacrysts in volcanic rocks of the Cameroon volcanic chain - constraints on magma genesis and crustal contamination, *Contr. Mineral. and Petrol.*, 147:129-144.
- Robinson, R. A. and R. H. Stokes. 1959. *Electrolyte Solutions*. London: Butterworths.
- Roeder, P.L. and R.F. Emslie. 1970. Olivine-liquid equilibrium. Contrib. Mineral. Petrol. 29: 275-289.
- Spencer, K. and D. H. Lindsley. 1981. A solution model for coexisting iron-titanium oxides. *Am. Mineral.* 66: 1189-1201.
- Stolper, E. and H. Y. McSween. 1979. Petrology and origin of the shergottite meteorites. *Geochem. Cosmochim. Acta*. 43: 1475-1498.

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

- Thompson, J. B. 1967. Thermodynamic properties of simple solutions, in *Researches in Geochemistry, vol.* 2, ed. P. H. Abelson. New York: Wiley and Sons.
- Thompson, J. B. and D. R. Waldbaum, 1969. Mixing properties of sanidine crystalline solutions: III. Calculations based on two-phase data. *Am. Mineral.* 54: 811–838.
- Truesdell, A. H. and B. F. Jones. 1974. WATEQ, a computer program for calculating chemical equilibria in natural waters. *J. Res. U. S. Geol. Surv.* 2: 233-248.
- Wood, B. J., 1987. Thermodynamics of multicomponent systems containing several solutions, in *Ther-modynamic Modeling of Geological Materials: Minerals, Fluids and Melts*, eds. I. S. E. Carmichael and H. P. Eugster. 71-96. Washington: Mineral. Soc. Am.
- Wood, B.J. and S. Banno. 1973. Garnet-orthopyroxene and orthopyroxene-clinopyroxene relationships in simple and complex systems. *Contrib. Mineral. Petrol.* 42:109-124.

Problems

1. Kyanite, and alusite, and sillimanite (all polymorphs of Al_2SiO_5) are all in equilibrium at 500°C and 376 MPa. Use this information and the adjacent table to construct an approximate temperature-pressure phase diagram for the system

	$\bar{\mathrm{v}}$	S
φ	(cm^3)	(J/K-mol)
kyanite	44.09	242.30
andelusite	51.53	251.37
sillmanite	49.90	253.05

kyanite-sillimanite-andalusite. Assume ΔV and ΔS are independent of temperature and pressure. Label each field with the phase present.

2. Show that:
$$\overline{G}_{excess} = (W_{G_1}X_2 + W_{G_2}X_1)X_1X_2$$
 may be written as a 4 term power expansion, i.e.:
 $\overline{G}_{ex} = A + BX_2 + CX_2^2 + DX_2^3$

3. Construct G-bar–X diagrams for a regular solution with W= 12 kJ (W is the interaction parameter in a non-ideal solution) at 100° temperature intervals from 200 to 700° C. Sketch the corresponding phase diagram.

4. Interaction parameters for the enstatite-diopside solid solution have been determined as follows: $W_{H-En} = 34.0 \text{ kJ/mol}$, $W_{H-Di} = 24.74 \text{ kJ/mol}$ (assume W_V and W_S are 0).

a.) Use the asymmetric solution model to calculate ΔG_{real} as a function of X_2 (let diopside be component 2) curves for this system at 100 K temperature from 1000 K to 1500 K. Label your curves.

b.) What is the maximum mole fraction of diopside that can dissolve in enstatite in this temperature range:?

c.) Sketch the corresponding T-X phase diagram.

5. Sketch G-bar–X diagrams for 1600° C, 1500° C, 1300° C, and 1250° C for the system Diopside-Anorthite (Figure 4.8). Draw tangents connecting the equilibrium liquids and solids.

6. Suppose you conduct a 1 atm melting experiment on a plagioclase crystal. Predict the mole fractions of anorthite in the liquid and solid phases at a temperature of 1425° C. Assume both the liquid and solid behave as ideal solutions. Albite melts at 1118° C, anorthite at 1553°C. ΔH_m for albite is 54.84 kJ/mol; ΔH_m for anorthite is 123.1 kJ/mol.

7. Given the following 2 analyses of basaltic glass and coexisting olivine phenocrysts, determine the K_D for the MgO \rightleftharpoons FeO exchange reaction, and calculate the temperatures at which the olivine crystallized using both MgO and FeO. Assume Fe₂O₃ to be 10 mole% of total iron (the analysis below includes only the total iron, calculated as FeO; you need to calculate from this the amount of FeO by substracting an appropriate amount to be assigned as Fe₂O₃). Note that the mole % Fo in olivine is equivalent to the mole % Mg or MgO. (*HINT: you will need to calculate the mole fraction of MgO and FeO in the liquid*).

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

Sample	TR3D-1	DS-D8A	
1	(wt % oxide)	(wt % oxide)	
SiO ₂	50.32	49.83	
Al ₃ O ₂	14.05	14.09	
ΣFe as FeO	11.49	11.42	
MgO	7.27	7.74	
CaO	11.49	10.96	
Na ₂ O	2.3	2.38	
K ₂ O	0.10	0.13	
MnO	0.17	0.20	
TiO ₂	1.46	1.55	
olivine			
Mole % Fo (=mole % Mg)	79	81	

Glass (liquid) composition:

8. Determine the temperature and oxygen fugacity of equilibration for the following set of coexisting iron-titanium oxides in lavas from the Azores:

	titanomagnetite s.s. phase mole percent magnetite	ilmenite s.s. phase mole percent hematite	
G-4 groundmass	29.0	10.3	
SJ-8 phenocrysts	41.9	13.0	
SM-28 microphenocrysts	54.5	7.0	
T-8 groundmass	33.7	8.1	
F-29 microphenocrysts	36.2	6.0	

Make a plot of f_{O2} vs. temperature using your results and compare with Fig. 3.21. What buffer do the data fall near?

9. Starting from equations 4.54, 4.56 and 4.18, use the fundamental relationships between free energy, entropy, enthalpy, and the equilibrium constant to derive the temperature dependence of the titanomagnetite-ilmentite exchange (equation 4.57).

10. In a melt having a composition, in wt %, of:

SiO_2	58.12% TiO ₂	0.92%				
Al_2O_3	16.47% Fe ₂ O ₃	1.82%				
MgO	5.62%	FeO	9.94%	CaO	7.11%	
	1 1	•	11 1.1 .			

Use the Ghiorso regular solution model and the interaction parameters in Table 4.3 to: a.) calculate the \overline{G}_{ex} and $\Delta \overline{G}_{mixing}$ for this composition at 1300°C. b.) calculate the activity of Si₄O₈ at this temperature.

Geochemistry

CHAPTER 4: APPLICATIONS OF THERMODYNAMICS

11. An analysis of an oil field brine from Mississippi with a temperature of 150° C is shown to the right. Calculate the *activities* of these species at that temperature using the Truesdell-Jones equation.

12. Show that for a strong electrolyte, i.e., one in which dissociation is complete and:

 $m_{-} = v_{-}m$ and $m_{+} = v_{+}m$ where m is the molality of the solute component $A_n + B_n$ -, that:

$$m_{\pm} = m \left(\nu_{+}^{\nu^{+}} \nu_{-}^{\nu^{-}} \right)^{1/\nu}$$

Problem 11 Species Conc g/kg Na⁺ 63.00 K⁺6.15 Mg²⁺ Ca²⁺ 2.77 44.6 Cl-_____ 200.4 SO_4^2 0.13 HCO 0.03

13. Mean ionic activity coefficients were measured for the following solutions at an ionic strength of 3: $\gamma_{\text{KCl}} = 0.569$, $\gamma_{\text{NaCl}} = 0.734$, $\gamma_{\text{Na_2CO_3}} = 0.229$. Assuming $\gamma_{\text{Cl}-} = \gamma_{\text{K}-} = \gamma_{\pm\text{KCl}}$ what is the activity coefficient of CO_3^{2-7} ?

14. Calculate the electrostatic and solvation contributions to the mean ionic activity coefficient of MgCl₂ at concentrations of 0.0033, 0.01, 0.033, 0.05, 0.1, 0.33, 0.5, and 1 using the Debye-Hückel equation and Robinson and Stokes (equ. 4.92) equ. respectively. Plot results, as well as $\gamma_{\text{elect}+\text{solv}} = \gamma_{\text{elect}}*\gamma_{\text{solv}}$ as a function of ionic strength.

15. Calculate the mean ionic activity coefficient for NaCO₃ using the Debyeγ± Hückel and Truesdell-Jones equations and compare your results with the ob-I, m observed served values to the right. Overall, which fits the data better? 0.001 0.003 0.887 0.006 0.847 0.01 0.015 0.78 0.03 0.716 0.06 0.644 0.1 0.15 0.541 0.3 0.462 0.6 0.385 1 1.5 0.292 3 0.229

0.229

6

CHAPTER 5: KINETICS

CHAPTER 5: KINETICS: THE PACE OF THINGS

5.1 INTRODUCTION

Thermodynamics concerns itself with the distribution of components among the various phases and species of a system at equilibrium. *Kinetic* concerns itself with the *path* the system takes in achieving equilibrium. Thermodynamics allows us to predict the equilibrium state of a system. Kinetics, on the other hand, tells us how and how fast equilibrium will be attained. Although thermodynamics is a macroscopic science, we found it often useful to consider the microscopic viewpoint in developing thermodynamics models. Because kinetics concerns itself with the path a system takes, what we will call *reaction mechanisms*, the microscopic perspective becomes essential, and we will very often make use of it.

Our everyday experience tells one very important thing about reaction kinetics: they are generally slow at low temperature and become faster at higher temperature. For example, sugar dissolves much more rapidly in hot tea than it does in ice tea. Good instructions for making ice tea might then incorporate this knowledge of kinetics and include the instruction to be sure to dissolve the sugar in the hot tea before pouring it over ice. Because of this temperature dependence of reaction rates, low temperature geochemical systems are often not in equilibrium. A good example might be clastic sediments, which consist of a variety of phases. Some of these phases are in equilibrium with each other and with porewater, but most are not. Another example of this disequilibrium is the oceans. The surface waters of the oceans are everywhere oversaturated with respect to calcite, yet calcite precipitates from seawater only through biological activity. At a depth of 2500 m, the ocean is undersaturated with calcite, yet calcite shells of micro-organisms persist in sediments deposited at these depths (though they do dissolve at greater depths). Thus, great care must be used in applying thermodynamics to such systems. Even in the best of circumstances, thermodynamics will provide only a limited understanding of low temperature geochemical systems. A more complete understanding requires the application of kinetic theory. Indeed for such systems, kinetics is the deciding factor controlling their state and evolution. Even in metamorphic systems, with temperatures in the range of 300-700° C, kinetics factors are crucially important in determining their final states.

High temperature geochemical systems, such as magmas, are more likely to be in equilibrium, and thermodynamics provides a reasonable understanding of these systems. However, even at high temperatures, kinetic factors remain important and can inhibit equilibrium. One obvious example of disequilibrium at high temperature is the formation of volcanic glasses. Thermodynamics predicts that magmas should crystallize as they cool. But where cooling is rapid enough, this does not occur. Glasses, which in many ways are simply extremely viscous liquids, form instead.

It is perhaps ironic that it is kinetic factors, and a failure to achieve equilibrium, that in the end allow us to use thermodynamics to make statements about the Earth's interior. As we pointed out in the preceding chapter, if equilibrium were always achieved, the only rocks we could collect at the surface of the Earth (which is, after all, the only place we can collect them) would consist of quartz, clays, serpentine, etc.; their petrology would tell us nothing about their igneous or metamorphic histories. Fortunately, kinetic factors allow the original minerals and textures of gneisses, peridotites, lavas, etc. to be preserved for our study.

The foregoing might suggest that kinetics and thermodynamics are entirely unrelated subjects, and further, that what we have learned about thermodynamics is of little use in many instances. This is certainly not the case. As we shall see, transition state theory provides a very strong link between kinetics and thermodynamics. What we have learned about thermodynamics will prove very useful in our brief study of kinetics. Furthermore, chemical systems are always governed by a combination of thermodynamics and kinetics, so a full understanding of the Earth requires the use of both thermodynamic and kinetics tools. The goal of this chapter is to add the latter to our geochemical toolbox.

Geochemistry

CHAPTER 5: KINETICS

5.2 REACTION KINETICS

5.2.1 Elementary and Overall Reactions

In thermodynamics, we found that the equilibrium state of a system is entirely independent of the path taken to reach that state. The goal of kinetics is a description of the manner in which the equilibrium state is achieved. This description is inherently path-dependent. Consider for example, the weathering of anorthite. We can write an *overall* reaction for this process as:

$$CaAl_{2}Si_{2}O_{8} + 3H_{2}O + CO_{2(g)} \rightarrow CaCO_{3} + 2Al(OH)_{3} + 2SiO_{2(qz)}$$
 5.1

In nature, however, this process will involve several intermediate steps. These intermediate steps can include:

$$H_2O + CO_{2(g)} \rightarrow H_2CO_{3(aq)}$$
 5.2

$$H_2CO_3 \rightarrow HCO_{3(aq)} + H^+$$
 5.3

$$CaAl_2Si_2O_8 + H_2O + 2H^+ \rightarrow Si_2Al_2O_5(OH)_4 + Ca^{2+}_{(aq)}$$
 5.4

$$H_2O + Si_2Al_2O_5(OH)_4 \rightarrow 2SiO_{2(qz)} + 2Al(OH)_3$$
 5.5

$$\mathrm{HCO}_{3(\mathrm{aq})}^{-} \rightarrow \mathrm{CO}_{3(\mathrm{aq})}^{2-} + \mathrm{H}^{+}$$
 5.6

$$CO_{3(aq)}^{2-} + Ca_{(aq)}^{2+} \rightarrow CaCO_3$$
5.7

In thermodynamics, equation 5.1 is a perfectly adequate description of the reaction. In kinetics, a description of an *overall* reaction such as 5.1, requires a knowledge of the path taken, that is a knowledge of the steps involved. Reactions 5.2 through 5.7 thus describe the overall reaction 5.1. Reactions 5.2, 5.3, and 5.6 are *elementary reactions* in that they involve only one step and the reaction as written describes what occurs on the microscopic level. The remaining reactions are not elementary in that they each consist of a number of more elementary steps.

5.2.2 REACTION MECHANISMS

Reaction 5.4 describes the breakdown of anorthite to form kaolinite plus a free calcium ion. This reaction involves profound structural changes in the solid phase that are not described by equation 5.4. A full kinetic description of 5.4 will require some knowledge of the steps involved in these structural changes. One possibility is that all components are in solution at an intermediate state:

$$CaAl_{2}Si_{2}O_{8} + 6H_{2}O + 2H^{+} \rightarrow 2H_{4}SiO_{4(aq)} + 2Al(OH)_{2(aq)}^{+} + Ca_{(aq)}^{2+} + 2OH^{-}$$
 5.4a

$$2H_4SiO_{4(aq)} + 2Al(OH)_{2(aq)}^+ + 2OH^- \rightarrow Si_2Al_2O_5(OH)_4 + 5H_2O$$
 5.4b

Reaction 5.5, the breakdown of kaolinite to quartz and gibbsite, could involve SiO_2 dissolving, subsequently precipitating as opaline silica, and later transforming to quartz:

$$Si_2Al_2O_5(OH)_4 + 5H_2O \rightarrow 2H_4SiO_{4(aq)} + 2Al(OH)_3$$
 5.5a

$$H_4SiO_{4(aq)} \rightarrow SiO_{2(opal)} + 2H_2O$$
 5.5b

$$SiO_{2(opal)} \rightarrow SiO_{2(az)}$$
 5.5c

The description of an overall reaction in terms of elementary reactions is called the *reaction mechanism*. The rates of truly elementary reactions are path-independent because there is only one possible path. In this sense, elementary reactions are somewhat analogous to state functions in thermodynamics. Clearly then, an important step in any kinetic study is determination of the reaction mechanism, that is, to describe the process in terms of elementary reactions. As we shall see, there may be more than one possible path for an overall reaction, and that several paths may be simultaneously involved. Kinetics can only provide an accurate description of a process if all these paths are known.

Geochemistry

CHAPTER 5: KINETICS

5.2.3 REACTION RATES

Consider a reaction such as the precipitation of dolomite from a solution. We can describe this as:

$$Ca^{2+} + Mg^{2+} + 2CO_3^2 \rightleftharpoons CaMg(CO_3)_2$$

We *define* the *rate* of this reaction, \Re , as the rate at which dolomite is produced:

$$\Re \equiv \frac{d[\text{CaMg(CO_3)}_2]}{dt}$$

Clearly, if dolomite is to be formed, CO_3^{2-} must be consumed in this reaction twice as fast as Ca or Mg. For every mole of Ca or Mg consumed, exactly two moles of CO_3^{2-} will also be consumed and one mole of dolomite produced. This being the case, we could equally well express the reaction rate as:

$$\Re = -\frac{1}{2} \frac{d[\operatorname{CO}_3^{2^-}]}{dt} \qquad \text{or} \qquad \Re = -\frac{d[\operatorname{Ca}^{2^+}]}{dt} = -\frac{d[\operatorname{Mg}^{2^+}]}{dt}$$

We can now formulate the general rule. For any reaction such as:

$$aA + bB \rightarrow cC + dD$$
 5.8

The *reaction rate*, \Re , is defined as the change in composition of the reaction mixture with time:

$$\Re = -\frac{1}{a}\frac{d[A]}{dt} = -\frac{1}{b}\frac{d[B]}{dt} = \frac{1}{c}\frac{d[C]}{dt} = \frac{1}{d}\frac{d[D]}{dt}$$
5.9

The brackets denote the concentrations of the species and the negative sign indicates that reactants are consumed as the reaction proceeds. Thus *the rate of a reaction is simply the rate at which a reactant is consumed or product produced* divided by its stoichiometric coefficient.

5.2.3.1 The Reaction Rate for an Elementary Reaction: Composition Dependence

Reaction rates will, in general, depend on the concentration of the reactant. To understand this, consider the reaction:

$$N^{\circ} + O_2 \rightleftharpoons NO + O$$
 5.10

This reaction between free nitrogen atoms and oxygen molecules occurs in the stratosphere (where N° is produced by high energy collisions involving N₂) and contributes to the production of nitrous oxide. Let's assume that reaction 5.10 is an adequate description of this reaction. In other words, we are assuming that 5.10 is an elementary reaction and the reaction mechanism for the production of NO from nitrogen and oxygen gas is collision of a N° molecule and O₂ molecule. For the reaction to occur, the nitrogen and oxygen molecules must collide with enough kinetic energy that the mutual repulsion of the electron clouds is overcome and the electrons can be redistributed into new covalent orbits. The repulsive force represents an energy barrier, E_B , which will prevent low energy nitrogen and oxygen atoms



distributed into new covalent orbits. Figure 5.1. A nitrogen atom approaching an oxygen molecule The repulsive force represents an energy barrier, E_{B} , which will prevent low energy nitrogen and oxygen atoms from reacting. Figure 5.1 illustrates is repealed by electrostatic repulsion of the electron cloud of the oxygen. Otherwise, it will not approach closely enough so that its electrons can combine with those of oxygen.

W. M. White Chapter 5: Kinetics

this point. The reaction rate will therefore depend on 1) the number of collisions per unit time, and 2) the fraction of N and O molecules having energy greater than the barrier energy.

Let's first consider the number of collisions per unit time. In order for a 'collision' to occur, the electron clouds must overlap, that is, they must approach within $(r_N +$ r_{O_2}), where r_N and r_{O_2} are the radii of the nitrogen and oxygen molecules. To make things simple, imagine the oxygen to be fixed and the nitrogen in motion. In other words, our reference frame will be that of the oxygen molecules. We can imagine the nitrogen sweeping out a cross-section with radius (r_N $+ r_{O_2}$) as it travels. If the nitrogen



Figure 5.2. A nitrogen atom will sweep out a volume $V = v\pi(r_N + r_{O_2})^2 t$ in time *t*. Whether a collision occurs will depend on whether the center (indicated by black dot) of an oxygen atom falls within this volume.

is travelling at velocity v, in time t, it will sweep out a cylindrical volume (Figure 5.2):

$$V = v\pi (r_N + r_{O_2})^2 t 5.11$$

Whether a collision occurs will depend on whether the *center* of an oxygen molecule falls within this volume (Figure 5.2). The number of collisions that will occur in this time will be:

$$C = n_0 v \pi (r_N + r_{02})^2 t$$
 5.12

where n_O is the number of oxygen molecules per unit volume. The number of collisions per unit time is then simply:

$$\frac{C}{t} = n_0 v \pi (r_N + r_{O_2})^2$$
 5.13

If there are n_N nitrogen atoms and the average velocity between nitrogen and oxygen molecules is \overline{v} , then the number of collisions per unit time is:

 $\dot{c} = n_N n_O \overline{v} \pi \left(r_N + r_{O_2} \right)^2$ $k = \overline{v} \pi \left(r_N + r_{O_2} \right)^2$ 5.14

If we let

then the rate at which collisions occur is:

 $\dot{c} = kn_N n_{O_2} \tag{5.15}$

Thus we see that *the reaction rate in this case will depend on the concentration of nitrogen, oxygen and a constant* that depends on the nature of the reactants. This is a general result.

5.2.3.2 The Reaction Rate for an Elementary Reaction: Temperature Dependence

We now need to estimate the fraction of nitrogen and oxygen atoms having at least the barrier energy, E_B . For simplicity, we will assume that oxygen and nitrogen molecules have an identical energy distribution. The Boltzmann Distribution Law, which we encountered in Section 2.6.4.1 (equ. 2.84), can be written to express the average number of molecules having energy level ε_i as:

$$n_i = A e^{-\varepsilon_i / kT}$$
 5.16

CHAPTER 5: KINETICS

W. M. White

where k is Boltzmann's constant and A is a constant (comparing with equation 2.84, we see that A = n/Q where n is the total number of molecules in the system and Q is the *partition function*). In plain English, this equation tells us that the number of molecules in some energy level *i* decreases exponentially as the energy of that level increases (Figure 2.9). We want to know the number of molecules with energy greater than E_B . In this case we are dealing with translational energy. The quantum spacings between translational energy levels are so small that they essentially form a continuum, allowing us to integrate equation 5.16. Fortunately for us, the integration of 5.16 from $\varepsilon = E_B$ to infinity has a simple solution:

$$A\int_{E_B}^{\infty} e^{-\varepsilon_i/kT} d\varepsilon = AkT e^{-E_B/kT}$$
5.17

The *fraction* of molecules with energy greater than E_B is just:

$$\frac{A\int_{E_B}^{\infty} e^{-\varepsilon_i/kT} d\varepsilon}{A\int_0^{\infty} e^{-\varepsilon_i/kT} d\varepsilon} = \frac{AkTe^{-E_B/kT}}{AkT} = e^{-E_B/kT}$$
5.18

The rate of reaction will be the rate of collision times the fraction of molecules having energy greater than E_{B} :

$$\Re = n_N n_O \overline{\nu} \pi \left(r_N + r_{O_2} \right)^2 e^{-E_B/kT}$$
 5.19

Now we just need to find a value for velocity. The average velocity can be calculated from the Maxwell-Boltzmann Law^{*}, which gives the distribution of velocities of molecules in a gas. Doing so, we find that the average velocity is:

$$\overline{v} = \sqrt{\frac{8kT}{\pi\mu}}$$
5.20

where μ is the reduced mass, $\mu = m_N m_{O_2} / (m_N + m_{O_2})$. Substituting 5.20 into 5.19, our equation for the reaction rate is:

$$\Re = n_N n_O \pi \left(r_N + r_{O_2} \right)^2 \sqrt{\frac{8kT}{\pi\mu}} e^{-E_B/kT}$$
 5.21

$$k = \pi \left(r_N + r_{O_2} \right)^2 \sqrt{\frac{8kT}{\pi\mu}} e^{-E_B/kT}$$
 5.22

our reaction rate equation is:

$$\Re = kn_N n_{Q_2}$$
 5.23

Thus the reaction rate in this case depends on the concentration of nitrogen and oxygen and a constant k, called the *rate constant*[†], which depends on temperature, properties of the reactants, and the barrier energy.

In a more rigorous analysis we would have to take into consideration atoms and molecules not being spherically symmetric and that, as a result, some orientations of the molecules are more likely to result in reaction than others. In addition, a head-on collision is more likely to result in reaction than a glancing blow, so the collision cross section will be less than $\pi(r_N + r_{O_2})^2$. These factors can, however, be accounted for by multiplying by a constant, called a *stearic factor*, so the form of our equation, and the temperature dependence, would not be affected. Values of stearic factors for various reactions range over many orders of magnitude and can be quite small. In rare circumstances, they can be greater than 1 (implying an effective collision cross section greater than the combined atomic radii).

^{*} So-called because Maxwell proposed it and Boltzmann proved it rigorously.

⁺ To distinguish the rate constant, *k*, from Boltzmann's constant, k, we will always write the former in lower case italics and the latter in roman typeface.

W. M. White Chapter 5: Kinetics

Temperature occurs in 2 places in equation 5.22; however, the square-root dependence is slight compared to the exponential one. For example, consider a temperature change of 300 K to 325 K. For a reaction with an activation energy of 25 kJ, the exponential temperature dependence results an increase in reaction rate of more than a factor of 2, whereas the square root dependence increases the reaction rate by only 4%. Hence the temperature dependence can be essentially expressed as:

$$k \propto e^{-E_B/\mathbf{k}T}$$
 ‡

The temperature dependence of the rate constant is most often written as:

$$k = A e^{-E_B / kT}$$
 5.24

which is the important *Arrhenius relation*[¶]. It expresses the rate constant in terms of the barrier, or *activation*, energy (also often written as E_A or E^*), and A, a proportionality constant sometimes called the *frequency factor* (because it depends on the frequency of collisions), and temperature. (We can replace k, Boltzmann's constant with R, the gas constant, if we deal in moles rather than atoms.)





Fig. 5.3. a. Relative change in the reaction rate as a function of activation energy at 300 K. b. Change in the reaction rate for the same as a function of temperature with an activation energy of 15 kJ.

constant is illustrated in Figure 5.3. We see that the reaction rate falls off by a factor of 10^2 as temperature is decreased from 500 to 200 K. This confirms our everyday experience that reaction rates are extremely temperature dependent. Table 5.1 lists some examples of activation energies for geochemical reactions.

The pre-exponential factor, *A*, is often assumed to be independent of temperature. Comparison of 5.24 with 5.23 shows, however, that it need not be. In the case of an elementary gas phase reaction, we would predict a dependence on the square root of temperature. Other kinds of reactions show other kinds temperature dependencies of the frequency factor, however. A more accurate expression of temperature dependence of the reaction rate is:

$$x = AT^n e^{-E_B/kT}$$
 5.25

where the exponent *n* can be any number. Nevertheless, the temperature dependence of the frequency factor is usually small and it can often be safely neglected, as in our example above.

[‡] This form of the temperature dependence of reaction rate was first proposed by Jacobus H. Van't Hoff (1852-1911), who deduced it by analogy to equation 3.96, the *Van't Hoff equation*. Van't Hoff was born in Rotterdam and in 1878 was appointed professor of geology, mineralogy, and chemistry at the University of Amsterdam. In 1896 he moved to the University of Berlin, where he remained for the rest of his life. He won the Nobel Prize for Chemistry in 1901.

[¶] Named for Svante August Arrhenius (1859-1927) because Arrhenius provided the theoretical justification for Van't Hoff's proposal. Arrhenius's PhD dissertation, completed in 1884 at the University of Uppsala in Sweden, was rated fourth class by the committee of examiners, implying great things were not expected of him. The old boys must have been a little surprised nineteen years later when Arrhenius won the Nobel Prize for chemistry. Among Arrhenius's other contributions were the ionic theory of electrolyte solutions and the greenhouse theory of climate: that the CO_2 concentration in the atmosphere could be an important control on global temperature, and that anthropogenic burning of fossil fuel could lead to global warming.

W. M. White

CHAPTER 5: KINETICS

5.3.2.3 A General Form of the Rate Equation

In general, the rate of a reaction such as:

$$aA + bB \rightarrow cC + dD$$

$$\Re = ka_A^{n_A} a_B^{n_B} a_C^{n_C} a_D^{n_D}$$

$$5.26^{\pounds}$$

can be expressed as:

where *k* is the *rate constant* and a_A , etc. are activities (we will often use the simplifying assumption of ideality and replace these by concentrations). The exponents n_A , n_B , etc. can be any number, including zero. The sum of the exponents n_A , n_B , ... is the *order of the reaction*. In general, the value of the exponents must be determined experimentally, though their values can be predicted if the reaction mechanism is known, as we saw in the above example.

Just as the mole fraction was the unit of choice for thermodynamics, *moles per volume, or moles per area in the case of reactions taking place on surfaces, is the unit of choice for kinetics.* Thus wherever more than one phase is involved, one concentration should be expressed in moles per unit area or volume.

There are several simplifications of equation 5.26 for *elementary reactions*. First, the rate of reaction is independent of the concentration of the products so the exponents of the products will be 0. Indeed, *one of the criteria for an elementary reaction is that the product not influence the reaction rate*. Second, *the values of the exponents for the reactants are the stoichiometric coefficients of the reactant species*. Thus if the reaction can be written in terms of a series of elementary reactions, then the exponents for the rate equation can be deduced from those of the component elementary reactions.

A further simplification may be made where one of the reactants is in sufficient abundance that its

Example 5.1. Rate of Hydration of CO_{2(aq)}

The rate for the hydration of CO₂ (i.e., CO₂ + H₂O \rightleftharpoons H₂CO₃) has been found to follow the first order rate law:

$$-\frac{d[\text{CO}_{2(aq)}]}{dt} = k[\text{CO}_{2(aq)}] \quad 5.27$$

At 25° C, *k* has been determined to be 0.014 sec⁻¹. Make a graph showing how the concentration of CO₂ will change with time as the reaction proceeds, assuming an equilibrium (i.e., final) CO₂ concentration of 0.

Answer: Since we are not given the absolute concentrations, we cannot determine the absolute change. We can, however, determine relative change. To do so, we just integrate 5.27:

$$-\int \frac{d[\operatorname{CO}_{2(aq)}]}{[\operatorname{CO}_{2(aq)}]} = k \int_{0}^{t} dt$$

With some rearranging, we obtain:

$$\frac{[\text{CO}_{2(aq)}]}{[\text{CO}_{2(aq)}]_0} = e^{-kt}$$
 5.28



Figure 5.4. Progress in the reaction $CO_{2(aq)} + H_2O \rightarrow H_2CO_3$ with time, as measured by decrease in $[CO_2]$.

Figure 5.4 shows our result. It is apparent that this is a fast reaction. We can assume that equilibrium will prevail on most time scales of interest to us.

 $^{^{\}pounds}$ Don't confuse this equation, which expresses the way in which reaction rates depend on concentrations, with equation 5.9, which is the definition of the reaction rate.

W. M. White Chapter 5: Kinetics

$$CO_2 + H_2O \rightarrow H_2CO_3$$

The rate of this reaction will be:

$$-\frac{d[\mathrm{CO}_2]}{dt} = k[\mathrm{CO}_2][\mathrm{H}_2\mathrm{O}]$$

which is a second order reaction. However, in aqueous solution, H_2O will always be present in great excess over CO_2 and its abundance will not be significantly changed by this reaction. This allows us to treat the reaction as if it were first order and to define a *pseudo-first order* rate constant, k^* , as:

$$k^* = k [H_2O]$$

Since $[H_2O]$ is constant, it follows that k^* is as well. The reaction rate can then be written as:

$$\frac{d[\mathrm{CO}_2]}{dt} = k^*[\mathrm{CO}_2]$$

In Examples 5.1 and 5.2 we have used just such a *pseudo-first order* rate constant.

5.2.4 RATES OF COMPLEX REACTIONS

Deciding whether a reaction is elementary is not always straightforward. Consider the reaction:

$$2NO_2 \rightarrow 2NO + O_2 \qquad 5.29$$

On a microscopic basis, we might describe this reaction as the collision of two NO_2 molecules to form two NO molecules and an O_2 molecule. Since no intermediate steps occur, this would appear to be an elementary reaction. The rate equation for this reaction has been experimentally determined to be:

$$-\frac{d[\mathrm{NO}_2]}{dt} = 2k[\mathrm{NO}_2]^2$$

This has the predicted form for an elementary reaction of second order; thus experiment confirms that reaction 5.29 is elementary.

Now consider the reaction:
$$2O_3 \rightarrow 3O_2$$

We might reason that this reaction requires only the collision of two ozone molecules with no intermediate products and that the reaction is therefore primary. However, the experimentally determined rate law is:

$$\frac{1}{3}\frac{d[O_2]}{dt} = -\frac{1}{2}\frac{d[O_3]}{dt} = k\frac{[O_3]^2}{[O_2]}$$
5.30

Since the rate depends on the concentration of the product, the reaction is not elementary and must involve intermediate steps.

5.2.4.1 Chain Reactions and Branching

Many overall reactions involve a series of sequential elementary reactions, or steps, each of which must be completed before a subsequent reaction can occur. Such reactions are termed *chain reactions*. It

162

TABLE 5.1. ACTIVATION ENERGIES OF SOME GEO-CHEMICAL REACTIONS

Geochemistry

Reaction	E _A
	kJ/mol
$Mg_{3}Si_{4}O_{10}(OH)_{2} \rightarrow 3MgSiO_{3} + SiO_{2} + H_{2}O$	371.8
$CaCO_3 + SiO_2 \rightarrow CaCO_3 + CO_2$	225.0
$2CaCO_3 + Mg^{2+} \rightarrow (CaMg)CO_3 + Ca^{2+}$	$\frac{1}{4}\frac{4}{100_2} = k[C$
$NaAlSi_2O_6H_2O + SiO_2 \rightarrow NaAlSi_3O_8 + H_2O$	106 <i>.</i> \$t
$C_2H_4 + H_2 \rightarrow C_2H_6$	102.8
$CaF_2 \rightarrow Ca^{2+} + 2F^+$	73.0
$MgSiO_3 + 2H^+ + H_2O \rightarrow Mg^{2+} + H_4SiO_4$	49.0
$SiO_{2(qz)} + 2 H_2O \rightarrow H_2SiO_4$	40.6
$SiO_{2(am)} + 2 H_2O \rightarrow H_2SiO_4$	35.8
$H_2SiO_4 \rightarrow SiO_{2(qz)} + 2 H_2O$	28.4
$Mg_2SiO_4 + 4H^+ \rightarrow 2Mg^{2+} + H_2SiO_4$	21.7
$CaCO_3 \rightarrow Ca^{2+} + CO_3^{2+}$	20.1
$O + O_3 \rightarrow 2O_2$	13.4

рΚ

4.5

Oxidation Rate Constants

7.9 ×

25

7.9 ×

 Fe^{2+} Species k

7.4

 (s^{-1})

 10^{-6}

 10^{6}

Equilibrium Constants

 $Fe^{2+} + H_2O \rightleftharpoons FeOH^+ + H^+$

 $FeOH^+ + H_2O \rightleftharpoons Fe(OH)_2 + H^+$

Fe²⁺

FeOH⁺

Fe(OH)₂

Reaction

CHAPTER 5: KINETICS

Example 5.2: Oxidation of Ferrous Iron

Given the adjacent equilibrium and pseudo-first order rate constants for the oxidation of three species of ferrous iron (Fe²⁺, Fe(OH)⁺, and Fe(OH)₂) to ferric iron in the adjacent table, calculate the o verall rate o f oxidation of ferrous iron at pH 2, 6, and 8 assuming a total Fe²⁺ concentration of 1×10^{-6} M.

Answer. The overall oxidation fate can be written as	Answer:	The overall	oxidation	rate can	be written	as:
--	---------	-------------	-----------	----------	------------	-----

$$\frac{d\Sigma Fe^{2+}}{dt} = k_1 [Fe^{2+}] + k_2 [FeOH^-] + k_3 [Fe(OH)_2] \qquad 5.31$$

at Thus to calculate the rate, we will have to calculate the concentrations of the various species. These are given by:

$$[FeOH^{-}] = K_1 \frac{[Fe^{2+}]}{[H^{+}]}$$
 and $[Fe(OH)_2] = K_2 \frac{[Fe^{2+}]}{[H^{+}]^2}$

We can substitute these expressions into 5.32 we have:

$$\frac{d\Sigma Fe^{2+}}{dt} = [Fe^{2+}] \left(k_1 + \frac{k_2 K_1}{[H^+]} + \frac{k_2 K_2}{[H^+]^2} \right)$$
 5.32

Since the total Fe^{2+} is the same at all three pH's the concentration of the Fe^{2+} ion must vary. So we need to calculate the concentration of ionic Fe^{2+} at these pH's. The total Fe^{2+} is:

$$\Sigma Fe^{2+} = [Fe^{2+}] + [FeOH^{-}] + [Fe(OH)_{2}]$$

or:

so that:

$$\Sigma F e^{2+} = [F e^{2+}] \left(1 + \frac{K_1}{[H^+]} + \frac{K_2}{[H^+]^2} \right)$$
$$[F e^{2+}] = \frac{\Sigma F e^{2+}}{1 + \frac{K_1}{[H^+]} + \frac{K_2}{[H^+]^2}}$$

We can now calculate the rates. Substituting in appropriate values into equ. 5.32, we find the rate is 0.0031 M/sec, 0.9371 M/sec, and 7.89 M/sec at a pH of 2, 4, and 8 respectively. Thus the combination of the different rate constants and the pH dependency of the Fe speciation results in a very strong pH dependence of the oxidation rate.

is also possible that the path of an overall reaction may include two or more alternative elementary reactions, or sequences of elementary reactions, that occur simultaneously. These alternatives paths are called *branches*. The combustion of hydrogen is a good example because it is a chain reaction involving several branches.

Experiments have shown that the reaction rate for the combustion of hydrogen is not simply:

$$\frac{d[H_2O]}{dt}] = k[H_2]^2[O_2]$$
5.33

and therefore $2H_2 + O_2 \rightarrow 2H_2O$ is not an elementary reaction. If it were an elementary reaction, equation 5.33 predicts that its rate should continuously decrease through the course of the reaction (provided temperature is held constant!) since the reactants will be consumed and their concentrations will decrease. In actuality the rate of this reaction can increase rapidly, sometimes catastrophically (even at constant temperature), as it proceeds. Evidently, the reaction mechanism is more complex. Indeed it appears to involve several steps. The final step of this reaction is:

$$OH + H_2 \rightarrow H_2O + H$$
 5.34

This is an elementary reaction, depending only on the concentration of the two reactants. However, one of the reactants, OH, and one of the products, H, are not among the original constituents of the gas.

CHAPTER 5: KINETICS

W. M. White

Rather, they are created by intermediate steps. Species that do not appear in the overall reaction are termed *reactive intermediates*.

The first step in the combustion of hydrogen is breakup of the hydrogen molecule, forming highly reactive atomic H:

$$H_2 \rightarrow 2H$$
 5.34a

The next step is reaction of the atomic hydrogen with an oxygen molecule:

$$H + O_2 \rightarrow OH + O$$
 5.34b

Reactions 5.34a and 5.34b are an example of a *chain reaction*.

Since 5.34 is an elementary reaction, the reaction rate can be written as:

$$\frac{d[\mathrm{H}_{2}\mathrm{O}]}{dt} = k[\mathrm{OH}][\mathrm{H}_{2}]$$

This is also the rate of the overall reaction. Thus the overall reaction will depend on the availability of OH. What makes the combustion of hydrogen particularly interesting is that there are several ways in which OH may be created. Reaction 5.34b is one way. The monatomic oxygen created in this reaction, however, provides two alternate mechanisms for the creation of the OH complex:

$$H_2 + O \rightarrow OH + H$$
 5.34c

and
$$H + O \rightarrow OH$$
 5.34d

Reactions 5.34b through 5.34d represent alternative reaction paths or *branches*. Notice that the final step also provides an alternative mechanism, or branch, for the production of monatomic hydrogen.

The branching that occurs provides the potential for a "runaway" or explosive reaction. This is apparent if we simply sum reactions 5.34 and 5.34b through 5.34d:

$$4 \times [OH + H_2 \rightarrow H_2O + H]$$

$$+ 2 \times [H + O_2 \rightarrow OH + O]$$

$$+ H_2 + O \rightarrow OH + H$$

$$+ H + O \rightarrow OH$$

$$5H_2 + 2O_2 \rightarrow 4H_2O + 2H$$
5.34e

Each cycle of these reactions produces four water molecules plus two hydrogens. Since the rate of the overall reaction, i.e., the production of water, depends on [OH], which in turn depends on [H], the reaction will accelerate with time. (Actually, the combustion of hydrogen is a *very* complex reaction. When all the elementary reactions are written down, including the reverse reactions and reactions with the container wall, they fill an entire page. Interestingly, it displays this runaway behavior only under certain combinations of T, P, and container size and shape. The latter dependence results from reactions with, or catalyzed by, the container wall. Under certain conditions, it will become steady state; i.e., the creation and consumption of water balance to produce a constant concentration of water.)

5.2.4.2 Rate-Determining Step

It often happens that the reaction rate of a chain, or sequential, reaction, is controlled by a single step that is very much slower than the other steps. For example, how quickly you can buy a pencil at the campus bookstore on the first day of class will probably be controlled entirely by how quickly you can get through the checkout line. Such a step is called the *rate-determining step*. Once the rate of this step is determined, the rates of all other steps are essentially irrelevant.

Now consider a reaction that can occur through two branches. For example,

$$A \xrightarrow{1} B$$
 and $A \xrightarrow{2} B$
te is then: $\frac{d[A]}{dt} = -(k_1 + k_2)[A]$ 5.35

The reaction ra

Geochemistry

CHAPTER 5: KINETICS

If one path is very much faster than the other, then the fastest of the two will always be taken. Thus for branched reactions, the fastest branch determines the reaction mechanism. Mathematically, we may say that if $k_1 >> k_2$ then $(k_1 + k_2) \approx k_1$ and therefore:

$$\frac{d[A]}{dt} \approx -k_1[A]$$

In our analogy above, if an express checkout is available, you would certainly take it. In this case, the slowness of the regular checkout line becomes irrelevant for determining how quickly you can buy your pencil. To sum up, we may say that when reactions occur *in series*, then *the slowest reaction is the rate determining step*. *When parallel*, or branched, reaction paths are available, then *the fastest path is rate-determining*.

5.2.5 Steady-State and Equilibrium

Many geochemical systems are steady state ones, that is, time-invariant systems, or approximately so. The equilibrium state is also a steady state, but not all steady state systems are necessarily equilibrium ones. We may say then that steady state is a necessary, but not sufficient, condition for *equilibrium*. Let's consider how a system will approach the steady-state and equilibrium.

Consider the elementary reaction: $A \rightarrow B$

Suppose that this reaction does not entirely consume A, but reaches a steady state where the concentration of A is $[A]_s$, the subscript *s* denoting the steady state. In this case, we can express the reaction rate as:

$$\frac{d[A]}{dt} \approx k([A]_s - [A])$$
5.36

where $[A]_s$ is the steady state concentration of A. The reaction rate is 0 when $[A] = [A]_s$.

To see how the concentration will vary before steady state is achieved, we integrate 5.36:

$$\ln\left(\frac{[A]_{s}-[A]}{[A]_{s}-[A]^{\circ}}\right) = -kt$$

where [A]° is the initial concentration of A. This may be written as:

$$\frac{[A]_{s} - [A]}{[A]_{s} - [A]^{\circ}} = e^{-kt}$$
5.37

The denominator is a constant (for a given set of initial conditions), so we can rewrite 5.37 as:

$$[A]_{s} - [A] = Ce^{-kt}$$

The excess concentration of A, i.e., $[A]-[A]_s$ declines as e^{-t} , so that steady state is approached asymptotically. An effective steady state will be achieved when t >> 1/k. As in Example 5.1, the reaction rate decreases exponentially with time, i.e.:

$$\frac{d[A]}{dt} = kCe^{-kt}$$

Now suppose that in addition to the reaction: $A \rightarrow B$, the reaction $B \rightarrow A$ also occurs and that both are first order elementary reactions. The rates of reaction will be:

$$\frac{d[A]}{dt} = -k_{+}[A] + k_{-}[B]$$
5.38

Here we are using k_+ for the rate constant of the forward reaction and k_- for the rate constant of the reverse reaction. Assuming the system is closed and that no other processes affect the concentrations of A and B, then:

$$[A] + [B] = \Sigma AB$$

where ΣAB is the total of A and B and is a constant. Equ. 5.38 can therefore be written as:

Chapter 5: Kinetics

M. White

$$\frac{d[A]}{dt} = -k_{+}[A] + k_{-}[\Sigma AB - A] = -(k_{+} + k_{-})[A] + k_{-}\Sigma AB$$
5.39

The concentration of at some time τ , $[A]_{\tau}$ is obtained by integrating 5.39:

$$\int_{A^{\circ}}^{A_{\tau}} \frac{d[A]}{-(k_{+}+k_{-})[A]+k_{-}\Sigma AB} = \int_{0}^{\tau} dt$$
$$\frac{-(k_{+}+k_{-})[A]_{\tau}+k_{-}\Sigma AB}{-(k_{+}+k_{-})[A]^{\circ}+k_{-}\Sigma AB} = e^{-(k_{+}+k_{-})\tau}$$

which yields:

Since $[A]^{\circ} + [B]^{\circ} = \Sigma AB$, we can also express this as:

$$\frac{-k_{+}[A]_{\tau} + k_{-}[B]_{\tau}}{-k_{+}[A]^{\circ} + k_{-}[B]^{\circ}} = e^{-(k_{+} + k_{-})\tau}$$
5.40

Thus in the general case, the concentrations of A and B will depend on their initial concentrations. However, for $\tau = \infty$, or as a practical matter when $\tau \gg (k_+ + k_-)$, then a steady-state will be achieved where 5.40 reduces to:

$$k_{+}[A]_{\infty} = k_{-}[B]_{\infty}$$
 5.41

Example 5.3: Racemization of Amino Acids

Amino-acids are nitrogen-containing organic molecules that are essential to life. The chemical properties of amino acids depend not only on their composition, but also on their structure. Twenty

different amino acids are used in building proteins. Amino acid comes in two forms, which can be distinguished by the direction in which they rotate polarized light. Interestingly enough, organisms synthesize only the form that rotates polarized light in a counterclockwise manner, labeled the L-form (Fig. 5.5a). After death of the organism, however, the amino acid can spontaneously convert to its mirror image, the D-form (Fig. 5.5b), corresponding to clockwise rotation of light. This process is termed racemization. Racemization is a first order reaction and rate constants for this process have been determined for a number of amino acids in various substances. This provides a means of dating sediment. Given that the rate constant k_{+} for the 1isoleucine \rightarrow d-alloisoleucine reaction is 1.2×10^{-7} y⁻¹ and for the d-alloisoleucine \rightarrow l-isoleucine is 9.6 \times 10⁸ y⁻¹, what is the age of a sediment whose d-alloisoleucine/l-isoleucine ratio is 0 .1? Assume that the total isoleucine is conserved and an initial disoleucine concentration of 0.

Answer: This is a special case of equation 5.40 where $[B]^{\circ}$ is 0 and $[A] + [B] = [A]^{\circ}$. L etting γ be the r atio [B]/[A] (d-alloisoleucine/l-isoleucine) and substituting into to 5.40, we obtain:

$$\frac{-k_{+} + \gamma k_{-}}{-k_{+}(1+\gamma)} = e^{-(k_{+}+k_{-})t}$$

Substituting values and solving for *t*, we find the age is 8.27×10^3 yr. Of course, racemization rates, like all reaction rates depend on temperature, so this result assumes constant temperature.



Figure 5.5. S tructure of Lisoleucine and D-alloisoleucine. Solid wedge shapes indicate bonds coming out of the plane of the paper, hashed wedge shapes indicate bonds behind the paper.



CHAPTER 5: KINETICS

5.3 Relationships between Kinetics and Thermodynamics

5.3.1 Principle of Detailed Balancing

Equation 5.41 describes the relation between the concentration of reactant and product of a reversible reaction after infinite time, i.e., in the steady state. This then is just the state the reaction will obtain in the absence of constraints and external disturbance. This is precisely the definition of equilibrium we decided on in Chapter 2. It follows that $[A]_{\infty}$ and $[B]_{\infty}$ are then also the equilibrium concentrations. Thus we see, as we stated in Chapter 2, that equilibrium is not necessarily a static state on the microscopic scale. Rather, it is a steady state where the forward rate of reaction is equal to the reverse rate. Formally, we may say that for an elementary reaction such as:

$$A \rightleftharpoons B$$

at equilibrium the following relation must hold:

$$k_{+}[A]_{eq} = k_{-}[B]_{eq}$$
 5.42

where k_+ and k_- are the rate constants for the forward and reverse reactions respectively. This is known as the *principle of detailed balancing*, and it establishes an essential link between thermodynamics and kinetics. This link is apparent when we combine equation 5.42 with equation 3.85 to obtain:

$$\frac{k_{+}}{k_{-}} = \frac{[B]_{eq}}{[A]_{eq}} = \mathbf{K}^{app}$$
5.43

It is apparent from equation 5.43 that if the equilibrium constant and one of the rate constants for a reaction are known, the rate constant for the reverse reaction may be deduced. Furthermore, if the form of the rate law for either the forward or reverse reaction is known, the other can be deduced. This is a trivial point for elementary reactions since rate laws for such reactions are readily obtained in any case. The importance of this point is that it holds for overall reactions as well as elementary ones. For example, consider the serpentinization of olivine:

$$2MgSiO_4 + H_2O + 2H^+ \rightleftharpoons Mg_3Si_2O_5(OH)_4 + Mg^{2+}$$

This is not an elementary reaction as several intermediate steps are involved, as in the example of the weathering of anorthite discussed earlier. Nevertheless, if olivine, serpentine and water can be assumed to be pure phases and have unit activity, the equilibrium constant for this reaction is:

$$K^{app} = \frac{[Mg^{2+}]}{[H^{+}]^{2}}$$
 5.44

The relation between the forward and reverse reaction rate constants must be:

$$k_{\rm [Mg^{2+}]=k_{\rm +}[{\rm H^+}]}$$

Suppose that experiments show that the rate law for the forward reaction is:

$$\frac{d[Mg^{2+}]}{dt} = k[Ol][H^+]$$

where [*Ol*] is the specific area (area per solution volume) of olivine in the experiment. From equation 5.9, we can express the rate for the reverse reaction as:

$$\frac{d[Ol]}{dt} = 2\frac{d[Mg^{2+}]}{dt} = k[Ol][H^+]$$

Using equation 5.43 to obtain a substitution for *k*, we find that the rate law for the reverse reaction, i.e., for the formation of olivine from serpentine must be:

$$\frac{d[Ol]}{dt} = 2k' \frac{[Mg^{2+}]}{[H^+]^2} [Ol][H^+] = 2k' \frac{[Mg^{2+}]}{[H^+]} [Ol]$$
 5.45

where k' is the rate constant for the reverse reaction.

© W. M. White 2005

October 11, 2005

Geochemistry

CHAPTER 5: KINETICS

5.3.2 Enthalpy and Activation Energy

The principle of detailed balancing allows us to relate the activation energy in the Arrhenius relation (equation 5.24) to the heat (enthalpy) of reaction. Recall that the temperature dependence of the equilibrium constant was:

$$\ln \mathbf{K} = -\frac{\Delta H_r^o}{RT} + \frac{\Delta S_r^o}{R}$$
(3.95)

If pressure is held at the reference pressure, then the *van't Hoff* equation states:

$$\left(\frac{\partial \ln K}{\partial T}\right)_{P} = \frac{\Delta H_{r}^{o}}{RT^{2}}$$
(3.96)

Integrating, we get:

$$\ln \mathbf{K} = \frac{-\Delta H_r^o}{RT} + \mathbf{K}^o$$
 5.46

where K° is the integration constant. Thus the temperature dependence of the equilibrium constant may be expressed as:

$$\mathbf{K} = \mathbf{K}^{\circ} e^{-\Delta H_r^{\circ}/RT}$$
 5.47

A similarity to the Arrhenius relation is immediately apparent. If we write the Arrhenius relations for the forward and reverse reactions and combine them with equations 5.43 and 5.47, we obtain:

$$\frac{k_{+}}{k_{-}} = \frac{A_{+}e^{-E_{+}/RT}}{A_{-}e^{-E_{-}/RT}} = \mathbf{K} = \mathbf{K}^{\circ}e^{-\Delta H_{r}^{\circ}/RT}$$
5.48

(For simplicity and clarity, here, and in the subsequent discussion of Transition State Theory, we assume ideal behavior, and therefore that activities equal concentrations and that $K^{app} = K$.) From this we can see that:





$$\frac{A_+}{A_-} = K^\circ \qquad 5.49$$

and
$$E_+ - E_- = \Delta H_r^\circ \qquad 5.50$$

This relationship is illustrated in Figure 5.6.

Comparing equation 5.46 with 3.95, we find that the factor K° must be related to the entropy change of reaction, and further that the coefficients of the Arrhenius relation must be related to entropy by:

$$\frac{A_+}{A_-} = \mathrm{K}^\circ = e^{\Delta S_r^\circ / R} \quad 5.51$$

Indeed, it can be shown that Arrhenius coefficient,

or frequency factor, is related to entropy as:

$$A_{+} = \frac{kT}{h} e^{\Delta S_{+}^{*}/R}$$
 5.52

where $\Delta S_{+}^{'}$ is the entropy difference between the initial state and the activated state (discussed below) and *h* is Planck's constant. The ratio kT/h has units of time⁻¹ and is called the *fundamental frequency*.

5.3.3 Aspects of Transition State Theory

In the above discussion, we have already made implicit use of *Transition State Theory*. Transition State Theory postulates that an *elementary* reaction such as:

Geochemistry

CHAPTER 5: KINETICS

$$A + BC \rightarrow AC + B$$
 5.53

proceeds through the formation of an *activated complex* ABC*, also called a reactive intermediate. Thus reaction 5.70 can be described by the mechanism:

$$A + BC \rightarrow ABC^*$$
 5.53a and $ABC^* \rightarrow AC + B$ 5.53b

The activated complex ABC* is assumed to be in thermodynamic equilibrium with both reactants and products. Hence it is possible to define an equilibrium constant for 5.53a (assuming ideal behavior) as:

$$K^* = \frac{[ABC^*]}{[A][BC]}$$
$$\Delta G^* = -RT \ln K^*$$

as well as a free energy change:

and enthalpy and entropy changes: $\Delta G^* = \Delta H^* - T \Delta S^*$

Though do so here would take us too far afield, it can be shown from a statistical mechanical approach that the rate constant for 5.53 is:

$$k = \kappa \frac{\mathbf{k}T}{h} \mathbf{K}^*$$
 5.54

where kT/h is the fundamental frequency as we defined it above, and κ is a constant, called the transmission coefficient, whose value is often close to 1. Equation 5.54 is known as the *Eyring Equation*[£]. It is then easily shown that the rate constant is:

$$k = \frac{kT}{h} e^{\Delta S^* / R} e^{-\Delta H^* / RT} = \frac{kT}{h} e^{-\Delta G^* / RT}$$
 5.55

Thus if the nature of the activated complex is understood, the rate constant can be calculated. For example, we saw that the partition function is related to entropy and energy (it is also easily shown that is related to Gibbs Free Energy and enthalpy as well). The rate constant can be calculated from partition functions of the activated complex and reactants.

Now consider that reaction 5.53 is reversible so that:

$$A + BC \rightleftharpoons AC + B$$

and that the reverse reaction proceeds through the same activated complex ABC*.

The net rate of reaction is: $\Re_{net} = \Re_{+} - \Re_{-}$

If ΔG is the free energy difference between product and reactant, then the free energy difference between the product and the activated complex must be $\Delta G - \Delta G^*$. From this it is readily shown (Problem 5.3) that the ratio of the forward and reverse reaction rates is:

$$\frac{\mathfrak{R}_{+}}{\mathfrak{R}_{-}} = e^{-\Delta G/RT}$$
 5.57

where ΔG is the actual free energy difference between products and reactants. The negative of ΔG in this context is often called the *affinity of reaction, reaction affinity,* or simply *affinity* and is designated A_r (for clarity, however, we shall continue to designate this quantity as ΔG). Substituting 5.57 into 5.56 and rearranging, we have:

$$\mathfrak{R}_{net} = \mathfrak{R}_{+} (1 - e^{\Delta G/RT)}$$
 5.58

If the forward reaction is an elementary one, then \Re_+ will be:

5.56

 $^{^{\}pounds}$ Named for Henry Eyring (1901-1981) who formulated transition state theory in 1935. It was evidently an idea whose time had come, because M. G. Evans and M. Polanyi independently developed the same theory in a paper published the same year. Eyring, who was born in Juarez, Mexico, received his PhD from the University of California at Berkeley in 1929. He worked in the University of Wisconsin, the Kaiser Wilhelm Institut in Berlin (working with Polanyi) and Princeton University before becoming professor of chemistry at the University of Utah in 1946, where he remained for the rest of his life.

CHAPTER 5: KINETICS

W. M. White

$$\Re_{+} = k_{+}[A][BC]$$

where k will be as defined in 5.24.

It must be emphasized that *equations* 5.57 *and* 5.58 *apply to elementary reactions only*. However, a similar equation may be written for *overall* reactions:

$$\Re = k(1 - e^{n\Delta G/RT})[A]^{n_A}[B]^{n_B}\dots$$
5.59

where the *n*'s can be any real number, and [A], [B], ... are the concentrations (surface areas for solids) of the reactants (Lasaga, 1981b). Using the Arrhenius expression for *k* (equ. 5.24), equation 5.59 becomes: $\Re = Ae^{-E_A/RT} (1 - e^{n\Delta G/RT}) [A]^{n_A} [B]^{n_B} \dots 5.60$

Equation 5.60 links kinetics and thermodynamics and forms the basis of *irreversible thermodynamics*.

If the system is not far from equilibrium, then $\Delta G \ll RT$ and we may approximate e^x by 1 + x, so that for an elementary reaction:

$$\Re_{net} \approx -\frac{\Re_+ \Delta G}{RT}$$
 5.61

Thus close to equilibrium, the reaction rate will vary linearly with ΔG , slowing as equilibrium is approached. Substituting A_r for $-\Delta G$, 5.61 can also be written as:

$$\Re_{net} \approx -\frac{\Re_+ A_r}{RT}$$
 5.62

At this point, you might think, "this is all fine and good, but how do I calculate ΔG ?" There are several approaches to estimating the value of ΔG under non-equilibrium conditions. For the first method, let's return to the relationship between activities, ΔG° , and *K*. In Chapter 3, we found we could express the relationship between chemical potential and activities *at equilibrium* as:

$$\sum_{i} v_i \mu_i^o + RT \ln \prod_{i} a_i^o = 0 \tag{3.84}$$

At equilibrium, the first term on the left is ΔG° and the second term is *RT* ln K. *Under non-equilibrium conditions*, however, the product of activities will not be equal to K and equation 3.84 will not be equal to 0. Rather, it will have some finite value, which is ΔG . We define a quantity Q as:

$$Q \equiv \prod_{i} a_{i}^{o}$$
 5.63

Q is called the *reaction quotient* (Chapter 3). Though equation 5.63 has the same form as our definition of the equilibrium constant (equ. 3.85), there is an important difference. K defines the relationship between activities *at equilibrium*. In defining *Q*, we impose no such condition, so that *Q* is simply the product of activities. At equilibrium Q = K, but not otherwise. Under non-equilibrium conditions, we can express equation 3.86 as:

$$\Delta G^{\circ} + RT \ln Q = \Delta G \qquad 5.64$$

 \mathbf{O}

Since ΔG° is equal to $-RT \ln K$, equ. 5.64 can be written as:

$$RT \ln Q - RT \ln K = \Delta G$$
 or: $\frac{Q}{K} = e^{\Delta G/RT}$ 5.65

Substituting 5.65 into 5.58, we have for an elementary reaction:

$$\Re_{net} = \Re_{+}(1 - \frac{Q}{K})$$
 5.66

Thus we expect reaction rates to decrease as $Q \rightarrow K$.

To arrive at the second method of estimating ΔG , we recall that ΔG may be written as $\Delta H - T\Delta S$. At equilibrium:

$$\Delta G_{eq} = \Delta H_{eq} - T_{eq} \Delta S_{eq} = 0$$

CHAPTER 5: KINETICS

Example 5.4: Estimating ΔG^* for the Aragonite-Calcite Transition

Aragonite is the high-pressure form of $CaCO_3$. Upon heating at 1 atm, it will spontaneously revert to calcite. Carlson (1980) heated aragonite crystals containing calcite nuclei to a series of temperatures for fixed times on the heating stage of a microscope, then measured the growth of the calcite nuclei and from that calculated growth rates shown in the adjacent table. Using these data, determine the value of ΔG^* for this reaction.

Answer: This is a reversible reaction, so we have to consider that both the forward and reverse of the aragonite \rightarrow calcite reaction will occur. According to transition state theory, the rate

constant for the forward reaction is

$$k_{+} = \frac{kT}{h} e^{-\Delta G^{*}/RT} \qquad (5.55)$$

From equ. 5.60, the rate of the net reaction is:

$$\Re_{net} = \frac{kT}{h} e^{-\Delta G^{*/RT}} \left(1 - e^{\Delta G/RT}\right)$$

This rate expression has units of time⁻¹, but Carlson's results are given in units of distance/time. How do we reconcile these? We might guess in this case that fundamental frequency, the pre-exponential term, ought to be multiplied by some sort of *fundamental distance*. A fundamental distance is this case would be lattice spacing, which for aragonite is about 5 Å (or 5×10^{-10} m). Thus if λ is the lattice spacing, we have

$$\Re_{net} = \frac{\lambda kT}{h} e^{-\Delta G^{*/RT}} \left(1 - e^{\Delta G/RT}\right) \text{m/s}$$

Solving for ΔG^* :

$$\Delta G^* = -RT[\ln \Re_{net} - \ln \frac{\lambda kT}{h} - \ln(1 - e^{\Delta G/RT})]$$

To determine ΔG^* , we have to calculate ΔG , which we can do using the thermodynamic data in Table 2.2 and equat ion 2.133. Our spreadsheet is shown below. Calculating the average ΔG^* for the 4 measurements, we find $\Delta G^* = 161$ kJ. We can then use ΔG^* to predict the reaction rates. A comparison between the measured and predicted reaction rates is shown in Figure 5.8.

R	8.314	J/mol-K		h	6.63E-34	J-sec		
k	1.38E-23	J/K		λ	5.00E-10	m		
lk/h	1.04E+01	m/K-sec						
					ln(1-exp			
Т°С	R m/sec	ln R	Т, К	ΔG, J	$(\Delta G/RT))$	ln(lkT/h)	∆G* kJ	-In Rcalc
455	7.45E-09	18.7152	728	-2828.8	-0.985	8.933	161.38	18.6485
435	3.63E-09	19.4352	708	-2709.3	-0.997	8.906	160.95	19.4397
415	1.61E-09	20.2452	688	-2592	-1.010	8.877	160.80	20.2756
395	6.24E-10	21.1952	668	-2476.6	-1.022	8.847	161.17	21.1605
375	2.72E-10	22.0252	648	-2363	-1.035	8.817	160.58	22.0986
						ΔG^*	160.98	ave

Aragon Transiti	ite to Calcite on Rates
T ℃	R(m/sec)
455	7.45×10^{-09}
435	3.63×10^{-09}
415	1.61×10^{-09}
395	6.24×10^{-10}
375	2.72×10^{-10}



Figure 5.8. Comparison of observed and predicted rates of the aragonite \rightarrow calcite reaction. Data (circles) from Carlson (1980).

Geochemistry

CHAPTER 5: KINETICS

where the subscript *eq* denotes the quantity when products and reactants are at equilibrium. Under non-equilibrium conditions, $\Delta H - T\Delta S$ will have some finite value. We can make use of this and write:

$$\Delta G = \Delta H - T\Delta S - (\Delta H_{eq} - T_{eq}\Delta S_{eq}) = \Delta H - \Delta H_{eq} - (T\Delta S - T_{eq}\Delta S_{eq})$$
 5.67

For temperatures close to the equilibrium temperature, ΔH and ΔS may be considered constant, i.e., independent of temperature, so that 5.67 simplifies to:

$$A_r = -\Delta G = \Delta T \Delta S \qquad 5.68$$

where $\Delta T = (T - T_{eq})$ and is sometimes called the *temperature overstep*. This may be substituted into 5.61, so that close to equilibrium we have

$$\Re_{net} = \frac{-\Re_+ \Delta S(T - T_{eq})}{RT_{eq}}$$
 5.69

Wood and Walther (1983) used this equation to analyze experimental reaction rate studies of a variety of silicate-aqueous fluid reactions. They found that essentially all the experimental data could be fit to this equation if \Re_+ is given by:

$$\Re_{\perp} = -k A$$

where A is the surface area of the solid phase and *k* is the rate constant. Furthermore, the temperature dependence of the rate constant could be expressed as:

$$\log k = -2900/T - 6.85$$
 5.70

This is illustrated in Figure 5.7. The data show a scatter of $>\pm 1$ order of magnitude about the line, so clearly the equation cannot be used for exact prediction of reaction rates. As Kerrick et al. (1991) point out, this approach has limits and cannot be applied to reactions involving carbonates. Nevertheless, Wood and Walther's work provides a useful way to estimate the order of magnitude of silicate dissolution rates.



Figure 5.7. Log of the rate constant vs. inverse of temperature for a variety of silicate and aluminate dissolution reactions. Wood and Walther (1983) extracted reaction rate data from both studies of the rates of mineral dissolutions (labeled "Dissolution" and phase equilibria studies (labeled "Phase Equilibria Studies"). Notice that the rate constant has units of mole of oxygen per cc per second. From Wood and Walther (1983).

Geochemistry

CHAPTER 5: KINETICS

5.4 Diffusion

We can't stir geochemical reactions. However, nature often provides advection to transport components. In nature, the driving force of advection is gravity: fluids (including the mantle, which behaves as a fluid on geological time scales) move upward or downward because they are lighter or heavier than their surroundings. Just as it does when we stir the tea, advection serves to transport reactants and therefore speed reactions. However, advection is rarely effective on very small scales. On these scales, diffusion is usually the process responsible for transport of chemical components. Except in gases, diffusion is too slow to transport components more than a few meters (and generally less). Consequently, chemical transport generally involves both diffusion and advection: advection for large-scale transport and diffusion for small-scale transport. In this section, we discuss the nature of diffusion and develop the tools necessary to treat it. Because advection depends on the physical properties of materials (density, viscosity), we will not treat it here.

5.4.1 Diffusion Flux and Fick's Laws

Fick's Law[§], or *Fick's First Law*, states that at steady state, the *flux*, J, of some species through a plane is proportional to the concentration gradient normal to that plane:

$$J = -D\frac{\partial c}{\partial x}$$
 5.71

The minus sign indicates diffusion is toward the region of lower concentration. The proportionality coefficient, D, is the *diffusion coefficient*. J has units of mass/area-time, e.g., moles/ m^2 -sec. If concentration is expressed per unit volume, as is often preferred in kinetics, the diffusion coefficient has units of m^2 -sec⁻¹. The diffusion coefficient must be empirically determined and will depend on the nature of the diffusing species, the material properties of the system in which diffusion is taking place and, as usual, temperature.

Strictly speaking, equation 5.71 is applicable to diffusion in only one dimension. A more general expression of Fick's First Law, applicable in 3 dimensional space is:

J =
$$-D\nabla \mathbf{c}$$

T c is: $\vec{\nabla}c = \frac{\partial c}{\partial x}\mathbf{x} + \frac{\partial c}{\partial y}\mathbf{y} + \frac{\partial c}{\partial z}\mathbf{z}$

5.72

where ∇c is:

and x, y, and z are unit vectors in the respective directions. Provided the concentration gradient is unidirectional (this will be the case, for example, in an accumulating sequence of sediments), equation 5.72 can be reduced to equation 5.71 simply by choosing the x-direction to be the direction of the concentration gradient.

Now let's consider how concentration will change with time as a consequence of diffusion. Imagine a volume enclosed in a cube of dimension dx (Fig. 5.10). Further imagine a diffusion flux of a species of interest through the plane and into the volume at xand a flux out of the volume through the plane at x+dx. Suppose n_x atoms per second pass through the plane at x and n_{x+dx} atoms per second pass through



Figure 5.10. A volume of dimension dx with fluxes through the planes at *x* and dx.

^sNamed for Adolf Fick (1829-1901). Fick was born in Kassel, Germany and earned an MD from the University of Marburg in 1851. Fisk's interest in diffusion through cell membranes led him to formulate the laws that bear his name. It was actually the second law that was published first, in an 1855 paper titled *Über Diffusion*. Fick deduced it by analogy to Fourier's equation for thermal diffusion.

CHAPTER 5: KINETICS

M. White

the plane at x+dx. The fluxes at the two planes are thus $J_x = n_x/dx^2$ -sec and $J_{x+dx} = n_{x+dx}/dx^2$ -sec. Conservation of mass dictates that the increase in the number of atoms, dn, within the volume is just what goes in less what goes out. Over a time dt this will be $dn = (J_x - J_{x+dx})dt$. The change in the concentration over this time is just this change in the number of atoms per unit volume:

$$dc = \frac{(J_x - J_{x+dx})dt}{dx}$$

and the rate of change of concentration is:

$$\frac{dc}{dt} = -\frac{(J_{x+dx} - J_x)}{dx}$$

If we specify that we are interested in the change in concentration at some fixed point x and some fixed time t, in the limit of infinitesimal dt and dx, this equation can be written as:

$$\left(\frac{\partial c}{\partial t}\right)_{x} = -\left(\frac{\partial J}{\partial x}\right)_{t}$$
 5.73

Equation 5.73 is called the equation of continuity since it follows from mass conservation. Now since the flux is given by Fick's First Law, we can write:

$$\left(\frac{\partial c}{\partial t}\right)_{x} = -\left(\frac{\partial\left(-D\frac{\partial c}{\partial x}\right)}{\partial x}\right)$$

Simplifying, we arrive at Fick's Second Law:

$$\left(\frac{\partial c}{\partial t}\right)_{x} = D\left(\frac{\partial^{2} c}{\partial x^{2}}\right)_{t}$$
5.74

Equation 5.74 tells us that rate of change with time of the concentration at any point is proportional to the second differential of the diffusion profile. Fick's second law is illustrated in Figure 5.11.

5.4.1.1 Solutions to Fick's Second Law

There is no single solution (i.e., function expressing c(t,x)) for equation 5.74; rather there are a number of possible solutions, and the solution appropriate to a particular problem will depend on the boundary conditions. Let's consider a few of the simpler ones. In all cases, we assume that the system is uniform in composition in the *y* and *z* directions so diffusion occurs only in the *x* direction.

As a first case, consider a thin film of some diffusing species sandwiched between layers of infinite length having concentration c=0. This might represent a 'doped' layer in a diffusion experiment in the laboratory. In nature, it might represent a thin sedimentary horizon enriched in some species such as iridium (such as the iridium-enriched layer in many sediments at the Cretaceous-Tertiary Bound-



Figure 5.11. Three possible concentration gradients. In a, $\partial^2 c/\partial x^2 = 0$ and therefore $\partial c/\partial t = 0$. Thus for a gradient that is straight, the concentration at any point remains constant (even though there is diffusion along the gradient). This is therefore the steady state case. In case b, $\partial^2 c/\partial x^2 > 0$ and hence the concentration at any point increases with time. In case c, $\partial^2 c/\partial x^2 < 0$ and therefore the concentration at any point decreases with time. Both cases b and c will tend, with time, toward the steady state case, a.

W. M. White Chapter 5: Kinetics

ary). Diffusion will cause the species to migrate away from x = 0 as time passes. Mathematically, this situation imposes certain boundary conditions on the solution of 5.74. We take the position of enriched horizon to be 0, and we seek a solution to 5.74 such that at t = 0 c = 0 everywhere except x = 0. At some time t > 0, our function should have the property that c approaches 0 as x approaches infinity. We further require that the total amount of the species remain constant, i.e.:

$$M = \int_{-\infty}^{+\infty} c dx$$

where M is the total amount of substance in a cylinder of unit cross section and length *x*.

The solution is given by Crank (1975) as:

$$c(x,t) = \frac{M}{2(\pi D t)^{1/2}} e^{-x^2/4Dt} \qquad 5.75$$

Figure 5.12 shows how the concentration profile changes with time under these circumstances. It is interesting to note that these profiles are the same as those of a "normal" statistical distribution



Figure 5.12. Concentration profiles at 3 different times resulting from outward diffusion from an thin film of the diffusing species. Note that the area under the curve remains constant through time.

same as those of a "normal" statistical distribution error curve with a standard deviation $\sigma = (2Dt)^{1/2}$. Suppose a boundary condition is imposed that diffusion can occur only in the positive direction. We can treat this case as if the diffusion in the negative direction is reflected at the plane x = 0. The solution is obtained by superimposing the solution for the negative case on the positive one:

$$c(x,t) = \frac{M}{(\pi Dt)^{1/2}} e^{-x^2/4Dt}$$
 5.76

Now consider a situation where the diffusing species has an initial uniform concentration C_0 between x = 0 and $x = -\infty$, and 0 concentration between x = 0 and $x = \infty$. In the laboratory, this circumstance might arise if we place 2 experimental charges adjacent one and other: one having been 'doped' with the species of interest. In nature, a somewhat analogous situation might be a layer of fresh water overlying a formation brine in an aquifer, or river water overlying seawater in an estuary, or two adjacent crystals.

The solution to this case may be found by imagining the volume between x = 0 and $x = -\infty$ as being composed of an infinite number of thin films of thickness $\delta\xi$ (Fig, 5.13). The concentration of the diffusing species at some point x_p at time *t* is then the sum of the contributions of each imaginary thin film (Crank, 1975). The mathematical solution is obtained by integrating the contribution of all such films:

$$c(x,t) = \frac{C_o}{2(\pi D t)^{1/2}} \int_x^\infty e^{-\xi^2/4Dt} d\xi \qquad 5.77$$

or defining $\eta = \xi / 2\sqrt{Dt}$:

$$c(x,t) = \frac{C_o}{\pi^{1/2}} \int_{x/2\sqrt{Dt}}^{\infty} e^{-\eta^2} d\eta \qquad 5.78$$



Figure 5.13. An extended initial distribution can be thought of as consisting of an infinite number of thin films of thickness $\delta\xi$. The concentration of a diffusing species at some point x_p is the contribution from each film from distances x to infinity (after Crank, 1975).
W. M. White

CHAPTER 5: KINETICS

The integral in 5.78 may be written as:

$$\int_{\frac{x}{2\sqrt{Dt}}}^{\infty} e^{-\eta^{2}} d\eta = \int_{0}^{\infty} e^{-\eta^{2}} d\eta - \int_{0}^{\frac{x}{2\sqrt{Dt}}} e^{-\eta^{2}} d\eta \quad 5.79$$

This integral has the form of a standard mathematical function called the *error function*, which is defined as:

$$erf(x) = \frac{2}{\pi^{1/2}} \int_0^x e^{-\eta^2} d\eta$$
 5.80

Substituting 5.80 into 5.78 and since $erf(\infty) = 1$, equation 5.78 becomes:

$$c(x,t) = \frac{C_o}{2} \left\{ 1 - erf\left(\frac{x}{2\sqrt{Dt}}\right) \right\}$$
 5.81

Values for the error function may be found in mathematical tables. The error function is also a standard function in some spreadsheets such as Microsoft Excel^{TM*}. Alternatively, it may be approximated as:

$$erf(x) \cong \sqrt{1 - \exp(-4x^2/\pi)} \qquad 5.82$$



Fig. 5.14. Distribution of a diffusing species initially confined to $-\infty < x < 0$ at three times after diffusion begins.

Figure 5.15 shows how the concentration profile will appear at different times. Since erf (0) = 0, the profiles have the interesting property that $c = C_o/2$ at x = 0 at all times.

A similar approach can be used for a diffusing species initially confined to a distinct region, for example: -h < x < +h. Examples might be sedimentary or metamorphic layers of finite thickness or a compositionally zoned crystal (Example 5.6). Again, the layer is treated as a series of thin films, but the integration in equation 5.77 is carried out from -h to +h. The result is:

$$c(x,t) = \frac{C_0}{2} \left\{ erf\left(\frac{h-x}{2\sqrt{Dt}}\right) + erf\left(\frac{h+x}{2\sqrt{Dt}}\right) \right\}$$
 5.83

5.4.2 Diffusion in Multicomponent Systems and The Diffusion Coefficient

There are 2 important constraints on diffusion that we have thus far ignored. First, diffusion differs from other kinds of transport in that we specify that there is *no net transport of material across the bound-ary of interest*. If there is net transport, we are, by definition, dealing with flow or advection rather than diffusion. If this constraint is to be satisfied, movement of one species through a plane must be accompanied by movement of one or more other species in the opposite direction.

The second constraint is electrical neutrality. Diffusion of even small quantities of an ion will quickly lead to the development of a large electric potential. The force associated with the potential would prevent any further diffusion of that ion in that direction. Thus diffusion of an ionic species must be coupled with diffusion of an equal quantity of charge in the opposite direction. In addition to these constraints, we must recognize that diffusion in some cases will lead to non-ideal mixing and the finite enthalpy and volume changes that accompany such situations.

With this in mind, we can recognize four classes of situations and four kinds of diffusion coefficients:

1.) Tracer, or Self- Diffusion, in which the net mass and charge fluxes associated with the diffusing species is sufficiently small that they can be safely ignored. There maybe no significant concentration gradient. This situation occurs when, for example, an experimental charge is doped with a radioactive

^{*} The error function in Excel, ERF(), is an add-in function found among the "analysis tools". ERF() does not properly treat the case where x < 0. The error function has the property that erf(-x) = -erf (x). In working with Excel, test for a negative value of x and where x is <0 replace ERF (-x) with -ERF(x). IF functions in Excel have the format "IF(logical_test, value_if_true, value_if_false)". So, for example, us a statement such as "=IF(x<0,-ERF(-x),ERF(x))".

W. M. White

CHAPTER 5: KINETICS

isotope in sufficiently small amounts such that the concentration of the element, and hence its chemical potential, does not vary significantly. This is the simplest situation, and the one that we have dealt with thus far.

2.) Chemical Diffusion refers to non-ideal situations where chemical potential rather than concen-

Example 5.6. Diffusion in a Crystal

Igneous crystals are often zoned as a result of changes in the composition of the magma. Suppose an olivine crystal of 2 mm diameter with a concentration of 2000 ppm Ni suddenly comes in contact with a magma in which its equilibrium concentration should be 500 ppm Ni. How long would it take for diffusion to homogenize the crystal at a temperature of 1250° C, assuming instantaneous equilibration at the crystal-liquid boundary?

Answer: We can treat the olivine crystal as a sphere. Radial symmetry then allows us to consider the problem as a function of radius. We need only consider the variation of concentration along 1 radial direction with 0 < x < r. Our boundary condition is that at x = r (the edge of the crystal) concentration is held constant by reaction with the liquid. We'll call this concentration C_r . The initial distribution is $c = C_i$ for 0 < x < r. According to Crank (1975), the solution is:

$$\frac{c-C_i}{C_r-C_i} = 1 + \frac{2r}{\pi x} \sum_{n=1}^{\infty} \frac{-1^n}{n} \sin\left(\frac{n\pi x}{r}\right) \exp\left(\frac{-Dn^2 \pi^2 t}{r^2}\right) \qquad 5.84$$

and the concentration at x = 0, C_0 , is:

$$C_0 = (C_r - C_i) \left\{ 1 + 2\sum_{n=1}^{\infty} -1^n \exp\left(\frac{-Dn^2 \pi^2 t}{r^2}\right) \right\} + C_i \qquad 5.85$$

From equation 5.71, we see that as the concentration gradient disappears, the rate of diffusion goes to 0. So the crystal approaches homogeneity only asymptotically, becoming homogeneous only at $t = \infty$, but it will become *essentially* homogeneous more quickly. *Essentially* homogeneous implies we could not detect a gradient. If our analytical precision is only 5%, we could not detect a gradient of less than 5%. So let's rephrase the question to ask, how will it take before the concentration gradient is less than 5%? We set $C_0/C_r = 1.05$ and substituting into 5.85 and rearranging, we obtain:

$$0.05 \le \left(1 - \frac{C_i}{C_r}\right) 2\sum_{n=1}^{\infty} -1^n \exp\left(\frac{-Dn^2 \pi^2 t}{r^2}\right)$$
 5.86



a rim concentration fixed at 500 ppm.

As it turns out, for relatively large values of t (Dt/r² > 0.1), the summation converges within 0.05% after the first term, so that 5.86 may be approximated by:

$$t \cong \frac{-r^2}{D\pi^2} \ln\left(\frac{0.05}{2(C_i/C_r-1)}\right)$$

For the value of $D = 10^{-12} \text{ cm}^2/\text{sec}$ given by Morioka and Nagasawa (1991), we find that about 154 years is required before the olivine homogenizes. If the olivine spent less than this time in c ontact with the magma, we would expect it to be zoned in Ni concentration. Figure 5.15 shows how the concentration profile of Ni would vary with time.

W. M. White

Geochemistry

Chapter 5: Kinetics

tration must be considered. In this circumstance, Fick's Laws can be rewritten as:

$$J = -L\frac{\partial\mu}{\partial x} \qquad 5.87 \quad \text{and} \qquad \left(\frac{\partial c}{\partial t}\right)_t = L\left(\frac{\partial^2\mu}{\partial x^2}\right)_t \qquad 5.87a$$

L is called the *chemical* or *phenomenological coefficient*. These equations must be used in situations where there is a significant change in composition of the material through which diffusion is occurring, such a chemically zoned liquid or solid, or across a phase boundary. For example, consider an olivine crystal in equilibrium with a surrounding basaltic liquid. There would be a significant change in the concentration of a species such as Mg at the phase boundary, and hence equation 5.71 would predict that, even at equilibrium, diffusion of Mg out of the olivine and into the liquid should occur, but, of course, this will not be the case. However, by specifying that the olivine and the basalt are in equilibrium, we are specifying that the chemical potential of Mg is the same in the olivine and in the melt, and thus equation 5.87 correctly predicts that no diffusion will occur at equilibrium.

Example 5.7. Equilibration between a Mineral Grain and Pore Water

Imagine a grain of calcite in an accumulating sediment surrounded by pore water. Assume that the distribution coefficient of Sr between calcite and water is 100, that the calcite has an initial Sr concentration of 2000 ppm and that a constant Sr concentration of 10ppm is maintained in the water, and that the grain is spherical. If the diffusion coefficient for Sr in calcite is 10⁻¹⁵ cm²/sec and radius of the calcite grain is 1 mm, how will the average concentration of Sr in the grain change with time?

Answer: T h is problem is similar to the previous example (Example 5.6). This time, however, we want to know the average concentration of the

grain. The mass of Sr at time *t* in a sph erical shell of thickness *dr* is:

$$M(t) = c(r,t)4\pi r^2 dr$$

The average concentration of Sr in the grain at time t is then obtained by integrating and dividing by the volume:

$$\overline{C}(t) = \frac{1}{\frac{4}{3}\pi a^3} \int_0^a c(r,t) 4\pi r^2 dr \quad 5.88$$

where c(r,t) is given by equation 5.85 i n Example 5.6. The solution is (Albarède, 1995):

$$\overline{C}(t) = \frac{6C_i}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n} e^{-n^2 \pi^2 Dt/a^2} + C_0 \qquad 5.89 *$$

where C_i is the initial concentration in the calcite and C_0 is the concentration at the edge of the crystal, which will be in equilibrium with the pore water. The solution is shown Figure 5.16. The grain reaches equilibrium with the pore water within 100,000 to 200,000 years.

* The summation in equation 5.89 is slow if Dt/a^2 is small. An alternative solution to equ. 5.89 is:



Figure 5.16. Change in the bulk concentration of Sr in a 2 mm diam eter calcite grain assuming and initial concentration of 2000 ppm, a constant concentration in the porewater of 10 p pm and a calcite/water distribution coefficient of 100.

$$\overline{C}(t) = 1 - \frac{6\sqrt{Dt}}{a} \left[\sum_{n=1}^{\infty} \frac{e^{-na}/\sqrt{Dt}}{\sqrt{\pi}} - \frac{na}{\sqrt{Dt}} \left\{ 1 - erf\left(\frac{na}{\sqrt{Dt}}\right) \right\} \right] + 3\frac{Dt}{a^2} \qquad 5.90$$

The two solutions give identical results. They differ only in the ease of computation. For large values of Dt/a^2 the summation in 589 is preferred. For small values of Dt/a^2 5.90 is preferred.

CHAPTER 5: KINETICS

W. M. White

3.) Multicomponent Diffusion refers to situations where the concentration of the species of interest is sufficiently large that its diffusion must be coupled with diffusion of other species in the opposite direction to maintain electrical neutrality and/or constant volume. In such a circumstance, the diffusion of any one species is related to the diffusion of all other species. For example, had we considered the diffusion of Mg in olivine in Example 5.6, it would have been necessary to consider the diffusion of Fe in the opposite direction. Under such circumstances, the flux of species *i* is computed as:

$$J_i = -\sum_{k=1}^n D_{i,k} \frac{\partial c_k}{\partial x}$$
 5.91

where D_{ik} is the *interdiffusion coefficient* describing the interaction of species *i* and *k*, and *n* is the number of components in the system. Equation 5.91 is known as the Fick-Onsager Law. The interdiffusion coefficient is related to the tracer diffusion coefficient as:

$$D_{i,k} = \frac{n_i D_i + n_k D_k}{n_i + n_k}$$
5.92

where n_i is the mole fraction of *i* and D_i is the tracer diffusion coefficient for *i*.

The complete solution for diffusion flux in the system is:

$$\mathbf{J} = \mathbf{D}\mathbf{C}$$
 5.93

where \mathbf{J} is the flux vector, \mathbf{D} is the diffusion coefficient matrix, or tensor, and \mathbf{C} is the concentration gradient vector: . . .

$$\begin{bmatrix} J_1 \\ J_2 \\ \vdots \\ J_{n-1} \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} & \dots & D_{1n-1} \\ D_{21} & D_{22} & \dots & D_{2n-1} \\ \vdots & \vdots & & \vdots \\ D_{n-11} & D_{n-12} & \dots & D_{n-1n-1} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_{n-1} \end{bmatrix}$$

where C_i is $\partial c_i / \partial x$ (this is the same as equation 5.93, but in matrix notation). Interdiffusion can result in diffusion up a concentration gradient, because as we can see from equation 5.91, the flux of species idepends on the concentration gradients of all species, not just its own.

4.) Multicomponent Chemical Diffusion refers to situations where both the chemical potential and the diffusion of other species must be considered. In this case, the diffusion flux is calculated according to equation 5.93, but the diffusion coefficient matrix, D, must be calculated as:

$$\mathbf{D} = \mathbf{L}\mathbf{G}$$
 5.94

Where L is the matrix of phenomenological coefficients and the elements of G, the thermodynamic matrix, are functions of the derivatives of activity with respect to concentration. The paper of Liang et al. (1997) provides an excellent review of the theory, and the experiments of Watson (1982) provides a good example of the complexity and non-intuitive nature of multicomponent chemical diffusion. In this experiment, a quartz sphere was dissolved in molten basalt. All elements except SiO₂ were diffusing out of the basalt into SiO₂-rich liquid.

Example 5.8. Interdiffusion Coefficients

Calculate the diffusion flux for Mn in a garnet given the interdiffusion coefficient matrix below if the concentration gradients are .023 mol/cm, -0.009 mol/cm and -.015 mol/cm for Mn²⁺, Mg²⁺, and Fe²⁺ respectively.

Diffusion	Coefficient	Matrix	for	GARNET

	Mn	Mg	Fe
Mn	8.38×10 ⁻²⁰	-9.91×10 ⁻²⁰	-4.68×10 ⁻²¹
Mg	-2.78×10 ⁻²⁰	7.26×10 ⁻²¹	-8.81×10 ⁻²³
Fe	-7.16×10 ⁻²⁰	-4.81×10 ⁻²³	1.19×10 ⁻²⁰

from Demspter (1975).

Answer: we calculate the diffusion flux for Mn using equation 5.91. We find that J_{Mn} is 3×10^{-21} $mole/cm^2-sec.$

W. M. White Chapter 5: Kinetics

Figure 5.17 shows concentration profiles in which Na and K are actually diffusing in the direction of higher concentration.

Although diffusion is treated differently in different circumstances, it is important to bear in mind that the mechanism on a microscopic scale is always the same: it results from the random motion of atoms or molecules. Reference is sometimes made to the 'driving force' of diffusion, taken either as the concentration or chemical potential gradient. While it may sometimes be convenient to think in terms of 'driving forces', these forces are not real.

To demonstrate this point, we can derive Fick's first law just from a consideration of random atomic motion. Consider two adjacent lattice planes in a crystal spaced a distance dxapart. Let the number of atoms of the element of interest at the first plane be n_1 and the number of atoms at the second be n_2 . We assume that atoms can change position randomly by jumping to an adjacent plane and that this occurs with an average frequency v (i.e., 1 jump of distance dx every 1/v sec). We further assume that there are no external forces, so that a jump in any direction has equal probability. At the first plane there will be $vn_1/6$ atoms that jump to the second plane (we divide by 6 because there are 6 possible jump directions: up, down, back, front, right, left). At the second plane there will be $vn_2/6$ atoms that jump to first plane. The net flux from the first plane to the second is then:

$$J = \frac{v n_1 / 6 - v n_2 / 6}{dx^2} = \frac{v}{6} \frac{(n_1 - n_2)}{dx^2} \qquad 5.95$$

The concentration, *c*, is the number of atoms of interest per unit volume, i.e., n/dx^3 , so we may substitute cdx^3 for the number of atoms in 5.95:

$$J = \frac{v}{6} \frac{(c_1 - c_2)dx^3}{dx^2} = \frac{v}{6}(c_1 - c_2)dx$$

The concentration gradient is just $(c_2-c_1)/dx$, so letting $dc = -(c_1 - c_2)$, we have:

$$J = -\frac{v dx^2}{6} \frac{dc}{dx}$$

If we let $D = vdx^2/6$ then we have Fick's first law:

$$J = -D\frac{ac}{dx}$$

Hence *D* is related to the jump frequency, v, and square of the jump distance (*dx*).

We see that there is a net diffusion, not because of the presence of a force, but only because there are more atoms at one point than at an adjacent one. In the absence of a concentration gradient (i.e., n_1 and



Figure 5.17. Electron microprobe traverses across boundary layers of melt in contact with dissolving quartz spheres. Quartz is to the left at zero microns. Note that K_2O and Na_2O are diffusing 'uphill', that is, toward higher K_2O and Na_2O concentrations in the SiO₂-rich liquid. From Watson (1982).

W. M. White Chapter 5: Kinetics

 n_2 the same, there would be $n\nu/6$ atoms moving from the first to the second plane and $n\nu/6$ atoms moving from second to the first plane. But if we cannot distinguish atoms originally at the first plane from those originally at the second, these fluxes balance and go unnoticed. If we could distinguish the atoms, we could detect a flux even in the absence of a concentration gradient. Other factors, such as pressure or stress gradient, presence of an electromagnetic field, a temperature gradient, or concentration gradients of other species may make a jump in one direction more probable than another, as can differences in chemical potential between the two planes. Such 'forces' will affect the diffusion flux and terms must be added to the diffusion equations to account for them. The point we are stressing here is that diffusion can occur in the absence of all such forces.

5.4.3 Diffusion in Solids and the Temperature Dependence of the Diffusion Coefficient

We can imagine four ways in which diffusion might take place in solids (Figure 5.18):

1.) *Exchange*: the interchange of position of two atoms in adjacent sites.

2.) *Interstitial*: in which an atom moves from one interstitial eite to enother

site to another.

3.) *Interstitialcy*: in which an atom is displaced from a lattice site into an interstitial site.

4.) *Vacancy*: in which an atom moves from a lattice site to a vacancy, creating a vacancy behind it. Mechanisms 1 and 3 involve displacement of two atoms and therefore have high activation energies. Since interstitial sites are likely to be small, mechanism 2 will apply mainly to small atoms (H and He, for example). Thus we are left with mechanism four as a principal mechanism of diffusion in solids.

Hence, diffusion in solids is a bit like a game of checkers: an atom can generally only travel by moving from lattice site to lattice site. Furthermore, it can only move to a vacant lattice site (and one of the correct type). In general, lattice vacancies are of two types: permanent and temporary. Permanent vacancies can arise from defects or through the presence of impurities, for example through substitution of a doubly charged ion for a singly charged one with a vacancy providing charge balance. Temporary sites arise from thermal agitation causing the volume of the solid to be slightly greater than the ideal volume by forcing atoms onto the surface. The number of the former is temperature independent, the latter are temperature dependent.

Let's attempt to calculate a diffusion coefficient *ab initio* for the simple one-dimensional case of tracer diffusion in a solid occurring through the vacancy mechanism. Since a certain minimum energy is required to get an ion out of the lattice site 'energy well' we would expect the number of the temporary vacancies to have a temperature dependence of the form of equation 5.16, the Boltzmann distribution law. Thus the number of lattice vacancies can be written as:

$$N_{vac} = N_{perm} + ke^{-E_H/RT}$$

where k is some constant and E_H is an activation energy needed to create a vacancy or 'hole'.

The probability, \mathcal{P} , of an atom making a successful jump to a vacant site is found by multiplying the number of attempts, \aleph , by the fractions of atoms having sufficient energy to get out of the well:

$$\mathcal{P} = \aleph e^{-E_B / RT}$$
 5.96

The number of attempts is simply the vibration rate, v, times the number of holes:

$$\aleph = \nu[N_{vac}] = \nu \Big[N_{perm} + k e^{-E_H / RT} \Big]$$
5.97

Combining 5.96 and 5.97 we have:

$$\mathcal{P} = vN_{perm}e^{-E_B/RT} + vke^{-(E_H + E_B)/RT}$$



mechanisms in solids. After Broecker

and Oversby (1971).

W. M. White

Geochemistry

CHAPTER 5: KINETICS

The diffusion rate should be the number of jumps times the distance per jump, *d*:

$$\Re = v dN_{perm} e^{-E_B/RT} + v dk e^{-(E_H + E_B)/RT}$$
5.98

or

$$\Re = m e^{-E_B / RT} + n e^{-(E_H + E_B) / RT}$$
 5.99

where *n* and *m* are simply two constants replacing the corresponding terms in equation 5.98. Thus diffusion rates generally will have a temperature dependence similar to equation 5.24. At low temperature, the permanent vacancies will dominate and diffusion rate equation will look like:

$$\Re \simeq m e^{-E_B/RT}$$
 5.100

At higher temperature where thermally generated vacancies come into play, the latter term in 5.99 dominates and the diffusion rate equation will look like:

$$\Re \cong ne^{-(E_H + E_B)/RT}$$
 5.101

Diffusion that depends on thermally created vacancies is sometimes called *intrinsic* diffusion, while that depending on permanent vacancies is called

that depending on permanent vacancies is called *extrinsic* diffusion. The boundary between these regions will vary, depending on the nature of the material and the impurities present. For NaCl, the transition occurs around 500° C, for silicates it generally occurs above 1000° C. Where the diffusion mechanism changes, a break in slope can be observed on a plot of $\ln D$ vs. 1/T. For example, Figure 5.19 shows how the diffusion coefficient might change based on equations 5.100 and 5.101.

Combining E_B and E_H into a single activation energy term, E_A , which is the energy necessary to create the vacancy and move another atom into it, a typical expression for temperature dependence of the diffusion coefficient in solids is:

$$D = D_o e^{-E_a/RT}$$
 5.102

where D_o is again called the frequency factor. As we have seen, it will depend on vibrational frequency and the distance of the interatomic jump.

Experimental observation supports our theoretical expectation of an exponential temperature dependence of diffusion, for example, in a series of measurements of the diffusion coefficient, D, at various temperatures (Figure 5.19). Taking the log of both sides of equation 5.102, we obtain:

$$\ln D = \ln D_o - \frac{E_A}{RT}$$

Thus on a plot of ln *D* versus reciprocal temperature, data for diffusion of a given element in a given substance should plot alone a line with slope E_A/R and intercept D_0 (Figure 5.20). Values for the activation energy are generally similar for most elements (typically 50 to 200 kJ), but the frequency factor varies widely. Table 5.2 list frequency factors and activation energies for several elements in various geological materials.

The pressure dependence of the diffusion coef-



Fig. 5.19. Arrhenius plot illustrating how the change in diffusion mechanism from intrinsic to extrinsic can result in a break in slope.



Figure 5.20. Schematic plot of log diffusion rate against inverse of temperature (Arrhenius plot) for two elements: Cs and Sr.

W. M. White Chapter 5: Kinetics

ficient is:

$$D = D_{o}e^{-(E_{A} + (P - P_{ref})\Delta V)/RT} 5.103$$

where ΔV is the "activation volume".

5.4.4 Diffusion in Liquids

In both liquids and solids, diffusion coefficients depend on both the nature of the diffusing species and the nature of the media. As you might expect, small atoms generally diffuse more rapidly than large ions. The value of the diffusion coefficient in liquids can be estimated in a number of ways. Based on a model of molecular motion in a non-ionic liquid composed of molecules of diameter *d*, and assuming a kinetic energy per atom of 3kT/2 and a mean free path length of $2\alpha dT$, where α is the coefficient of thermal expansion, *m* is molecular mass and k is Boltzmann's constant, the diffusion coefficient may be estimated as (Kirkaldy and Young, 1985):

$$D \cong \alpha d \sqrt{\frac{8k}{\pi m}} T^{3/2} \qquad 5.104$$

This predicts a diffusion coefficient near the melting point of 10^{-4} cm/sec.

Diffusion coefficients in liquids are also commonly expressed in terms of viscosity. For uncharged species, the dependence of the diffusion coefficient on molecular radius and viscosity is expressed by the *Stokes-Einstein Equation*:

$$D = \frac{kT}{6\pi nr} \qquad 5.105$$

where r is molecular radius and η is viscosity. From this equation, we see that diffusion becomes more difficult as the liquid becomes more structured (polymerized) because the viscosity increases with increasing polymerization.

Because of the electrical neutrality effect, ion charge is important in diffusion of ions. In aqueous electrolytes, Table 5.2: Frequency factors & ActivationEnergies

Phase	Temp.	Do	EΔ
	Range ° C	$\rm cm^2/sec$	kJ/mol
Olivine	1125-1200	5.59×10^{-02}	264
Olivine	1000-1125	5.90×10^{-07}	133
Olivine	1000-1150	1.82×10^{-08}	143
garnet	750-1475	5.15×10^{-04}	145
garnet	750-1475	6.39×10^{-04}	157
β-Qz	1010-1220	3.70×10^{-13}	230
Anorthite	800-1300	9.00×10^{-10}	234
Spinel	1200-1600	2.40×10^{-06}	337
Basaltic melt	1250-1450	0.278	104
Basaltic melt	1250-1450	0.535	105
Basaltic melt	1250-1450	0.059	94
Basaltic melt	1300-1400	110	272
Basaltic melt	1250-1450	0.053	86
Andesitic melt	1300-1400	0.44	209
Andesitic melt	1300-1400	3.5	213
Andesitic melt	1300-1400	160	280
Plagioclase	700-1050	1.86×10^{-03}	266
K-feldspar	700-1050	1.82×10^{-01}	309
Plagioclase	700-1050	1.78×10^{-03}	265
albite	600-800	1.25×10^{-01}	100
orthoclase	800-870	6.00×10^{-06}	98
obsidian	500-980	1.50×10^{-08}	27
	Phase Olivine Olivine garnet garnet garnet β-Qz Anorthite Spinel Basaltic melt Basaltic melt Basaltic melt Basaltic melt Basaltic melt Andesitic melt Andesitic melt Andesitic melt Plagioclase K-feldspar Plagioclase albite orthoclase obsidian	Phase Temp. Range ° C Olivine 1125-1200 Olivine 1000-1125 Olivine 1000-1150 garnet 750-1475 garnet 750-1475 garnet 750-1475 garnet 750-1475 garnet 800-1300 Spinel 1200-1600 Basaltic melt 1250-1450 Basaltic melt 1250-1450 Basaltic melt 1250-1450 Basaltic melt 1300-1400 Basaltic melt 1300-1400 Basaltic melt 1300-1400 Andesitic melt 1300-1400 Plagioclase 700-1050 K-feldspar 700-1050 albite 600-800 orthoclase 800-870 obsidian 500-980	PhaseTemp. D_0 Range °C D_0 cm²/secOlivine1125-1200 5.59×10^{-02} Olivine1000-1125 5.90×10^{-07} Olivine1000-1150 1.82×10^{-08} garnet750-1475 5.15×10^{-04} garnet750-1475 6.39×10^{-04} garnet750-1475 6.39×10^{-04} β -Qz1010-1220 3.70×10^{-13} Anorthite800-1300 9.00×10^{-10} Spinel1200-1600 2.40×10^{-06} Basaltic melt1250-1450 0.278 Basaltic melt1250-1450 0.535 Basaltic melt1250-1450 0.059 Basaltic melt1300-1400110Basaltic melt1300-1400 0.44 Andesitic melt1300-1400 3.5 Andesitic melt1300-1400 1.86×10^{-03} Andesitic melt1300-1400 1.86×10^{-03} Flagioclase700-1050 1.86×10^{-03} K-feldspar $700-1050$ 1.78×10^{-03} albite $600-800$ 1.25×10^{-01} orthoclase $800-870$ 6.00×10^{-06}

Table 5.3 Trace Diffusion Coefficients for Ions in Infinitely Dilution Solution at 25°C.

Cation	D°	Anion	D°
	10 ⁻⁶ cm ⁻² sec ⁻		10 ⁻⁶ cm ⁻² sec ⁻
	1		1
H^{+}	93.1	OH⁻	52.7
Na^+	13.3	Cl-	20.3
K^+	19.6	I-	20.0
Mg^{2+}	7.05	SO_4^{2-}	10.7
Ca ²⁺	7.93	CO_3^{2-}	9.55
Sr ²⁺	7.94	HCO_3^{2-}	11.8
Ba ²⁺	8.48	NO ₃	19.0
Fe^{2+}	7.19	0	
La ³⁺	6.17		

From Lasaga (1997).

tracer diffusion coefficients depend on ion charge as:

Geochemistry

Chapter 5: Kinetics

W. M. White

$$D^{\circ} = \frac{RT\lambda^{\circ}}{|z|\mathcal{F}^2}$$
 5.106

where λ° is the limiting ionic conductance (conductance extrapolated to infinite dilution), \mathcal{F} is Faraday's constant, and *z* is the charge of the ion. The naught (°) denotes the standard state of infinite dilution. Table 5.3 lists values of D° for a few ions of geochemical interest. In dilution solution, diffusion coefficients depend approximately on the square root of ionic strength. In more concentrated solutions, diffusion coefficients show a complex dependence on ionic strength, the treatment of which is beyond the scope of this book. Discussions of this problem may be found in Anderson (1981), Tyrell and Harris (1984), and Lasaga (1997).

5.5 Surfaces, Interfaces, and Interface Processes

The properties of a phase at its surface are different from the bulk properties of the phase. This difference arises from the difference between the local environment of atoms on a surface or interface and those in the interior of a phase. An atom at the surface of a crystal is not surrounded by the same bonds and distribution of charges as it would be in the interior of the crystal lattice. Its potential energy must therefore be different. Here we define *surface* as the exterior boundary of a condensed phase (a solid or liquid) in a vacuum or gas. An *interface* is the boundary between two condensed phases, for example, between 2 crystals or between a mineral and water (the term 'surface' is, however, often used for what we have just defined as an interface). Surfaces, surface energies, and interfaces play an important role in many geochemical processes. All heterogeneous reactions, i.e., those involving more than 1 phase, must always involve interfaces or surfaces. Dissolution, melting, exsolution, and precipitation are examples of processes that, on an atomic scale, occur entirely at or near the interface between two phases. Surfaces can also play important roles as catalysts in many geochemical reactions.

On a microscopic scale, the reactivity of mineral surfaces will vary locally for several reasons. The first is the microtopography of the surface (Figure 5.21). For example, a single growth unit (which might be a single atom, an ion, or molecule and called an *adatom*), located on an otherwise flat surface will be particularly unstable because it is bonded to other units on only one side. A *step* (which might be formed through growth, dissolution, or *screw dislocation*) provides a more favorable growth site because the new unit is bound to other units on 2 sides. An even better site for growth is at a *kink*, where bonds may be formed on 3 sides. Conversely, a unit at a kink (with 3 exposed sides) is less stable than one at a step (with 2 exposed sides), which in turn is less stable than a unit on a flat surface, with only 1 exposed side. The point is, kinks and steps will be more reactive than other features, so surface reactions rates will depend in part on the density of these features.

Properties of mineral surfaces will also vary depending on the orientation of the surface relative to crystallographic axes. Most minerals grow or dissolve faster in one direction than in another. Most surface reactions involve the formation of new bonds between atoms of a mineral and atoms of the adjacent phase; the nature of the bonds that are possible will depend on the orientation of the surface relative to crystallographic axes. Reaction rates measured for one crystal face may not be representative of other faces.

Finally, almost all minerals have a variety of atoms and crystallographic sites, hence there will be a variety of bonds that are possible on any surface. We will discuss this aspect of surfaces in slightly more detail below.

5.5.1 The Surface Free Energy

In Chapters 2 and 3, we introduced the concept of molar quantities and partial molar quantities. For example, the molar volume of a substance was:

$$\overline{V} = \frac{V}{n}$$

and we defined the partial molar volume as:

CHAPTER 5: KINETICS

W. M. White

$$v_i^{\phi} = \left(\frac{\partial V}{\partial n_i}\right)_{T,P,n_j} \tag{3.11}$$

We now define 2 new quantities, the *molar surface area*:

$$\overline{A} = \frac{A}{n}$$
 5.107

and the *partial molar surface area*:

$$a_i^{\phi} = \left(\frac{\partial A}{\partial n_i}\right)_{T,P,n_j}$$
5.108

where *A* is the surface area of the phase. The molar volume or the molar Gibbs Free Energy of pure quartz depends only on temperature and pressure. Thus the molar volume of each (pure) quartz crystal is the same as that of every other (pure) quartz crystal at that temperature and pressure. Unlike other molar quantities, the *molar surface area and partial molar surface area depend on shape and size*, and are therefore not intrinsic properties of the substance. For a sphere, for example, the partial molar surface area is related to molar volume as:

$$a = \frac{\partial A}{\partial V} \frac{\partial V}{\partial n} = \frac{2v}{r}$$
 5.109

For other shapes, the relationship between *a* and *v* will be different.

Finally, we define the *surface free energy* of phase ϕ as:

$$\sigma^{\phi} \equiv \left(\frac{\partial G}{\partial A}\right)_{T,P,n}^{\varphi}$$
5.110

The surface free energy represents those energetic effects that arise because of the difference in atomic environment on the surface of a phase. Surface free energy is closely related to surface tension. The total surface free energy of a phase is minimized by minimizing the phase's surface area. Thus a water drop in the absence of other forces will tend to form a sphere, the shape that minimizes surface area. When surface effects must be considered we can revise the Gibbs Free Energy equation (equ. 3.14) to be:

$$dG = VdP - SdT + \sum_{i} \mu_{i} dn_{i} + \sum_{k} \sigma_{k} dA_{k}$$
5.111

where the last sum is taken over all the interfaces of a system. In this sense different crystallographic faces have different surface free energies. The last term in 5.111 increases in importance as size decreases. This is because the surface area for a given volume or mass of a phase will be greatest when particle size is small.

5.5.2 The Kelvin Effect

When the size of phases involved is sufficiently small, surface free energy can have the effect of displacing equilibrium. For an equilibrium system at constant temperature and pressure, equation 5.111 becomes:

$$0 = \sum_{i} \mathbf{v}_{i} \mu_{i}^{o} + RT \sum_{i} \mathbf{v}_{i} \ln a_{i} + \sum_{k} \sigma_{k} dA_{k}$$

The first term on the right is ΔG° , which according to 3.86 is equal to $-RT \ln K$. This is the 'normal' equilibrium constant, uninfluenced by surface free energy, so we'll call it K°. We will call the summation in the second term



Screw Dislocation

Figure 5.21. On a microscopic scale, the surface of a mineral exhibits a number of features. As a result, the local reactivity of the surface will be quite variable.

W. M. White

Geochemistry

CHAPTER 5: KINETICS

K^s, the equilibrium influenced by surface free energy. Making these substitutions and rearranging, we have:

$$\ln K^{s} = \ln K^{\circ} - \frac{\sum \sigma_{k} dA_{k}}{RT}$$
 5.112

Thus we predict that equilibrium can be shifted due to surface free energy, and the shift will depend on the surface or interfacial area. This is known as the *Kelvin Effect*.

There are a number of examples of this effect. For example, fine, and therefore high surface area, particles are more soluble than coarser particles of the same composition. Water has a surface free energy of about 70 mJ/m². Consequently, humidity in clouds and fogs can reach 110% when droplet size is small.

5.5.3 Nucleation and Crystal Growth

5.5.3.1 Nucleation

Liquids often become significantly oversaturated with respect to some species before crystallization begins. This applies to silicate liquids as well as aqueous solutions (surface seawater is several times oversaturated with respect to calcite). However, crystallization of such supersaturated solutions will often begin as soon as seed crystals are added. This suggests that nucleation is an important barrier to crystallization. This barrier arises because the formation of a crystal requires a local increase in free energy due to the surface free energy at the solid-liquid interface.

Let's explore how nucleation can be inhibited a bit further. For a crystal growing in a liquid, we can express the complete free energy change as:

$$dG_{tot} = \sigma dA + dG_{xt}$$
 5.113

where dG_{xt} refers to the free energy change of the crystallization reaction that apply throughout the volume of the crystal (i.e., free energy in the usual sense, neglecting surface effects).

Let's consider a more specific example, that of a spherical crystal of phase ϕ growing from a liquid solution of component ϕ . The free energy change over some finite growth interval is:

$$\Delta G_{tot} = 4\pi r^2 \sigma + \frac{4}{3}\pi r^3 \frac{\Delta G_{xt}}{\overline{V}} = 5.114$$

where *r* is the radius (we divide by \overline{V} to convert joules per mole to joules per unit volume). The first term is on the right is the surface free energy, and, although small, *is always positive*. At the point where the solution is exactly saturated ΔG will be 0. The net free energy, ΔG_{totr} is thus positive, so the crystal will tend to dissolve. In order for spontaneous nucleation to occur, the second term on the right must be negative and its absolute value must exceed that of the first term on the right of 5.114, i.e., the liquid must be supersaturated for nucleation to occur. Solving 5.114 for *r*, we find that ΔG_{tot} is ≤ 0 when $r \geq -3\sigma/\Delta G$.

How will ΔG vary with *r* up to this point? To answer this, we differentiate 5.114 with respect to *r*:

$$\frac{\partial \Delta G_{tot}}{\partial r} = 8\pi r\sigma + 4\pi r^2 \frac{\Delta G_{xt}}{\overline{V}} \qquad 5.115$$



Figure 5.22. Free energy as a function of crystal radius for small crystals forming near the saturation point. ΔT is the amount of undercooling (difference between temperature and saturation temperature).

CHAPTER 5: KINETICS

W. M. White

Since the volume free energy term is proportional to r^2 and the surface free energy term to r, the latter necessarily dominates at very small values of r. For small values of r, ΔG_{tot} will increase with increasing r because σ is always positive. In other words, near the saturation point where ΔG is small, very small crystals will become increasingly unstable as they grow. The critical value of r, that is the value at which ΔG will decrease upon further growth, occurs where $\partial G/\partial r = 0$. Solving equation 5.115, we find that

$$r_{crit} = -\frac{2\sigma}{\Delta G_{rt}/\overline{V}}$$
 5.116

For a solution that undergoes cooling and becomes increasingly saturated as a result (e.g., a magma or a cooling hydrothermal solution), we can use equation 5.68 to approximate the ΔG term (i.e., $\Delta G \cong -\Delta T \Delta S$, where ΔT is the difference between actually temperature and the temperature at which saturation occurs and ΔS is the entropy change of crystallization). Figure 5.22 shows the total free energy calculated in this way as a function of *r* for various amounts of undercooling.

The surface free energy term correlates with viscosity. Thus nucleation should require less supersaturation for aqueous solutions than silicate melts. Among silicate melts, nucleation should occur more readily in basaltic ones, which have low viscosities, than in rhyolitic ones, which have high viscosities. This is what one observes. Also, we might expect rapid cooling to lead to greater supersaturation than slow cooling. This is because there is an element of chance involved in formation of a crystal nucleus (the chance of bringing enough of the necessary components together in the liquid so that *r* exceeds r_{crit}). Slow cooling provides time for this statistically unlikely event to occur, and prevents high degrees of supersaturation from arising. With rapid cooling, crystallization is postponed until ΔG_r is large, when many nuclei will be produced. Let's briefly consider nucleation rates in more detail.

5.5.3.2 Nucleation Rate

The first step in crystallization from a liquid is the formation of small clusters of atoms having the composition of the crystallizing phase. These so-called *heterophase fluctuations* arise purely because of statistical fluctuations in the distribution of atoms and molecules in the liquid. These fluctuations cause local variations in the free energy of the liquid, and therefore their distribution can be described by the Boltzmann distribution law:

$$N_{i} = N_{i} e^{-\Delta G_{xt} / kT}$$

where N_i is the number of clusters per unit volume containing *i* atoms, N_v is the number of atoms per unit volume of the cluster, and ΔG_i is the difference between the free energy of the cluster and that of the liquid as a whole. The number of clusters having the critical size (r_{crit}) is:

$$N_{crit} = N_v e^{-\Delta G_{crit}/k}$$

where ΔG_{crit} is the total free energy (ΔG_{tot}) of clusters with critical radius obtained by solving equation 5.113 when $r = r_{crit}$. For spherical clusters, this is:

$$\Delta G_{crit} = \frac{16\pi}{3} \frac{\sigma^3 V}{\Delta G_{rt}^2}$$
 5.117

Substituting equation 5.68 for ΔG_{r} , we have:

$$\Delta G_{crit} = \frac{16\pi}{3} \frac{\sigma^3 \overline{V}}{\left(\Delta T \Delta S_{rt}\right)^2}$$
 5.118

If E_A is the activation energy associated with attachment of an additional atom to a cluster, then the probability of an atom having this energy is again given by the Boltzmann distribution law:

$$\mathcal{P} = e^{-E_A/k}$$

Now according to transition state theory, the frequency of attempts, v, to overcome this energy is simply the fundamental frequency v = kT/h. The attachment frequency is then the number of atoms adjacent to the cluster, N^* , times the number of attempts times the probability of success:

CHAPTER 5: KINETICS

W. M. White

$$N^* v \mathcal{P} = N^* \frac{\mathbf{k}T}{h} e^{-E_A/\mathbf{k}T}$$
 5.119

The nucleation rate, *I*, is then the attachment frequency times the number of clusters of critical radius:

$$I = N_{crit} N^* v \mathcal{P} = N^* \frac{kT}{h} e^{-E_A/kT} e^{-\Delta G_{crit}/kT}$$
5.120

Combining the frequency of attachment terms into a pre-exponential frequency factor A, and substituting 5.118 into 5.120 we have:

$$I = A e^{-16\pi\sigma^{3} \overline{V}^{2} / (3\Delta G^{2} kT)}$$
5.121

which is the usual expression for nucleation rate (e.g., McLean, 1965). If we substitute 5.68 into 5.121, we see that:

$$I = A e^{-16\pi\sigma^3 \overline{V}^2 / 3(\Delta S \Delta T)^2 kT} \quad \text{or} \qquad I \propto e^{-1/\Delta T^2} \qquad 5.122$$

Example 5.9. Nucleation of Diopside

The enthalpy of fusion of diopside is 138 kJ/mol and its melting temperature is 1665 K. Assuming an activation energy of 1×10^{-18} J, how will the nucleation rate of diopside crystals in a diopside melt vary with temperature for surface free energies of 0.02, 0.06, and 0.12 J/m²?

Answer: The one additional piece of information we need here is the molar volume, which we find to be 66 cc/ mol from Table 2.2. We can calculate ΔS_m from the relation:

$$\Delta S_m = \frac{\Delta H_m}{T_m}$$

Assuming $\Delta S_{m\nu}$, σ , and E_A are independent of temperature, we can use equations 5.120 and 5.118 to calculate the nucleation rate. The calculation for the 3 surface free energies is shown in Figure 5.23a. Nucleation will be experimentally observable when the nucleation rate reaches $\approx 10^{-10}$ m⁻², which corresponds roughly to 1 nuclei/cm²/hr. For a surface free energy of 0.02 J/m², the rate is reached only a few kelvins below the melting point. Further undercooling results in very high nucleation rates. For a surface energy of 0.06 J/m², an undercooling of 35 K is required, and an undercooling of 130 K is required at the highest value of surface energy. In the latter case, the rise in nucleation rate with undercooling is not nearly as steep.



Figure 5.23. Calculated nucleation rate of diopside in diopside melt as a function of temperature.

In Figure 5.23b, we see that the nucleation rate passes through a maximum and as undercooling proceeds further, the rate decreases. This decrease reflects the 1/T dependence of both exponential terms in equation 5.120, i.e., the formation and growth of h eterophase fluctuations will fall as temperatures falls. Observed nucleation rates show this maximum, but the "bell" is generally more symmetric and considerably narrower. This reflects the increasing viscosity of the melt, and therefore the decreasing mobility of atoms, i.e., diffusion of atoms to the proto-nuclei slows.

W. M. White Chapter 5: Kinetics

This implies nucleation rate will be a very strong function of "temperature overstepping" for relatively small values of ΔT , but will level off at higher values of ΔT . At low degrees of overstepping, nucleation rate will be nil, but will increase rapidly once a critical temperature is achieved, as is demonstrated in Example 5.9. A more detailed treatment of nucleation and growth of crystals in cooling magmas can be found in Toramaru (1991).

5.5.3.3 Heterogeneous Nucleation

The nucleation of diopside crystals from diopside melt is an example of homogenous nucleation, i.e., nucleation in a system where initially only one phase is present. Heterogeneous nucleation refers to the nucleation of a phase on a pre-existing one. Often the surface free energy between the nucleating



Figure 5.24. Illustration of the balance of forces as a spherical crystal or droplet of phase β crystallizes or condenses from phase α on a surface.

phase and the pre-existing surface is lower than between the nucleating phase and the phase from which it is growing. Hence heterogeneous nucleation is often favored over homogeneous nucleation. Perhaps the most familiar example is dew. Dew droplets appears on surfaces, such as those of grass, at significant lower relative humidity than necessary for fog or mist to form. The reason is that the surface free energy between grass and water is lower than between water and air. Another example is the clusters of crystals seen in igneous rocks. These result from one crystal nucleating on the other, again because the free energy of the crystal-crystal interface is lower than that of the crystal-magma interface.

Let's examine this in a more quantitative fashion. Consider a spherical cap of phase β nucleating from phase α on a flat surface (s) (Figure 5.24). The balance of surface forces at the three-phase contact is:

$$\sigma_{\alpha s} = \sigma_{\beta s} + \sigma_{\alpha \beta} \cos \theta \qquad 5.123$$
$$\cos \theta = \frac{\sigma_{\alpha s} - \sigma_{\beta s}}{\sigma_{\alpha \beta}}$$

and solving for θ

If the interfacial energy between the nucleating phase, β , and the surface (σ_{β_s}) is smaller than that between phase α and the surface (σ_{β_s}) , then the angle of intersection, θ , will be small so as to maximize the interfacial surface area between β and s for a given volume of β . In the limit where $\sigma_{\beta_s} << \sigma_{\alpha_s}$ then θ will approach 0 and β will form a film coating the surface. As σ_{β_s} approaches σ_{α_s} the nucleating phase will form more spherical droplets. If $\sigma_{\beta_s} \ge \sigma_{\alpha_s}$ then θ will be 90° or greater, and heterogeneous nucleation will not occur. To take account of the reduced interfacial energy between β and s, equation 5.117 becomes:

$$\Delta G_{crit} = \frac{16\pi}{3} \frac{\sigma_{\alpha\beta}^3 \overline{V}}{\Delta G_{xt}^2} \left(2 - 3\cos\theta + \cos^3\theta\right)$$
 5.124

In metamorphic reactions, nucleation will necessarily always be heterogeneous. Provided the necessary components of the nucleating phase are available and delivered rapidly enough by fluid transport and diffusion, interfacial energy will dictate where new phases will nucleate, nucleation being favored on phases where the interfacial energy is lowest. Where transport of components limit growth, however, this may not be the case, as phases will nucleate where the components necessary for growth are available. For example, experimental investigation of the reaction calcite + quartz \rightleftharpoons wollastonite + CO₂ revealed that in the absence of water, wollastonite nucleated on quartz. In experiments where water was present, it nucleated on calcite. SiO₂ is not significantly soluble in CO₂, so it could not be transported in the H₂O free experiments, hence nucleation could only occur where SiO₂ was available, i.e., at the surface of quartz, despite a probable higher interfacial energy.

Unfortunately, agreement between observed and predicted nucleation rates is often poor (Kirkpatrick, 1981; Kerrick et al., 1991). Equation 5.121 and Figure 5.23 show that the nucleation rate is a very

W. M. White Chapter 5: Kinetics

strong function of the surface free energy (I $\propto \exp(\sigma^3)$), and the poor agreement between theory and observation may reflect the lack of accurate data on surface free energy as well as the activation energy, E_A . However, it may also indicate that further work on the nucleation theory is required.

5.5.3.4 Diffusion-Limited and Heat-Flow Limited Growth

Two other kinetic factors affect crystallization. These are the local availability of energy and local availability of components necessary for crystal growth. The latter can be important where the crystal is of different composition than the liquid (almost always the case in nature except freezing of fresh water). Crystals can grow only as fast as the necessary chemical components are delivered to their surfaces. Where diffusion is not rapid enough to supply these components, diffusion will limit growth.

A second effect of slow diffusion is to change the *apparent* distribution coefficient, because the crystal 'sees' the concentrations in the



Figure 5.25. Variation of Sr and Rb concentrations from a plagioclase liquid interface. Solid curves show the variation of concentration. The crystal-liquid interface is at 0. Dashed lines show the concentrations at infinite distance from the interface (Sr_{∞} , Rb_{\infty}). Sr_{plag} and Rb_{plag} are the concentrations in the crystal.

adjacent boundary layer rather than the average concentrations in the liquid. Thus the crystal may become less depleted in elements excluded from the crystal, and less enriched in elements preferentially incorporated in it, than equilibrium thermodynamics would predict. For example, suppose a crystal of plagioclase is crystallizing from a silicate melt. Plagioclase preferentially incorporates Sr and excludes Rb. If diffusion of Sr and Rb to the crystal is slow compared with the crystal growth rate, the liquid in the boundary layer immediately adjacent the crystal will become impoverished in Sr and enriched in Rb. The crystal will grow in equilibrium with this boundary layer liquid, not the average magma composition, thus will be poorer in Sr and richer in Rb than if it grew in equilibrium with the average magma. Figure 5.25 illustrates this point. If however, growth rate of the crystal is very much slower than the transport of components to the crystal-liquid interface, this circumstance will not arise.

When crystals grow from a magma there will be a local increase in temperature at the crystal-liquid boundary, due to release of latent heat of fusion, ΔH_m , which will retard crystal grow. In most cases, however, advection and conduction of heat is probably sufficiently rapid that this is at best a minor effect. The effect is probably more important in prograde metamorphic reactions (e.g., dehydration reactions), which are usually endothermic and hence require a continuous supply of energy to maintain crystal growth. Where crystal growth and transport of components is sufficiently rapid, heat flow may limit rates of crystal growth. This is more likely to occur at high temperatures and in late stages of metamorphism when structures are already large (Fisher, 1978). We will explore the effects of crystal growth rates, diffusion rates, and heat flow on metamorphic processes in a subsequent chapter.

5.5.4 Adsorption

Many geochemically important reactions take place at the interface between solid and fluid phases, and inevitably involve adsorption and desorption of species onto or from the surface of the solid. Two types of adsorption are possible: physical and chemical. Physical adsorption involves the attachment of an ion or molecule to a surface through intermolecular or van der Waals forces. Such forces are relatively weak and heat of adsorption (ΔH_{ad}) relatively low (typically 4-12 kJ/mol). Chemical adsorption involves the formation of a new chemical bond between the adsorbed species and atoms on the surface of the solid. Heats of chemical adsorption are relatively large (>40 kJ/mol).

CHAPTER 5: KINETICS

W. M. White

Adsorption of ions and molecules on a solid surface or interface affects the surface free energy. The relationship between surface free energy and adsorbed ions can be expressed as:

$$d\sigma = -\sum_{i} \frac{n_{i,s}}{A} d\mu_{i} = -\sum_{i} \Gamma_{i} d\mu_{i}$$
 5.125

where $n_{i,s}$ is the number of mole of species *i* adsorbed at the surface, A is the surface area, and we define Γ_i as the *Gibbs adsorption density*. Because silicates and oxides generally have positive surface free energies, we can see that adsorption will decrease this energy and is therefore strongly favored.

5.5.4.1 The Relation between Concentration and Adsorption: Langmuir and Freundlich Isotherm

Consider the adsorption of aqueous species M at a surface site which we will denote as S. The reaction may be written as:

$M + S \rightleftharpoons M \cdot S$

 $\frac{d\Theta_M}{dt} = k_-\Theta_M$

We will denote the fraction of surface sites occupied by M as Θ_M , the rate constant for adsorption as k_+ , and that for desorption as k_- . The fraction of free sites is then $(1 - \Theta_M)$, and we explicitly assume that M is the only species adsorbed from solution. The rate of adsorption is then:

$$\frac{dM}{dt} = k_+[M](1 - \Theta_M)$$
5.126

At equilibrium, the rate of adsorption and desorption will be equal, so

$$k_{-}\Theta_{M} = k_{+}[M](1 - \Theta_{M})$$
5.128

Solving 5.128 for $\Theta_{M'}$ we obtain:

$$\Theta_{\rm M} = \frac{k_+/k_-[M]}{1 + k_+/k_-[M]}$$
 5.129

which expresses the fraction of site occupied by M as a function of the concentration of M. Since at equilibrium:

$$\mathbf{K}_{ad} = \frac{[M]_{ads}}{[M]_{aa}} = \frac{k_+}{k_-}$$
(5.43)

where K_{ad} is the equilibrium constant for adsorption, equation 5.129 becomes:

$$\Theta_M = \frac{K_{ad}[M]}{1 + K_{ad}[M]}$$
5.130

Equation 5.130 is known as the *Langmuir Isotherm*^{*}. Since this is a chapter on kinetics, we have derived it using a kinetic approach, but it is a statement of thermodynamic equilibrium and can be readily derived from thermodynamics as well. From the definition of Θ_M , we may also write the Langmuir isotherm as:

$$\Gamma_{M} = \Gamma_{M}^{\max} \frac{K_{ad}[M]}{1 + K_{ad}[M]}$$
5.131

5.127

^{*} An admittedly odd name for this equation. It is named for Irving Langmuir (1881-1957). Langmuir obtained a PhD from the University of Göttingen and spent most of his career working for General Electric Company. While trying extending the life of light bulbs, Langmuir carried out experiments on the adsorption of gases by metals. He developed this equation to describe his results. He won the Nobel Prize for Chemistry in 1932. The term "Isotherm" arises because such descriptions of adsorption are valid only for 1 temperature (i.e., K_{ads} is temperature dependent, as we would expect).

CHAPTER 5: KINETICS

A. White

Example 5.10. The Langmuir Isotherm

Consider a suspension of 1 mol/l of FeOH. Assuming an adsorption site density of 0.1 mol/mol and

K for adsorption of Sr on FeOH of 10⁵, how will the Sr adsorption density vary with the concentration of Sr in the solution? Assume that no other ions are present in the solution.

Answer: We can use equation 5.131 to solve this problem. $\Gamma_{\rm M}^{\rm max}$ in this case is 0.1 mol/mol. Using this value and K_{ads} of 10^5 in this equation, we obtain the result shown in Figure 5.26. The inset shows that at concentrations less than about 4 μ M, the adsorption density rises linearly with concentration. At higher concentrations, the adsorption density asymptotically approaches the maximum value of 0.1 mol.



where Γ_{M}^{max} is the maximum observed adsorption. Thus the Langmuir Isotherm predicts a maximum adsorption when all available sites are occupied by M. At large concentrations of M, then:

$$\Gamma_M = \Gamma_M^{\max}$$
 5.132

Where the concentration of M is small such that $K_{ad}[M] \ll 1$, equation 5.130 reduces to:

$$\Theta_M \cong \mathrm{K}_{\mathrm{ad}}[M]$$
 5.133

This equation simply says that the fraction of sites occupied by M is proportional to the concentration of M in solution.

The Freundlich Isotherm, which is purely empirical, is:

$$\Theta_{M} = \mathbf{K}_{\mathrm{ad}}[M]^{n}$$
 5.134

where *n* is any number. At low concentrations of M, the Langmuir isotherm reduces to the Freundlich isotherm with n = 1, i.e., the amount adsorbed is a linear function of the concentration in solution.

5.5.5 Catalysis

The International Union of Pure and Applied Chemistry (IUPAC) defines *catalyst* as follows: *A catalyst is a substance that increases the rate without modifying the overall standard Gibbs energy change in the reaction; the process is called catalysis, and a reaction in which a catalyst in involved is known as a catalyzed reaction.*

Another definition of a catalyst is a *chemical species that appears in the rate law with a reaction order greater than its stoichiometric coefficient.* This latter definition makes is clear that a catalyst may be involved in the reaction as a reactant, a product, or neither. If it is a reactant or product, its presence affects the reaction rate to a greater extent than would be predicted from the stoichiometry of the reaction.

We can distinguish two kinds of catalysis. Homogeneous catalysis refers to a situation in which the catalyst is present in the same phase in which the reaction is occurring (necessarily a solution). Examples of homogeneous catalysts of geochemical reactions include acids and a collection of organic mole-

M. White CHAPTER 5: KINETICS

cules called enzymes. Catalysis that occurs at the interface between two phases is referred to as heterogeneous catalysis. We will focus primarily on heterogeneous catalysis here. Heterogeneous catalysts are commonly simply surfaces of some substance. A familiar, but non-geochemical, example is the platinum in the catalytic converter of an automobile, which catalyzes the further oxidation of gasoline combustion products.

Catalysts work by providing an alternative reaction path with lower activation energy. In many cases, the lowering of the activation energy arises when reacting species are adsorbed. The heat liberated by the adsorption (ΔH_{ads}) is available to contribute toward the activation energy. For example, consider the reaction:

$$A + B \rightarrow C$$

having an activation energy E_A. A solid catalyst of this reaction would provide the following alternate reaction mechanism:

$$A + S \rightarrow A \cdot S$$

$$B + S \rightarrow B \cdot S$$

$$B \cdot S + A \cdot S \rightarrow C \cdot S$$

$$C \cdot S \rightarrow C + S$$

The net heat of adsorption for this process is:

$$\Delta H_{ad} = \Delta H_{ad}^{A} + \Delta H_{ad}^{B} - \Delta H_{ad}^{C}$$

Recalling that enthalpy is related to activation energy, we can write the activation energy for the catalyzed reaction as:

 $E_{\Lambda}^{cat} = E_{\Lambda} + \Delta H_{ad}$ 5.135

$$a_A = a_A + a_A + a_A + a_A$$
 so negative (i.e., heat liberated by adsorption), the acti-

If ΔH_{ad} is vation energy is lowered and the reaction proceeds at a faster rate than it otherwise would.

As we noted earlier, a surface will have a variety of sites for adsorption/desorption and surface reactions on a microscopic scale. Each site will have particular activation energy for each of these reactions. The activation energies for these processes will, however, be related. Sites with large negative adsorption energies also will be sites with low activation energies for surface reactions. On the other hand, if a site has a large negative adsorption energy, the desorption energy will be large and positive and desorption inhibited. If either the activation energy or the desorption energy is too large, catalysis of the overall reaction will be inhibited. What is required for fast overall reaction rates is a site where some compromise is achieved. In general, reaction and desorption energies will be related as:

$$\Delta G_r = -n\Delta G_d \tag{5.136}$$

where *n* is some constant. The presence of several sites on a solid surface results in several possible reaction paths. The fastest reaction path, that is the path that optimizes n, will dominate the reaction.

Surfaces of semiconductors (metal oxides and sulfides) can catalyze oxidation-reduction reactions (e.g., Wehrli, et al., 1989). For example, both TiO_2 and Al_2O_3 can catalyze the oxidation of vanadyl, V(IV), to vanadate, V(V). Figure 5.27 compares the rate of reaction in the presence of TiO₂ solid to the homogeneous reaction, demonstrating the reaction is substantially faster in the presence of TiO₂. The reaction mechanism for the surface catalyzed reaction may be described as follows (Figure 5.28):

$$2$$
>TiOH + VO²⁺ \rightleftharpoons >(TiO)₂VO + 2H⁺
>(TiO)₂VO + TiOH \rightleftharpoons >(Ti)₃VO₄²⁻ + H⁺





Figure 5.27. Oxygenation of vanadyl at pH 4 and $P_{O_2} = 1$ atm in experiments of Wehrli and Stumm (1988). After Wehrli et al. (1989).

W. M. White

CHAPTER 5: KINETICS



Figure 5.28. Mechanism of oxygenation of surface-bound vanadyl. In step (1) vanadyl is adsorbed at a TiO₂ surface ($a \rightarrow b$). Note that the vanadium is bound to two surface TiO groups. In step (2), the vanadium binds to a third surface oxygen, releasing an H⁺ ion ($b \rightarrow c$). In step (3), the vandate ion is replaced at the surface by 3 H⁺ ions ($c \rightarrow d$) (at intermediate pH, most vanadate will remain bound to the surface).

$$(Ti)_3VO_4^2 + H_2O + 2OH^- \rightleftharpoons 3 > TiOH + HVO_4^2$$

where >Ti indicates the Ti atom is part of a surface. The rate law for this reaction as determined by Wehrli and Stumm (1988) is:

$$-\frac{d\{V(IV)\}}{dt} = k\{VO(OTi<)\}[O_2]$$

where the {} brackets denote surface concentrations. The reaction is thus second order, depending on the concentration of surface bound V(IV) and dissolved O_2 . Wehrli and Stumm (1988) determined the rate constant for this reaction to be 0.051 $M_{.1}s_{.1}$ and the activation energy to be 56.5 kJ/mol at pH 7.

The surface catalyzed reaction is essentially independent of pH, whereas the reaction in homogeneous solution is strongly pH dependent. The rate law for the latter can be written as:

$$-\frac{d\{\mathcal{V}(IV)\}}{dt} = k\{\mathcal{VO}^+\}[\mathcal{O}_2][\mathcal{H}^+]$$

The apparent rate constant for this reaction is $1.87 \times 10^{-6} \text{ s}^{-1}$ and the apparent activation energy is 140 kJ/mol. Part of the difference in the activation energies can be accounted for as the energy of the hydrolysis reaction:

$$VO^{2+} + H_2O \rightleftharpoons VO(OH)^+ + H^+$$

which is the first step in the homogeneous reaction. This energy is 54.4 kJ/mol. Wehrli and Stumm (1988) speculated that the remainder of the difference in activation energy is the energy required for the transition from the octahedral structure of the dissolved VO(OH)⁺ ion to the tetrahedral structure of the dissolved vanadate ion.

W. M. White

CHAPTER 5: KINETICS

5.6 Kinetics of Dissolution and Leaching

5.6.1 Simple Oxides

The rate of dissolution of non-ionic solids are generally controlled by surface reactions at the solidwater interface. Absorption of ions to the surface of the solid play a critical role in the dissolution process. Adsorption of H^+ and OH^- ions at the surface appears to dominate dissolution reactions; however, adsorption of other species, particularly organic ones such as carboxylic acids, can be important as well.

Consider the example of a simple oxide (e.g., Al_2O_3) illustrated in Figure 5.29. As we noted earlier, "dangling" oxygens on surfaces in contact with aqueous solution will be protonated under most circumstances, i.e., an H⁺ ion will react with one of the surface O ligands to form a surface hydroxyl. Bonding of a single proton to a surface oxygen merely replaces the bond that the oxygen would have formed with a metal ion had it been located in the crystal interior. Addition of a second proton, i.e., protonation of the surface hydroxyl, however, has the consequence of weakening metal-oxides bonds.

In the case of a trivalent ion such as Al_2O_3 , protonation of three such bonds effectively frees the ion from the lattice structure. We can expect, therefore, that the dissolution rate will be proportional to finding 3 protonated ligands surround a single surface metal ion.

The concentration of surface-bound protons, $[H^+]_s$ can be related to the concentration of H^+ in solution through an absorption equilibrium constant K_{ad} , so that:

$$\mathrm{H}^{+}]_{\mathrm{s}}/S = \mathrm{K}_{\mathrm{ad}}[\mathrm{H}^{+}]_{\mathrm{aq}}$$

where S is the density of surface sites. The probability of finding a metal surrounded with three protonated ligands is then proportional $([H^+]_s/S)^3$. Thus we expect the dissolution rate to be proportional to the third power of the surface protonation:



Figure 5.29. Cartoon of proton-promoted dissolution of an oxide such as Al_2O_3 at a surface step. After Stumm and Wollast (1990).

CHAPTER 5: KINETICS



Figure 5.30. Log of the rate of Al_2O_3 dissolution plotted against the log of the concentration of surface protons. The slope of 3 indicates a rate law with third order dependence on the surface concentration of protons. From Stumm and Wollast (1990).

than that of simple oxides because they typically contain several metals bound in different ways. This can result in incongruent dissolution, i.e., some metal ions may be released to solution more rapidly than others (though experiments suggest dissolution is most often congruent). A related, and particularly important, factor is lattice structure, in particular the degree to which the individual silica tetrahedra share oxygens. There is a complete range among silicates in this respect, from orthosilicates, such as olivine, in which no oxygens are shared, to the tecto-, or framework-, silicates, such as quartz and the feldspars, in which all oxygens are shared. As we discussed in Chapter 4, shared oxygens are termed bridging, and non-shared ones *non-bridging* oxygens. Sharing of oxygens increases the degree of *polymerization* of the structure.

The degree of polymerization is important in the context of dissolution because the non-bridging bonds are much more reactive than the bridging ones. Minerals with highly polymerized structures, such as feldspars, dissolve slowly and are subject to leaching, as components (particularly the network-modifiers) may be dissolved out leaving the silicate framework still partially intact. Silicates with a low fraction of shared oxygens dissolve more rapidly and more uniformly. An example is olivine, whose structure is illustrated in Figure 5.31a. Once the Mg ions surrounding it are removed, the individual silica tetrahedra are no longer bound to the

Figure 5.30 shows that this is indeed the case for Al_2O_3 .

5.137

Deprotonation of surface OH groups will occur at high pH through the following reaction:

S-OH +

where S—O denotes a surface-bound oxygen. This deprotonation disrupts metal-oxygen bonds through polarization of electron orbitals. As a result, dissolution rates will also increase with increasing pH in alkaline solutions. Adsorption of protons at the surface is thought to be fast, hence equilibrium between adsorbed and aqueous protons is quickly attained. Thus detachment of the metal species becomes the rate-determining step. Other ligands, particularly organic ones such as oxalates, will have a similar effect. The overall dissolution rate is given by (Stumm and Wollast, 1990) as:

$$\Re = k_H \{ [H_s^+] / S \}^i + k_{OH} \{ [OH_s^-] / S \}^i + k_L \{ [L_s] / S \}^i + k_{H_2O} \}$$

where *i* is the charge on the metal ion.

 $OH^- \rightleftharpoons S - O^- + H_2O$

6.5.2 Silicates

Surface protonation and deprotonation also play a dominant role in silicate dissolution (e.g., Blum and Lasaga, 1988). However, the dissolution of silicates is somewhat more complex a. Olivine Structure



Figure 5.31. Comparison of olivine (forsterite) and feldspar (albite) structures. In feldspar, all oxygens are shared by adjacent tetrahedra, in olivine none are; instead the excess charge of the SiO_4^{4-} units is compensated by 2 Mg²⁺.

W. M. White CHAPTER 5: KINETICS

mineral, and are free to form H₄SiO₄ complexes in the solution (a more likely mode of dissolution is replacement of Mg²⁺ by 2H⁺; in essence, this produces a free H₄SiO₄ molecule). In contrast, removal of Na⁺ by H⁺ in albite (Figure 5.31b) leaves the framework of tetrahedra largely intact. The rate of weathering can also be affected by the Al/Si ratio, as the silicate groups are less reactive than the aluminate ones. Thus the dissolution rate of plagioclase depends on the ratio of the anorthite (CaAl₂Si₂O₈) to albite (NaAlSi₃O₈) components, with calcic plagioclase weathering more rapidly (e.g., Oxburgh, et al., 1994).

Some idea of the role these factors play can be obtained from Table 5.3, which lists the mean lifetimes of a 1 mm crystal for a variety of minerals in contact with a solution of pH 5 based on experimentally determined dissolution rates.

There are four important classes of reactions From Lasaga et al. (1994). involved in silicate dissolution and leaching: hydration, ion exchange, leaching, and hydrolysis.

TADLE 5.3.	Dissolution	RATES AND MEAN
Lifetimes of	CRYSTALS AT	25°C and pH 5
Mineral	Log Rate	Mean Lifetime
	$(mol/m^2/s)$	years
Quartz	-13.39	34,000,000
Kaolinite	-13.28	6,000,000
Muscovite	-13.07	2,600,000
Epidote	-12.61	923,000
Microcline	-12.50	579,000
Albite	-12.26	-575,000
Sanidine	-12.00	291,000
Gibbsite	-11.45	276,000
Enstatite	-10.00	10,100
Diopside	-10.15	6,800
Forsterite	-9.50	2,300
Nepheline	-8.55	211
Anorthite	-8.55	112
Wollastonite	-8.00	79

Hydration simply implies the addition of water to the structure. The effect of hydration may range from simply relaxation of the polymeric structure (e.g., swelling of vermiculite) to disruption of bridging bonds, to dissolution of polymeric fragments.

lon exchange involves replacement of a network-modifying cation by hydrogen ions. For example, Wollast and Chou (1992) showed that when freshly ground albite is mixed with water, there is an increase in Na_{aq}^{+} and an increase in pH, corresponding to the consumption of H_{aq}^{+} . This reaction may be represented as:

$$NaAlSi_3O_8 + H^+ \rightleftharpoons HAlSi_3O_8 + Na^+$$

This reaction, as well as the replacement of Mg²⁺ by H⁺ in forsterite, has been shown to be largely reversible. Wollast and Chou (1992) found that ion exchange occurs to a depth of about 20 Å in albite, corresponding to a depth of 2 or 3 unit cells.

Leaching involves the removal of an ion without replacement by an ion from solution. The consumption of H⁺ observed by Wollast and Chou (1992) was less than the production of Na⁺, so that much of the Na⁺ loss from the albite appears to result from leaching rather than ion exchange, i.e.:

$$NaAlSi_3O_8 \rightleftharpoons AlSi_3O_8^- + Na^+$$

This reaction, of course, results in the production of negative charge on the surface. Wollast and Chou (1992) found that the extent of leaching could be related to both pH and Na_{ao}^+ concentration:

$$X_{X^{-}} \approx 10^{-5.01} a_{H^{+}}^{-0.35} a_{Na^{+}}^{-0.65}$$
5.138

where X_{X^-} is the mole fraction of negatively charged surface species. Thus according to equ. 5.138, leaching increases with increasing pH and decreases with increasing aqueous Na concentration.

As we noted, structure affects the rate and degree of leaching. Sheet silicates (micas, clays, talc, serpentine) have relatively open structures though which water and solutes can be transported deeply into the structure, resulting in leaching of cations, including octahedrally coordinated Al and Mg. While the feldspar structure is not open, preferential removal of aluminate groups and charge-balance cations produces a porous structure, allowing penetration of water. Thus deep (400 Å) Na-, Ca-, and Al-poor, and Si-, H-rich layers have been observed on experimentally reacted plagioclase (Casey and Bunker, 1990).

Hydrolysis refers to the surface protonation and deprotonation reactions we have already discussed in the context of oxide dissolution. Hydrolysis has the effect of breaking of covalent metal-oxide bonds

CHAPTER 5: KINETICS

M. White

in the polymer structure by replacing them with an O–H bonds. The effect is the same as replacing one of the oxygens in the tetrahedron by an OH group. Where a bridging oxygen is involved, hydrolysis decreases the degree of polymerization of the structure and eventually leads to its destruction. Complete hydrolysis of a silica tetrahedron results in the formation of a free H_4SiO_4 molecule.

This process appears to be of critical importance in the dissolution of silicates, as was the case for oxides. Many silicate dissolution experiments have shown a dependence of dissolution rate on pH of the form:

$$\Re = ka_{ii}^n \qquad 5.139$$

with the value of *n* less than one. Blum and Lasaga (1988, 1991) showed that the dissolution rate of albite can be directly related to the surface concentration of positive species ($S-OH_2^+$) under acidic conditions and to the concentration of negative surface species ($S-O^-$) under basic regions. In other words:

under acidic conditions:

$$\Re = k_i [\mathbf{S} - \mathbf{OH}_2^+]$$
 5.140a

and under basic conditions:

$$\Re = k_2[\mathbf{S} - \mathbf{O}^-] \qquad 5.140\mathbf{b}$$

This dependence explains the fractional dependence of dissolution rate on pH. The reason is that the concentrations of $S-OH_2^+$ and $S-O^-$ species can be related to pH through Freundlich isotherms (equation 5.134). In the case of albite dissolution, Blum and Lasaga (1991) found:

$$[S - OH_2^+] = K_1 a_{H^+}^{0.52}$$
 5.141a

$$\left[S - O^{-}\right] = \mathbf{K}_{2} a_{OH^{-}}^{0.37} = \mathbf{K}_{3} a_{H^{+}}^{-0.37} \quad 5.141 \mathrm{t}$$

for acidic and basic conditions respectively. This relationship is shown in Figure 5.32a. Substituting 5.141a and b into 5.140a and b, we expect:

$$\Re = k_3 a_{H^+}^{-0.52}$$
 5.142a

and
$$\Re = k_4 a_{\mu^+}^{-0.37}$$
 5.142b

This matches well the pH dependencies determined experimentally by Chou and Wollast (1985), who found the exponents in equations 5.136 were 0.49 and -0.30 for the acidic and basic conditions respectively, as is shown in Figure 5.32b. Although the experimental data have been questioned, a similar relationship between abundance of surface species and dissolution rate has been claimed for olivine.

Ganor et al. (1995) demonstrated that the dissolution rate of kaolinite $(Al_2Si_2O_5[OH]_4)$ also shows a fractional



Figure 5.32. (a) Relationship between pH and the absolute value of net surface charge on dissolving albite particles. A Freundlich isotherm (dashed red lines) can be fit to the data. (b) Relationship between log of the albite dissolution rate and pH determined by Chou and Wollast (1985). The slope passing through the step parts of the curve (dashed red lines) is similar to slopes in (a), indicating adsorption and desorption of protons controls the dissolution rate. From Blum and Lasaga (1991).

W. M. White

Geochemistry

CHAPTER 5: KINETICS



Figure 5.33. The structure of kaolinite. Kaolinite consist of a layer of Si-tetrahedra linked to a layer of Al-Octahedra through a shared oxygen.

exponential dependence on pH (equation 5.140) with the value of the exponent, n, being 0.4 ± 0.2 for the pH range 3 to 4. Consistent with earlier studies, they concluded the form of the rate equation reflected the equilibrium adsorption of protons on the mineral surface. Their interpretation of the details of the reaction mechanism, however, was somewhat different than the interpretation of the mechanism of Al_2O_3 dissolution of Stumm and Wollast (1990) discussed above. Ganor et al. (1995) argued that simultaneous protonation of all metal-oxide bonds was unlikely and furthermore unnecessary, since stepwise breaking of metal oxide bonds would

do the job. They argued that hydrogen ion-mediated breaking of Al–O–Si bonds was the critical and rate-determining step in kaolinite dissolution. In support of this hypothesis, they point out that *ab initio* (i.e., from first principles) calculations show that the activation energy for hydrolysis of this bond is lower than that of Si–O–Si bonds. The kaolinite structure (Figure 5.33) consists of alternating sheets of Si-tetrahedra and Al-octahedra, with each Si-tetrahedron sharing an oxygen with an Al in the octahedral layer. Breaking these Al–O–Si bonds effectively "unzips" the octahedral and tetrahedral sheets. Subsequent hydrolysis of the individual metal-oxygen bonds is then fast.

In focussing on the effects of pH in our discussion of dissolution thus far, we have implicitly assumed that dissolution reactions take place far from equilibrium. We have also ignored the effects of temperature. Clearly, temperature and the approach to equilibrium must be taken into account in a full treatment of dissolution. Furthermore, other dissolved species might either catalyze or inhibit dissolution reactions. Lasaga et al. (1994) have proposed the following rate equation to take account of these additional factors:

$$\Re = k_0 e^{-E_A / RT} f(\Delta G) A_{\min} a_{H^+}^n \prod_i a_i^{m_i}$$
 5.143

The $k_0 e^{-E_d/RT}$ term is the usual Arrhenius expression for temperature dependence. A_{min} is the surface area of the dissolving mineral, the term $a_{H^+}^n$ takes account of the pH dependence, and the $a_i^{m_i}$ terms take account of the inhibitory or catalytic effects of other ions; n and m_i may take any value. The $f(\Delta G)$ term is some function of ΔG that expresses the dependence of the rate on the deviation from equilibrium. For instance, we saw that transition state theory predicts that $f(\Delta G)$ for an elementary reaction takes the form of equation 5.59, i.e.:

EKGIES TOK DISSOLUTION NEACTIONS			
Mineral	E _a	pН	
	kJ//mol	-	
Albite	54.4	neutral	
Albite	32.2	Basic	
Albite	117.2	<3	
Epidote	82.9	1.4	
Kaolinite	29.3	3–4	
Microcline	52.3	3	
Quartz	71.2	7	
Sanadine	54.0	3	
Wollastonite	79.1	3-8	

 Table 5.4.
 Apparent Activation En

 Encies for Dissolution Reactions

modified from Lasaga et al. (1994).

For an overall reaction, $f(\Delta G)$ might have the form (equation 5.60):

Even if the exact form of the rate equation is not known, an apparent activation energy can be calculated to express the temperature dependence of reaction rate. In that case, however, the activation energy is valid only under a specific set of conditions. Values of such apparent activation energies for a few minerals are listed in Table 5.4.

Equation 5.143 predicts that the dissolution rate will slow as equilibrium between mineral and solution is approached, i.e., as the concentration of dissolved

W. M. White Chapter 5: Kinetics

components increases. This will occur when the rate of dissolution exceeds the rate of transport because the concentrations of dissolution products will build up at the mineral-water interface. In the steady-state, mass balance requires that the rate of dissolution (i.e., the rate at which aqueous species are produced at the surface) and transport (the rate at which components are removed from the solution adjacent the dissolving surface) must be equal. Thus overall weathering rates are controlled by a combination of surface kinetics and transport kinetics. In each individual situation, one or the other can be the rate-limiting step.

Surface reactions are most often rate-limiting in dissolution and weathering of silicate minerals at low temperature (25°C). Dissolution of readily soluble minerals (e.g., halite) and even moderately soluble minerals (e.g., gypsum) are, by contrast, usually limited by the rate at which the dissolving components can be transported away from the mineral-water interface. by advection and diffusion. As temperature increases, transport is increasingly likely to become rate-limiting. This is because the activation energy of diffusion in aqueous solution, typically 5-10 kJ/mol, is generally less than the activation energy of surface reactions (typically >30 kJ/mol; Table 5.4). Thus the diffusion rates increase more slowly with temperature than surface reaction rates. This point is illustrated for the case of calcite in Figure 5.33. At temperatures less than 75° C, growth and dissolution of calcite, a moderately soluble mineral, is effectively controlled by the surface reaction rate, while at temperatures greater than 125° C, diffusion is the rate-controlling step. Dissolution under hydrothermal and metamorphic conditions is most likely to be diffusion-controlled for most minerals (Guy and Schott, 1989).

5.7 Diagenesis

An introductory geology text might define diagenesis as the process through which a sediment is converted to a sedimentary rock. We will use diagenesis to refer to a number of physical and chemical processes that occur subsequent to deposition of a sediment, including compaction and expulsion of pore water, consumption of organic matter, and resulting changes in $p\varepsilon$. Some of the originally deposited phases dissolve in the pore water during diagenesis, other phases crystallize from the pore water. Some of these changes begin immediately after deposition, some only as a result of later deformation. Some occur as a result of moderately elevated temperature and pressure, though processes occurring at much higher temperature and pressure would be called metamorphism. Diagenesis and metamorphism form a continuum; though a geologist might volunteer a definite opinion on whether a particular specimen had been diagenetically or metamorphically altered, he would be hard pressed to come up with criteria to distinguish diagenesis from metamorphism that was not arbitrary. Here, we will briefly consider a few of these processes.



Figure 5.34. Log of steady state growth (a) and dissolution (b) of calcite as a function of temperature, comparing diffusion-controlled and surface reaction-controlled kinetics. The model assumes a 1 μ hydrodynamic boundary layer and saturation in the case of growth. From Murphy et al. (1989).

5.7.1 Compositional Gradients in Accumulating Sediment

Let's turn our attention to the early stages of diagenesis in a slowly accumulating sediment. Our first task is to decide upon a reference frame. There are two choices: we could choose a reference frame

Τ,

Concentration

 T_1

Concentration

W. M. White Chapter 5: Kinetics

fixed to a specific layer. In this case, the sedimentwater interface will appear to move upward with time. Alternatively, we can choose a reference frame that is fixed relative to the sediment-water interface, thus depth always refers to distance downward from that interface. As sediment accumulates, a given layer of sediment will appear to move downward in this reference frame. In this reference frame, we can express the change in concentration at some depth, x, as the sum of changes in the composition due to diagenesis plus the change in the composition of sediment flowing downward past our fixed reference point:

$$\left(\frac{\partial C}{\partial t}\right)_{x} = \frac{dC_{i}}{dt} - \omega \left(\frac{\partial C_{i}}{\partial x}\right)_{t} \qquad 5.144$$

where C_i is concentration some species *i*, and ω is the burial rate. The partial derivative on the left hand side refers to changes *at some fixed depth*, and the total derivative refers to diagenetic changes occurring in given layer, or horizon, undergoing burial. $(\partial C_i/\partial x)_t$ is the concentration gradient of *i* at some fixed time *t*. This equation allows us to convert a reference frame that is fixed relative to the sediment-water interface to one that is fixed relative to some sedimentary layer.

Now let's consider two extremes where equation 5.144 is particularly simple. In the first, steady-state is reached and there is no change with time, hence:

$$\left(\frac{\partial C}{\partial t}\right)_{x} = 0 \qquad 5.145$$

In other words, the concentration of i at some fixed depth below the water-sediment interface is constant. Under these circumstances then,

$$\frac{dC_i}{dt} = \omega \left(\frac{\partial C_i}{\partial x}\right)_t \qquad 5.146$$

This case is illustrated in Figure 5.35.

In the second extreme, there is no diagenesis and the composition of a given layer is determined only by what is initially deposited, thus:

$$\frac{dC_i}{dt} = 0 \qquad 5.147$$



Figure 5.35. Steady-state diagenesis. Concentration at a fixed depth x_1 below the surface remains constant, but layer A, whose depth increases with time due to burial, experiences a decreasing concentration with time. After Berner (1980).



Figure 5.36. Concentration profiles in a sediment in which the composition of the material changes with time, but there is no diagenesis. The composition of any given layer is fixed, but the composition at some fixed depth relative to the water-sediment interface, such as x_1 , changes with time. After Berner (1980).

The concentration change with time at some fixed depth is then due to change in the composition of the sediment moving downward past that point. Thus:

$$\left(\frac{\partial C_i}{\partial t}\right)_x = \omega \left(\frac{\partial C_i}{\partial x}\right)_t$$
 5.148

CHAPTER 5: KINETICS



Figure 5.37. Fluxes through a box in a sedimentary layer of unit lateral dimensions and thickness dx. Arrows labeled A, B, and D indicate advective, biodiffusive, and molecular diffusive fluxes. Loss or gain by the box due to these processes depends on the difference in the flux into and out of the box: dF/dx.

This case is illustrated in Figure 5.36.

The sediment consists both of solid particles and the water buried with the particles, the pore water. Assuming no other fluid is present (e.g., gas, petroleum) then the volume fraction of water in the sediment is equal to the porosity ϕ . The volume fraction of solids is then simply $1 - \phi$. Most sediments will undergo compaction as they are buried. This is due to the weight of overlying sediment (gravitational compaction). Gravitational compaction results in expulsion of pore water and in decrease in porosity with depth. In addition, dissolution and cementation will also affect porosity. Since the molar volume of a phase precipitating or dissolving (the most

important such phase is CaCO₃) will be different from its partial molar volume in solution, these processes will also result in motion of the pore water. When compaction occurs, the rate of burial of a sediment will not be equal to the sedimentation rate.

Now consider a box of sediment of thickness dx and unit length and width embedded within some sedimentary layer (Figure 5.37). We assume that the layer is of uniform composition in the lateral dimension, and therefore that there is no lateral diffusion, and that there is also no lateral advection of fluid. Within the box there are *C* moles of species *i*. If we chose our concentration units to be moles per volume, then the concentration is simply C_i .

Let's consider the processes that can affect the concentration of species *i* within the box. First of all, reactions occurring within the box might affect *i*. For example, oxidation and reduction will affect species such as Fe^{3+} , SO_4^{2-} , and Mn^{2+} , etc. If we are interested in the concentration of a dissolved species, then dissolution, crystallization, leaching, etc. will all change this concentration.

In addition to reactions occurring within the box, diffusion, advection, and bioturbation will also affect the concentration of *i* if there is a difference between the fluxes into and out of the box. Bioturbation is the stirring effect produced by the activity of animals that live in the sediment (the infauna). From a geochemical perspective, bioturbation is much like diffusion in that it results from the random motion of particles (even if these particles are of very different size from atoms and ions) and acts to reduce compositional gradients. Mathematically, we can treat the effect of bioturbation in a way similar to diffusion, i.e., we can define a bioturbation flux as:

$$J_B = -D_B \left(\frac{\partial C_i}{\partial x}\right)_t$$
 5.149

where D_B is the biodiffusion coefficient. Values of D_B for solid phases range from 10⁻⁶ cm²/sec in near shore clays to 10⁻¹¹ in deep sea pelagic sediments. The bioturbation coefficient will generally be different for solid species than for liquid ones. Since most animals live only in the upper few cm's or 10's of cm's, D_B will be a function of depth. In those circumstances, the time dependence of concentration is given by:

$$\left(\frac{\partial C}{\partial t}\right)_{x} = \left\{\frac{\partial (D_{B}(\partial C_{i} / \partial x))}{\partial x}\right\}_{t}$$
5.150

Since molecular diffusion through solids is much lower than through liquids, one can generally neglect diffusion in the solid and deal only with diffusion through the pore water. Because pore water

202

CHAPTER 5: KINETICS

W. M. White

only occupies a fraction, ϕ , of the total volume of sediment, the flux will be reduced accordingly. Thus the diffusion of a dissolved species will be:

$$J_{M} = -\phi D_{M} \left(\frac{\partial C_{i}}{\partial x} \right)$$
 5.151

where we have adopted the subscript M to denote molecular diffusion. Fick's second law becomes:

$$\left(\frac{\partial C_i}{\partial t}\right)_x = \frac{1}{\phi} \frac{\partial (\partial \phi D_M (\partial C_i / \partial x))}{\partial x}$$
 5.152

The advective flux is the product of the fluid (i.e., pore water) velocity times the concentration:

$$J_A = vC_i 5.153$$

To describe the rate of change in concentration in the box, we want to know the rate of reactions within it and the change in flux across it, as it is the change in flux that dictates what is lost or gained by the box. Combining all the fluxes into a single term, F_i , the rate of change of species *i* in the box is:

$$\frac{dC_i}{dt} = -\frac{\partial F_i}{\partial x} + \sum R_i$$
5.154

where the second term is the sum of the rates of all reactions affecting i. The flux term is negative because any decrease in flux over dx results in an increase in concentration within the box.

We can then use equation 5.144 to transform to a reference frame fixed relative to the sediment surface:

$$\left(\frac{\partial C_i}{\partial t}\right)_x = -\left(\frac{\partial F_i}{\partial x}\right)_t - \omega \frac{\partial C_i}{\partial x} + \sum R_i$$
5.155

The downward burial of sediment past point x can also be considered a flux. Combining this with the other flux terms, we have:

$$\left(\frac{\partial C_i}{\partial t}\right)_x = -\left(\frac{\partial F_i}{\partial x}\right)_t + \sum R_i$$
5.156

where F is the *net* flux of *i* in and out of the box and the last term is the rate of all internal changes, including chemical, biochemical, and radioactive, occurring within the box. Equation 5.156 is called the *Diagenetic Equation* (Berner, 1980). Let's now consider an example that demonstrates how this equation can be applied.

5.7.2 Reduction of Sulfate in Accumulating Sediment

Organic matter buried with the sediment will be attacked by aerobic bacteria until all dissolved O_2 is consumed. When O_2 is exhausted, often within tens of centimeters of the surface, consumption will continue anaerobically with sulfur in sulfate acting as the electron acceptor:

$$2\alpha CH_2O + SO_4^{2-} \rightarrow H_2S + 2HCO_3^{2-}$$
 5.157

where CH_2O represents organic matter generally and α is the number of organic matter carbon atoms reduced per sulfur atom. Let's assume that the rate of sulfate reduction depends only on the supply of organic matter and not on the abundance of sulfate. In this case:

$$\frac{d[CH_2O]}{dt} = -k[CH_2O]$$
5.158

We are greatly simplifying matters since there are a great variety of organic compounds in sediments each of which will have a different rate constant. To further simplify matters, we will assume (1) that conditions become anaerobic at the sediment-water interface, (2) that all consumption of organic matter occurs anaerobically, (3) that steady-state is achieved (i.e., $(\partial C/\partial t)_x = 0$), and (4) there is no compaction (and therefore no pore water advection) or bioturbation. Substituting 5.158 into 5.148, we have:

W. M. White

CHAPTER 5: KINETICS

$$-k[CH_2O] = \omega \left(\frac{\partial [CH_2O]}{\partial x}\right)_t$$
 5.159

Integrating, we obtain the concentration of organic matter as a function of depth:

$$f(x) = [CH_2O]^\circ e^{-kx/\omega}$$
 5.160

where $[CH_2O]^\circ$ is the organic matter concentration at the sediment-water interface (x = 0).

We can now also solve for the variation in concentration sulfate in the pore water. According to equation 5.9, the rate of sulfate reduction is related to the organic matter consumption rate as:

$$\frac{d\left[\mathrm{SO}_{4}^{2^{-}}\right]}{dt} = \frac{1}{2\alpha} \frac{d[\mathrm{CH}_{2}\mathrm{O}]}{dt} = \frac{k[\mathrm{CH}_{2}\mathrm{O}]^{\circ}}{2\alpha} e^{-kx/\omega} \qquad 5.161$$

Whereas the organic matter can be considered fixed in sediment, the sulfate is a dissolved species, so we must also consider diffusion. Making appropriate substitutions into 5.156, we have:

$$\phi D\left(\frac{\partial^2 \left[\operatorname{SO}_4^{2^-}\right]}{\partial x^2}\right) - \omega \left(\frac{\partial \left[\operatorname{SO}_4^{2^-}\right]}{\partial x}\right)_t - \frac{k \left[\operatorname{CH}_2 \operatorname{O}\right]^\circ}{2\alpha} e^{-kx/\omega} = 0$$

This is a second order differential equation and its solution will depend on the boundary conditions. Our boundary condition is that at x = 0, $C = C^{\circ}$. The solution under these conditions is:

$$[SO_4^{2-}] = \frac{\omega^2 [CH_2 O]^{\circ}}{2\alpha(\omega^2 + kD)\phi} (e^{kx/\omega} - 1) + [SO_4^{2-}]^{\circ} \quad 5.163$$

where $[SO_4^2]^\circ$ is the sulfate concentration at the surface.

Murray et al. (1978) applied this model to data from sediment cores taken from Saanich Inlet of British Columbia (Figure 5.38). Sedimentation rate, ω , was determined using ²¹⁰Pb (see Chapter 8) to be about 1 cm/yr, the factor α was independently estimated to be 0.5, and average porosity of 0.927. The value of D was taken to be 2.6×10^{-6} cm²/sec. They fit an exponential curve of the form $c = a e^{-bx}$ to the data and found a = 26.6 and b = 0.184. From this they determined the rate constant to be 6.1×10^{-9} sec⁻¹, and the initial concentration of metabolizable organic matter to be 380 mM/cm³ total sediment. The latter was somewhat larger than the value determined from the profile of total organic carbon in the core. To explain the discrepancy, the authors suggested methane is produced below the depth where sulfate is depleted. Methane then diffuses upward and is oxidized by sulfate-reducing bacteria.

References and Suggestions for Further Reading

Aagaard, P. and H. C. Helgeson. 1982. Thermodynamic and kinetic constraints on reaction rates among minerals and aqueous soluitons I. theoretical considerations. *Am. J. Sci.* 282: 237-285.

Albarède, F. 1995. Introduction to Geochemical Modeling. Cambridge: Cambridge Univ. Press.

- Anderson, D. E. 1981. Diffusion in Electrolyte Mixtures, in *Reviews in Mineralogy, Volume 8: Kinetics in Geochemical Processes* ed. A. Lasaga and R. Kirkpatrick. 211-260. Washington: Mineralogical Society of America.
- Berner, R. A. 1980. Early Diagenesis. Princeton: Princeton University Press.
- Berner, R. A. 1981. Kinetics of weathering and diagenesis, in *Reviews in Mineralogy, Volume 8: Kinetics in Geochemical Processes* ed. A. Lasaga and R. Kirkpatrick. 111-134. Washington: Mineralogical Society of America.

Blum, A. and A. Lasaga. 1988. Role of surface speciation in the low-temperature dissolution of minerals. *Nature*. 331: 431-433.



Fig. 5.38. Dissolved sulfate5cq6gentrations in sediments from the Saanich Inlet. Data are shown as open circles with error bars. Curve is fitted using equation 5.163 and parameters given in the text. From Murray et al. (1978).

W. M. White

CHAPTER 5: KINETICS

- Blum, A. E. and A. C. Lasaga. 1991. The role of surface speciation in the dissolution of albite. *Geochim. Cosmochim. Acta*. 55: 2193-2201.
- Broecker, W. S. and V. M. Oversby, 1971. *Chemical Equilibria in the Earth,* Chapter 5. New York: McGraw-Hill.
- Carlson, W. D. 1980. Experimental Studies of Metamorphic Petrogenesis, PhD thesis, Los Angeles: UCLA.
- Casey, W. H. and Bunker, 1990. The leaching of mineral and glass surfaces during dissolution, in *Reviews in Mineralogy, Volume 23: Mineral-Water Interface Geochemistry* ed. M. Hochella and A. White. 397-426. Washington: Mineralogical Society of America.
- Chermak, J. A. and J. D. Rimstidt. 1990. The hydrothermal transformation rate of kaolinite to muscovite/illite. *Geochim. Cosmochim. Acta*. 54: 2979-2990.
- Chou, L. and R. Wollast. 1985. Steady-state kinetics and dissolution mechanisms of albite. *Amer. J. Sci.* 285: 965-993.
- Crank, J. 1975. The mathematics of diffusion, 2nd edition, Oxford: Clarendon Press.
- Fisher, G. W., 1978. Rate laws in metamorphsim, Geochim. Cosmochim. Acta, 42: 1035-1050.
- Ganor, J., J. L. Mogollon and A. C. Lasaga. 1995. The effect of pH on kaolinite dissolution rates and on activation energy. *Geochim. Cosmochim. Acta*. 59: 1037-1052.
- Guy, C. and J. Schott. 1989. Multisite surface reaction versus transport control during the hydrolysis of a complex oxide. *Chem. Geol.* 78: 181-204.
- Heinrich, W., P. Metz and M. Gottshalk. 1989. Experimental investigation of the kinetics of the reaction 1 tremolite + 11 dolomite \rightleftharpoons 8 fosterite + 13 calcite + 9CO₂ + 1 H₂O. *Contrib. Mineral. Petrol.* 102: 163-173.
- Hochella, M. and White. A. (ed.), 1990. *Mineral-Water Interface Geochemistry: Reviews in Mineralogy, Volume 23:*. Washington: Mineralogical Society of America.
- Hofmann, A. W., 1980. Diffusion in natural silicate melts, a critical review, in *Physics of Magmatic Processes* ed. R. B. Hargraves. Princeton: Princeton Univ. Press.
- Kerrick, D. M., A. C. Lasaga, and S. P. Raeburn. 1991. Kinetics of heterogeneous reactions, in *Contact Metamorphism*, vol. 26, ed. D. M. Kerrick, 583-722. Washington: Mineralogical Society of America.
- Kirkaldy, J. S. and D. J. Young. 1985. Diffusion in the Condensed State. London: Institute of Metals.
- Kirkpatrick, R. J., 1981. Kinetics of crystallization of igneous rocks, in *Reviews in Mineralogy, Volume 8: Kinetics in Geochemical Processes* ed. A. Lasaga and R. Kirkpatrick.31-398. Washington: Mineralogical Society of America.
- Laidler, K. J. 1987. Chemical Kinetics. New York: Harper Collins.
- Lasaga, A. C., 1981a. Rate laws of chemical reactions, in *Reviews in Mineralogy, Volume 8: Kinetics in Geochemical Processes* ed. A. Lasaga and R. Kirkpatrick. 1-68. Washington: Mineralogical Society of America.
- Lasaga, A. C.1981b. Transition State Theory. in *Kinetics of Geochemical Processes, Reviews in Mineralogy 8*, ed. A. C. Lasaga and R. J. Kirkpatrick. 135-170. Washington: Mineralogical Society of America.
- Lasaga, A. C., 1997. Kinetic Theory in the Earth Sciences, Princeton NJ, Princeton University Press, 811.
- Lasaga, A. C. and R. J. Kirkpatrick (editors), 1981. *Reviews in Mineralogy, Volume 8: Kinetics in Geochemical Processes,* Washington: Mineralogical Society of America.
- Lasaga, A. C., J. M. Soler, G. J., T. E. Burch and K. L. Nagy. 1994. Chemical weathering rate laws and global geochemical cycles. *Geochim. Cosmochim. Acta*. 58: 2361-2386.
- Lennie, A. R. and D. J. Vaughan. 1992. Kinetics of the marcasite-pyrite transformation: an infrared study. *Am. Mineral.* 77: 1166-1171.
- Morioka, M. and Nagasawa, H., 1990. Ionic diffusion in olivine, in *Advances in Physical Geochemistry*, *Volume 8, Diffusion, Atomic Ordering, and Mass Transport* ed. ed. J. Ganguly. 176-197. New York: Springer-Verlag.
- Murphy, W. M., E. H. Oelkers and P. C. Lichtner. 1989. Surface reaction versus diffusion control of mineral dissolution and growth rates in geochemical processes. *Chem. Geol.* 78: 357-380.
- Murray, J. W., V. Grundmanis, and W. M. Smethie, 1978. Interstital water chemistry in the sediments of Saanich Inlet, *Geochim. Cosmochim. Acta*, 42: 1011-1026.

W. M. White

CHAPTER 5: KINETICS

- Oxburgh, R., J. I. Drever and T.-Y. Sun. 1994. Mechanism of plagioclase dissolution in acid solution at 25°C. *Geochim. Cosmochim. Acta*. 58: 661-670.
- Parks, G. A., 1990. Surface energy and adsorption at mineral/water interfaces: an introduction, in *Reviews in Mineralogy, Volume 23: Mineral-Water Interface Geochemistry* ed. M. Hochella and A. White .133-175. Washington: Mineralogical Society of America.
- Pearson, D. G., G. R. Davies and P. H. Nixon, 1995. Orogenic ultramafic rocks of UHP (diamond faces) origin, in R. G. Coleman and Wang, X. (ed.), *Ultrahigh Pressure Metamorphism*, 456-510.
- Schott, J., R. A. Berner and E. L. Sjöberg. 1981. Mechanism of pyroxene and amphibole weathering—I. Experimental studies of iron-free minerals. *Geochim Cosmochim Acta*. 45: 2123-2135.

Sposito, G., 1989. The Chemistry of Soils, New York: Oxford Univ. Press.

Stumm, W. (ed.), 1990. Aquatic Chemical Kinetics: Reaction Rates of Processes in Natural Waters, New York: Wiley-Interscience.

Stumm, W. and R. Wollast. 1990. Coordination chemistry of weathering: Kinetics of the surface-controlled dissolution of oxide minerals. *Rev. Geophys.* 28: 53.

Toramaru, A. 1991. Model of nucleation and growth of crystals in cooling magmas. *Contrib. Mineral. Petrol.* 108: 106-117.

Tyrell, H. J. V. and K. R. Harris. 1984. Diffusion in Liquids. London: Butterworths.

Watson, E. B., Basalt contamination by continental crust: some experiments and models, *Contrib. Mineral. Petrol.*, 80, 73-87, 1982.

- Wehrli, R. and W. Stumm. 1988. Oxygenation of vanadyl(IV): Effect of coordinated surface hydroxyl groups and OH⁻. *Langmuir*. 4: 753-758.
- Wehrli, B., B. Sulzberger and W. Stumm. 1989. Redox processes catalyzed by hydrous oxide surfaces. *Chem. Geol.* 78: 167-179.
- Wollast, R. and L. Chou. 1992. Surface reactions during the early stages of weathering of albite. *Geochim. Cosmochim. Acta*. 56: 3113-3121.

Wood, B. J. and J. V. Walther. 1983. Rates of hydrothermal reactions. Science. 222: 413-415.

- Liang, Y., F. M. Richter and L. Chamberlin, 1997. Diffusion in silicate melts: III. Empirical models for multicomponent diffusion, *Geochim Cosmochim Acta*, 61: 5295-5312.
- Zhong, S. and A. Mucci. 1993. Calcite precipitation in seawater using a constant addition technique: a new overall reaction kinetic expression. *Geochim. Cosmochim. Acta*. 57: 1409-1417.

Problems

1. (a) Assuming that the precipitation of calcite from aqueous solution occurs only through the reaction:

$$Ca^{2+} + CO_3^{2-} \rightarrow CaCO_{3(s)}$$

and that this reaction is *elementary*, write an equation for the rate of calcite precipitation.

(b) Assuming that the reaction above is reversible, i.e.:

$$Ca^{2+} + CO_3^{2-} \rightleftharpoons CaCO_{3(s)}$$

and still assuming that it is *elementary*, write an equation for the dependence of *net* rate of calcite precipitation on concentration and free energy change of reaction.

2. Zhong and Mucci (1993) found that at constant concentration of dissolved Ca^{2+} ([Ca^{2+}] ≈ 10.5 mmol/kg), the rate of calcite precipitation in seawater obeyed the following rate law:

$$\Re = K_f [\operatorname{CO}_3^{2-}]^3 - k_-$$

where $K_f = k_+ (a_{Ca^{2+}})^n \gamma_{CO_3}^3 = 10^{3.5}$ and $k_- = 0.29$ (\Re is in units of μ mol m⁻² h⁻¹).

(a) Is this rate law consistent with the mechanism of calcite precipitation in seawater being the elementary one described in problem 5.1 or with a more complex reaction mechanism? Justify your answer.

W. M. White

CHAPTER 5: KINETICS

(b) Using this rate law, predict the rate of calcite precipitation for concentrations of CO_3^{2-} of 0.04, 0.066, and 0.3 mmol/kg.

3. Ox	idation of methane	in the atmosphere	occurs through	a number of	RATES OF METHANE-
mecha	nisms, including rea	action with the hydr	roxyl radical:		HYDROXYL REACTION

$OH + CH_4 \rightleftharpoons H_2O + CH_3$

The rate of this reaction for a series of temperatures is shown in the adjacent table. Based on these data, estimate the activation energy and frequency factor for this reaction. (*HINT*: Try using linear regression).

4. Schrott et al. (1981) found that dependence on pH of the rate of dissolution of enstatite could be expressed as:

$$\Re = k a_{\mathrm{H}^{+}}^{\mathrm{n}}$$

where k shows a typical Arrhenius temperature dependence

where wono wou typical runnenaus temperature aepenaence.
a.) Reaction rates were measured at a series of pH values at constant tem-
perature (22° C). These data are shown in the adjacent table. Using these
data, estimate values of <i>k</i> and <i>n</i> for this temperature (<i>HINT</i> : try using linear
regression).

b.) Reaction rates were also determined at various temperatures at constant pH (6). Using these data, estimate the activation energy and frequency factor for the rate constant.

c.) Using your results from (a) and (b), estimate the rate of reaction (in moles Si released per sec per gram enstatite) at pH 4 and 30° C.

Marcasite and pyrite are polymorphs of FeS₂. Though pyrite has a 5. lower ΔG_f than marcasite, the latter often forms metastably. Lennie and Vaughan (1992) found that the kinetics of the marcasite to pyrite transformation follows a simple first order rate law:

$$-\frac{d\alpha}{dt} = k\alpha$$

where α is the volume fraction of marcasite and k has the usual Arrhenius temperature dependence with A = 2.76×10^{17} sec⁻¹ and E_A = 253 kJ/mol. Assuming a system consisting initially of pure marcasite, calculate the time required for one half of the marcasite to convert to pyrite (i.e., $\alpha = 0.5$) at 300° C and 350° C.

6. If ΔG is the free energy of reaction for the reaction:

$$A + BC \to AC + B \tag{1}$$

and assuming (1) this is an elementary reaction, (2) ideal behavior, and (3) it is a reversible reaction, show that the ratio of the forward and reverse rates of (1) is:

$$\frac{\mathfrak{R}_{+}}{\mathfrak{R}_{-}} = e^{-\Delta G/RT}$$
(5.57)

(HINT, start with equation 5.43).

7. On a temperature-pressure diagram, draw a line such that the time required for complete conversion of a 1 mm aragonite crystal to calcite will be complete within 10⁵ years. Assume spherically symmetric growth of calcite from a single nucleus in the center. Use the thermodynamic data in Table 2.2 and ΔG^* = 184 kJ (see Example 5.4).

1, C	ĸ	
25	6.60×15	
10	4.76×15	
0	3.76× ⁻¹⁵	
-10	2.93×15	
-25	1.93×15	

RATE OF ENSTATITE DISSOLUTION

	Rate
pН	moles Si/g-sec
1	2.75×10^{-10}
2	7.08×10^{-11}
6	2.82×10^{-13}

Т	Rate
°C	moles Si/g-sec
20	3.72×10^{-13}
50	2.34×10^{-12}
60	4.07×10^{-12}
75	8.13×10^{-12}

W. M. White

Geochemistry

CHAPTER 5: KINETICS

8. Using the data given in Example 5.5 for the reaction:

 $Ca_2Mg_5Si_8O_{22}(OH) + 11CaMg(CO_3)_2 \rightleftharpoons 8Mg_2SiO_4 + 13 CaCO_3 + 9CO_2 + H_2O_3 + 9CO_3 + 9CO_$

Tremolite + 11 Dolomite \Rightarrow 8 Forsterite + 13 Calcite + 9 CO₂ + H₂O

make a plot of the relative volume of dolomite (V_{Do}/V_{Do0}) as a function of time assuming an initial temperature of 620° C (the equilibrium temperature) and a heating rate of 0.1° per year.

(*HINT*: Because the reaction is fast, the overall temperature change will be small, so you may assume that T (i.e. absolute temperature) is constant. However, because the temperature is close to the equilibrium temperature, the change in the temperature overstep, ΔT , will be significant. Approximate ΔG in equation 5.72 as $\Delta S\Delta T$ and express ΔT as a function of time, $\Delta T = R_H t$ where R_H is the heating rate, then integrate.)

9. The transformation of kaolinite to illite (muscovite) may be written as:

$$K^+ + 1.5Al_2Si_2O_5(OH)_4 \rightleftharpoons H^+ + KAl_3Si_3O_{10}(OH)_2 + 1.5H_2O$$

Chermack and Rimstidt (1990) determined that the forward rate of reaction was:

$$\frac{d[\mathbf{K}^+]}{dt} = k_+[\mathbf{K}^+]$$

Forward and reverse rate constants for the reaction were determined to be:

 $\ln k_{+} = 12.90 - 1.87 \times 10^{4}/\text{T}$ and $\ln k_{-} = 6.03 - 1.21 \times 10^{4}/\text{T}$

a.) What are the activation energies for the forward and reverse reactions?

b.) What is the equilibrium constant for this reaction at 275 ° C ?

c.) Make a plot of log \Re_{net} vs. log ([H⁺]/[K⁺]) at 250° C assuming a K⁺ concentration of 2.0 x 10⁻⁶ M, ideal solution, that muscovite and kaolinite are pure phases, and that the forward and reverse reactions are elementary.

10. Using the data in Table 5.2, determine the diffusion coefficient for diffusion of Ba in an andesitic melt for $T = 1200^{\circ}$, 1100° , 1000° , and 900° C (assume the parameters are valid over this temperature range; remember to use thermodynamic temperature).

11. Using the data in Example 5.8, calculate the diffusion flux for Mg and Fe in garnet.

12. A remarkable feature of sediments recording the Cretaeceous-Tertiary Boundary is an enrichment in iridium (Ir) at the boundary, which is often marked by a boundary clay. Imagine a boundary clay 5 cm thick that initially has a uniform Ir concentration of 20 ppb. Assume and that sediments above and below the boundary clay contain negligible Ir. If the detection limit for Ir is 2 ppb, how thick would the Ir-enriched layer be after 60 million years if the diffusion coefficient for Ir is 10^{-15} cm²/sec?

13. Assuming a surface free energy of 10^4 J/cm^2 , $\overline{\text{V}} = 101 \text{ cc/mol}$, $\Delta H_m = 54.84 \text{ kJ/mol}$, and $T_m = 1118^\circ$ C, what is the critical radius for a spherical albite crystal growing in a pure albite melt that has been undercooled by 10° , 20° , and 30° ?

Make a plot of ΔG_{tot} as a function of crystal radius for each of these temperatures. (*HINT*: your scale should span only 10 or 20 use microns.)

14. Crystal growth and dissolution are reactions that involve both diffusion and surface reactions occurring in series (i.e., a component of a growing crystal must first be delivered to the surface, then incorporated in the growing crystal). Either of these processes can be the rate-limiting step at 25° C. Diffusion in aqueous solutions typically has an activation energy of 20 kJ/mol whereas surface reactions in aqueous solution typically have activation energies of 60-80 kJ/mol. Assuming the rates of diffusion and surface reaction for growth of a certain mineral from aqueous solution are approximately equal at 25° C, will diffusion or surface reaction be rate-limiting at 200° C?

W. M. White

Geochemistry

CHAPTER 5: KINETICS

15. Some anaerobic bacteria can utilize Mn^{4+} to oxidize organic matter. The reaction may be represented as:

 $2MnO_2(s) + CH_2O + H_2O \rightleftharpoons 2Mn^{2+} + CO_2 + 4OH^{-}$

In its oxidized form, Mn is highly insoluble and effectively immobile in sediment. However, in its reduced form, Mn is soluble and mobile. Imagine that at a depth of 50 cm in actively depositing marine sediments conditions become sufficiently reducing so that the reaction above occurs. Furthermore, assume that reaction is such that a constant concentration of 0.02 mM of Mn^{2+} is maintained at this depth and below. Above this depth, Mn^{2+} is oxidized and precipitated through reactions such as:

$$\mathrm{Mn}^{2+} + 2\mathrm{OH}^{-} + \frac{1}{2}\mathrm{O}_2 \rightleftharpoons \mathrm{MnO}_2(\mathrm{s}) + \mathrm{H}_2\mathrm{O}$$

Assuming (1) that the rate of the above reaction may be written as:

$$-\frac{d[\mathrm{Mn}^{2+}]}{dt} = k[\mathrm{Mn}^{2+}]$$

(2) *k* for this reaction is $1 \times 10^{-8} \text{ sec}^{-1}$, (3) the concentration of Mn^{2+} at the sediment-water interface is 0 and that diffusion from below is the sole source of Mn^{2+} between 0 and 50 cm, (4) D is 5×10^{-6} cm/sec, (5) a sedimentation rate of 1 cm/yr, (6) there is no advection, compaction, or bioturbation, and (7) a porosity of 0.85, make a plot of the concentration of dissolved Mn^{2+} vs. depth at steady state.

16. Diamond is remarkably stable at the surface of the Earth. Pearson et al. (1995) estimated that to convert 1 cc of diamond to graphite at 0.1Mpa and 1000°C would require 1 billion years, but only a million years would be required at 1200°C. From the difference in these rates, estimate the activation energy for the diamond-graphite transition.

CHAPTER O: AQUATIC CHEMISTRY

CHAPTER 6: AQUATIC CHEMISTRY

6.1 INTRODUCTION

and transport of dissolved and suspended matter. Beyond that, water is essential to life and central to human activity. Thus as a society, we are naturally very concerned with water quality, which in essence means water chemistry. Aquatic chemistry is therefore the principal concern of many geochemists.

In this chapter, we learn how the tools of thermodynamics and kinetics are applied to water and its dissolved constituents. We develop methods, based on the fundamental thermodynamic tools already introduced, for predicting the species present in water at equilibrium. We then examine the interaction of solutions with solids through precipitation, dissolution, and adsorption.

- Most reactions in aqueous solutions can be placed in one of the following categories:
- Acid-base, e.g., dissociation of carbonic acid:

$$H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

• Complexation, e.g., hydrolysis of mercury:

 $Hg^{2+} + H_2O \rightleftharpoons Hg(OH)^+ + H^+$

• Dissolution / Precipitation, e.g., dissolution of orthoclase:

$$AlSi_3O_8 + H^+ + 7H_2O \rightleftharpoons Al(OH)_3 + K^+ + 3H_4SiO_4$$

• Adsorption / Desorption, e.g., adsorption of Mn on a clay surface:

$$\equiv$$
S + Mn²⁺ \rightleftharpoons \equiv S-Mn

(where we are using \equiv S to indicate the surface of the clay).

In this chapter, we will consider these in detail. We return to the topic of aquatic chemistry in Chapter 13, we examine the weathering process, that is reaction of water and rock and development of soils, and the chemistry of streams and lakes.

6.2 Acid-Base Reactions

The hydrogen and hydroxide ions are often participants in all the foregoing reactions. As a result, many of these reactions are pH dependent. In order to characterize the state of an aqueous solution, e.g., to determine how much CaCO₃ a solution will dissolve, the complexation state of metal ions, or the redox state of Mn, the first step is usually to determine pH. On a larger scale, weathering of rock and precipitation of sediments depend critically on pH. Thus pH is sometimes called the *master variable* in aquatic systems. We note in passing that while pH represents the hydrogen ion, or proton concentration, the hydroxide ion concentration is easily calculated from pH since the proton and hydroxide concentrations are related by the dissociation constant for water, i.e., by:

$$K_W = a_{H^+} a_{OH^-}$$
 6.1

The value of K_W , like all equilibrium constants, depends on temperature, but is 10^{-14} at 25°C.

Arrhenius defined an *acid* as a substance that upon solution in water releases free protons. He defined a *base* is a substance that releases hydroxide ions in solution. These are useful definitions in most cases. However, chemists generally prefer the definition of Brønstead, who defined acid and base as proton donors and proton acceptors respectively. The strength of an acid or base is measured by its tendency to donate or accept protons. The dissociation constant for an acid or base is the quantitative measure of this tendency and thus is a good indication of its strength. For example, dissociation of HCI:

$$HCl \rightleftharpoons H^+ + Cl^-$$

W. M. White

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

has a dissociation constant:

$$K_{HCl} = \frac{a_{H^+} a_{Cl^-}}{a_{HCl}} = 10^3$$

HCl is a strong acid because only about 3% of the HCl molecules added will remain undissociated. The equilibrium constant for dissociation of hydrogen sulfide:

$$H_2S \rightleftharpoons H^+ + HS^-$$

 $K_{H_2S} = \frac{a_{H^+}a_{HS^-}}{a_{H_2S}} = 10^{-7.1}$

is:

Thus H₂S is a weak acid because very few of the H₂S molecules actually dissociate except at high pH.

Metal hydroxides can either donate or accept protons, depending on pH. For example, we can represent this in the case of aluminum as:

$$Al(OH)_{2}^{+} + H^{+} \rightleftharpoons Al(OH)^{2+} + H_{2}O$$
$$Al(OH)_{2}^{+} + OH^{-} \rightleftharpoons Al(OH)_{3}^{0} + H_{2}O$$

Compounds that can either accept or donate protons are said to be *amphoteric*.

Metals dissolved in water are always surrounded by solvation shells. The positive charges of the hydrogens in the surrounding water molecules are to some extent repelled by the positive charge of the metal ion. For this reason, water molecules in the solvation shell are more likely to dissociate and give up a proton more readily than other water molecules. Thus the concentration of such species will affect pH.

Most protons released by an acid will complex with water molecules to form hydronium ions, H_3O^+ or even $H_3O_2^+$. However, in almost all cases we need not concern ourselves with this and can treat the H^+ ion as if it were a free species. Thus we will use [H⁺] to indicate the concentration of $H^+ + H_3O^+ + H_3O_2^+ + ...$

6.2.1 PROTON ACCOUNTING, CHARGE BALANCE, AND CONSERVATION EQUATIONS

6.2.1.1 PROTON ACCOUNTING

Knowing the pH of an aqueous system is the key to understanding it and predicting its behavior. This requires a system of accounting for the H⁺ and OH⁻ in the system. There are several approaches to doing this. One such approach is the *Proton Balance Equation* (e.g., Pankow, 1991). In this system, both H⁺ and OH⁻ are considered components of the system, and the proton balance equation is written such that *the concentration of all species whose genesis through reaction with water caused the production of OH⁻ are written on one side, and the concentration of all species whose genesis through reaction with water caused the production of OH⁻ are written on the other side. Because water dissociates to form one H⁺ and one OH⁻, [H⁺] always appears on the left side and OH⁻ always appears on the right side of the proton balance equation. The proton balance equation for pure water is thus:*

$$[H^+] = [OH^-]$$
 6.2[†]

Thus in pure water the concentrations of H^+ and OH^- are equal.

Now, consider the example of a nitric acid solution. H⁺ will be generated both by dissociation of water and dissociation of nitric acid:

$$HNO_3 \rightleftharpoons H^+ + NO_3^-$$

Since one NO_3^- ion is generated for every H⁺, the proton balance equation becomes:

$$[H^+] = [OH^-] + [NO_3^-]$$
 6.3

[†] Be careful not to confuse algebraic expressions, written with an equal sign, such as the proton balance equation, with chemical reactions, written with the reaction symbol, \rightleftharpoons . In this case, it is obvious that this is not a balanced chemical reaction, but that will not always be the case.
Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

Next consider a solution of calcium carbonate. We specify the calcium and carbonate ions as components. Hydrogen ions may be generated by hydrolysis of calcium:

$$Ca^{2+} + H_2O \rightleftharpoons H^+ + Ca(OH)^+$$

and hydroxide ions may be generated by:

$$\mathrm{CO}_3^{2-} + \mathrm{H}_2\mathrm{O} \rightleftharpoons \mathrm{HCO}_3^{-} + \mathrm{OH}^{-}$$

The proton balance equation for this reaction is:

$$[H^{+}]+[HCO_{3}^{-}] = [OH^{-}]+[Ca(OH)^{-}]$$
6.4

Now consider a solution of a *diprotonic* acid such as H₂S. H₂S can undergo 2 dissociation reactions:

$$H_2S \rightleftharpoons H^+ + HS^-$$
 6.5

$$HS^- \rightleftharpoons H^+ + S^{2-}$$
 6.6

For every HS⁻ ion produced by dissociation of H_2S , one H⁺ ion would have been produced. For every S²⁻ ion, however, 2 H⁺ would have been produced, one from the first dissociation and one from the second. The proton balance equation is thus:

$$[H^+] = [OH^-] + [HS^-] + 2[S^{2-}]$$
6.7

An alternative approach to the proton balance equation is the *TOTH* proton mole balance equation used by Morel and Hering (1993). In this system, H^+ and H_2O are always chosen as components of the system but OH⁻ is not. The species OH⁻ is the algebraic sum of H_2O less H^+ :

$$H^{-} = H_2 O - H^{+}$$
 6.8

An implication of this selection of components is that when an acid, such as HCl is present, we choose the conjugate anion as the component, so that the acid, HCl is formed from components:

$$HCl = Cl^- + H^+$$

O

For bases, such as NaOH, we choose the conjugate cation as a component and the base, NaOH is formed from components as follows:

$$NaOH = Na^{+} + H_2O - H^{+}$$

Because aquatic chemistry almost always deals with dilute solutions, the concentration of H_2O may be considered fixed at a mole fraction of 1, or 55.4 M. Thus in the Morel and Hering system, H_2O is made an *implicit* component, i.e., its presence is assumed but not written, so that equation 6.8 becomes:

$$OH^- = -H^+$$
 6.9

The variable TOTH is the total amount of component H^+ , rather than the total of species H^+ . Every species containing the component H^+ contributes positively to TOTH while every species formed by subtracting component H^+ contributes negatively to TOTH. Because we create the species OH^- by subtracting component H^+ from component H_2O , the total of component H^+ for pure water will be:

$$TOTH = [H^+] - [OH^-]$$

Thus *TOTH* in this case is the difference between the concentrations of H^+ and OH^- . Of course, in pure water, $[H^+] = [OH^-]$, so *TOTH* = 0.

Now let's consider our example of the CaCO₃ solution. In addition to H⁺ and H₂O, we choose Ca²⁺ and CO₃²⁻ as components. In the proton mole balance equation, HCO₃⁻ counts positively (since HCO₃⁻ = CO₃²⁻ + H⁺) and CaOH⁺ (since CaOH⁺ = Ca²⁺ + H₂O - H⁺) negatively:

$$TOTH = [H^+] + [HCO_3^-] - [OH^-] - [Ca(OH)^+]$$
6.10

Comparing equations 6.10 and 6.4, we see that the *TOTH* is equal to the difference between the left and right hand sides of the proton balance equation, and that in this case TOTH = 0. This makes sense, because, having added neither [H⁺] nor [OH⁻] to the solution, the total of the component H the solution contains should be 0.

Now consider the dissolution of CO_2 in water to form carbonic acid:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3$$
 6.11

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

Under the right conditions of pH, this carbonic acid will dissociate to form bicarbonate ion:

$$H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$
 6.12

If we choose CO_2 as our component, bicarbonate ion would be made from components CO_2 , H_2O , and H^+ :

$$HCO_{3}^{-} = CO_{2} + H_{2}O - H^{+}$$

Thus in the *TOTH* proton mole balance equation, bicarbonate ion would count negatively, so *TOTH* is:

$$TOTH = [H^+] - [OH^-] - [HCO_3^-]$$
 6.13

Alternatively, had we defined CO_3^{2-} as a component, then species HCO_3^- is formed by the components: $HCO_3^- = H^+ + CO_3^{2-}$

In this case, the proton mole balance equation is:

$$TOTH = [H^+] - [OH^-] + [HCO_3^-]$$
 6.13a

Here we see that *TOT*H depends on how we define our components.

6.2.1.3 Conservation Equations

A further constraint on the composition of a system is provided by *mass balance*. Acid-base reactions will not affect the total concentration of a substance. Thus regardless of reactions 6.5 and 6.6, and any other complexation reactions, such as

$$Pb^{2+} + S^{2-} \rightleftharpoons PbS_{ac}$$

the total concentration of sulfide remains constant. Thus we may write:

$$\Sigma S = [H_2 S] + [HS^-] + [S^{2-}] + [PbS_{aq}] + \dots$$

We can write one mass balance, or *conservation*, equation for each component in solution. Of course for components, such as Na, that form only one species, Na⁺ in this case, the mass balance equation is trivial. Mass balance equations are useful for those components forming more than one species.

While the charge balance constraint is an absolute one and always holds, mass balance equations can be trickier because other processes, such a redox, precipitation, and adsorption, can affect the concentration of a species. We sometimes get around this problem by writing the mass balance equation for an element, since an elemental concentration is not changed by redox processes. We might also define our system such that it is closed to get around the other problems. Despite these restrictions, mass balance often provides a useful additional constraint on a system.

6.2.1.2 Charge Balance

As we saw in Chapter 3, solutions are electrically neutral; that is, the number of positive and negative charges must balance:

$$\sum_{i} m_i z_i = 0 \tag{6.14}$$

where *m* is the number of moles of ionic species *i* and *z* is the charge of species *i*. Equation 6.14 in known as the *charge balance equation* and is identical to equation 3.99. This equation provides an important constraint on the composition of a system. Notice that in some cases, the proton balance and charge balance equations are identical (e.g., equations 6.2 and 6.7).

For each acid-base reaction an equilibrium constant expression may be written. By manipulating these equilibrium constant expressions as well proton balance, charge balance, and mass balance equations, it is possible to predict the pH of any solution. In natural systems where there are many species present, however, solving these equations can be a complex task indeed. An important step in their solution is to decide which reactions have an insignificant effect on pH and neglect them.

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

6.2.3 THE CARDONATE SYSTEM

We now turn our attention to carbonate. Water at the surface of the Earth inevitably contains dissolved CO_2 , either as a result of equilibration with the atmosphere or because of respiration by organisms. CO_2 reacts with water to form *carbonic acid*:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3$$
 6.15

Some of the carbonic acid dissociates to form bicarbonate and hydrogen ions:

$$H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$
 6.16

Some of the bicarbonate will dissociate to an additional hydrogen ion and a carbonate ion:

$$\mathrm{HCO}_{3}^{-} \rightleftharpoons \mathrm{H}^{+} + \mathrm{CO}_{3}^{2-} \tag{6.17}$$

We can write three equilibrium constant expressions for these reactions:

$$\mathbf{K}_{sp} = \frac{a_{H_2CO_3}}{f_{CO_2}} \tag{6.18}$$

$$K_{1} = \frac{a_{H^{+}}a_{HCO_{3}^{-}}}{a_{H_{2}CO_{3}^{-}}}$$
6.19

$$K_2 = \frac{a_{H^+} a_{CO_3^{2-}}}{a_{HCO_3^{-}}}$$
 6.20

The equilibrium constants for these reactions are given in Table 6.1 as a function of temperature.

In the above series of reactions, we have simplified things somewhat and have assumed that dissolved CO_2 reacts completely with water to form H_2CO_3 . This is actually not the case, and much of the dissolved CO_2 will actually be present as distinct molecular species, $CO_{2(aq)}$. Thus reaction 6.15 actually consists of the two reactions:

$$\operatorname{CO}_{2(g)} \rightleftharpoons \operatorname{CO}_{2(aq)}$$
 6.15a

$$CO_{2(aq)} + H_2O \rightleftharpoons H_2CO_3$$
 6.15b

The equilibrium for the second reaction favors $CO_{2(aq)}$. However, it is analytically difficult to distinguish between the species $CO_{2(aq)}$ and H_2CO_3 . For this reason, $CO_{2(aq)}$ is often combined with H_2CO_3 when representing the aqueous species. The combined total concentration of $CO_{2(aq)} + H_2CO_3$ is sometimes written as $H_2CO_3^{-1}$. We will write it simply as H_2CO_3 .

The importance of the carbonate system is that by dissociating and providing hydrogen ions to solution, or associating and taking up free hydrogen ions, *it controls the pH of many natural waters*. Example

T (°C)	pK _{CO2}	pK_1	pK ₂	pK_{cal}	pK _{arag}	pK _{CaHCO3} **	pK _{CaCO3} ⁺
0	1.11	6.58	10.63	8.38	8.22	-0.82	-3.13
5	1.19	6.52	10.55	8.39	8.24	-0.90	-3.13
10	1.27	6.46	10.49	8.41	8.26	-0.97	-3.13
15	1.34	6.42	10.43	8.43	8.28	-1.02	-3.15
20	1.41	6.38	10.38	8.45	8.31	-1.07	-3.18
25	1.47	6.35	10.33	8.48	8.34	-1.11	-3.22
30	1.52	6.33	10.29	8.51	8.37	-1.14	-3.27
45	1.67	6.29	10.20	8.62	8.49	-1.19	-3.45
60	1.78	6.29	10.14	8.76	8.64	-1.23	-3.65
80	1.90	6.34	10.13	8.99	8.88	-1.28	-3.92
90	1.94	6.38	10.14	9.12	9.02	-1.31	-4.05

 Table 6.1. Equilibrium Constants for the Carbonate System

 $^{*}K_{CaHCO_{3}^{+}} = a_{CaHCO_{3}^{+}} / (a_{Ca^{2+}}a_{HCO_{3}^{-}})$

⁺ $K_{CaCO_3^0} = a_{CaCO_3^0} / (a_{Ca^2} + a_{CO_3^2})$

CHAPTER O: AQUATIC CHEMISTRY

Example 6.1. Proton, Mass, and Charge Balance Equations for Na₂CO₃ Solution

Write the proton, proton mass balance, charge balance, and carbonate conservation equations for a solution prepared by dissolving Na_2CO_3 in water. Assume that $NaCO_3$ dissociates completely and that the system is closed.

Answer: We begin with the proton balance equation. From the dissociation of water we have:

$$[H^+] = [OH^-]$$

In addition to this, hydroxide ions will also be generated by reaction between CO_3^{2-} and water:

$$CO_3^{2-} + H_2O \rightleftharpoons HCO_3^{-} + OH^{-}$$

and

$$HCO_3^- + H_2O \rightleftharpoons H_2CO_3 + OH$$

Since for each HCO_3^- formed, one OH^- must have formed and for each H_2CO_3 present two OH^- must have formed, the proton balance equation is:

$$[H^+] + [HCO_3^-] + 2[H_2CO_3] = [OH^-]$$
 6.21

Choosing CO_2 and sodium ions as components (in addition to H⁺ and H₂O), the three carbonate species are made from components as follows:

$$H_2CO_3 = H_2O + CO_2$$

 $HCO_3^- = H_2O + CO_2 - H^+$
 $CO_3^{2-} = H_2O + CO_2 - 2H^+$

In this case, the proton mole balance equation is:

$$TOTH = [H^+] - [HCO_3^-] - 2[CO_3^{2-}] - [OH^-]$$
6.22

The charge balance equation is:

$$[H^{+}] + [Na^{+}] = [OH^{-}] + [HCO_{3}^{-}] + 2[CO_{3}^{2-}]$$
6.23

The conservation equation for carbonate species is:

$$\Sigma CO_3 = [CO_3^{2-}] + [HCO_3^{-}] + [H_2CO_3]$$
 6.24

Since the dissolution of Na_2CO_3 produces two moles of Na^+ for every mole of carbonate species, we may also write:

$$[Na^{+}] = 2\Sigma CO_3 = 2([CO_3^{2-}] + [HCO_3^{-}] + [H_2CO_3])$$

6.2 shows that pure water in equilibrium with atmospheric CO_2 will be slightly acidic. The production of free H⁺ ions as a result of the solution of CO_2 and dissociation of carbonic acid plays an extremely important role is weathering.

Ground waters may not be in equilibrium with the atmosphere, but will nonetheless contain some dissolved CO_2 . Because of respiration of organisms in soil (mainly plant roots & bacteria) through which they pass before penetrating deeper, ground waters often contain much more CO_2 than water in equilibrium with the atmosphere. In addition, calcite and other carbonates are extremely common minerals in soils and in sedimentary, metamorphic, and altered igneous rocks. Ground waters will tend to approach equilibrium with calcite by either dissolving it or precipitating it:



Figure 6.1. Activities of different species in the carbonate system as a function of pH assuming $\Sigma CO_2 = 10^{-2}$. After Drever (1988).

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

Example 6.2. pH of Water in Equilibrium with the Atmosphere

What is the pH of water in equilibrium with the atmospheric CO_2 at 25°C, assuming ideal behavior and no other dissolved solids or gases present? The partial pressure of CO_2 in the atmosphere is $3.7 \times$ 10^{-4} .

Answer: In this case, the proton balance and charge balance equations are identical:

$$[H^{+}] = [OH^{-}] + [HCO_{3}^{-}] + 2[CO_{3}^{2-}]$$
6.25

We might guess that the pH of this solution will be less than 7 (i.e., $[H^+] > 10^{-7}$). Under those circumstances, the concentrations of the hydroxyl and carbonate ions will be much lower than that of the hydrogen and bicarbonate ions. Assuming we can neglect them, our equation then becomes simply:

$$[\mathrm{H}^+] \cong [\mathrm{HCO}_3^-] \tag{6.25a}$$

We can combine equations 6.18 and 6.19 to obtain an expression for bicarbonate ion in terms of the partial pressure of CO₂:

 $[\text{HCO}_{3}^{-}] = (K_{1}K_{CO_{2}}P_{CO_{2}})/[\text{H}^{+}]$

Substituting this into 6.25a and rearranging, we have:

$$\mathbf{H}^{+}]^{2} \cong \mathbf{K}_{1} \mathbf{K}_{CO_{2}} P_{CO_{2}}$$

IF Taking the negative log of this expression and again rearranging, we obtain:

$$pH \cong \frac{-\log \mathrm{K}_{1} - \log \mathrm{K}_{\mathrm{CO}_{2}} - \log P_{\mathrm{CO}_{2}}}{2}$$

Substituting values from Table 6.1, we calculate pH = 5.64. Looking at Figure 6.1, we can be assured that our assumption that carbonate and hydroxyl ion abundances are valid. Indeed, an exact solution using the Solver in ExcelTM differs from the approximate one by less than 0.0001 pH units.

$$CaCO_3 \rightleftharpoons Ca^{2+} + CO_3^{2-}$$
 6.27

Carbonate ions produced in this way will associate with hydrogen ions to form bicarbonate as in reaction 6.17 above, increasing the pH of the solution. Water containing high concentrations of calcium (and magnesium) carbonate is referred to as 'hard water'; such waters are generally somewhat alkaline.

Now suppose we have a known activity of all carbonate species in solution, say for example 10^{-2} :

$$a_{H_2CO_3} + a_{HCO_3^-} + a_{CO_3^{2-}} = \Sigma CO_2 = 10^{-2}$$
6.28

From this, and the dissociation constants, we can calculate the amount of each species present as a function of pH and temperature. For example, we can use the equilibrium constant expressions to obtain substitutions for the carbonic acid and carbonate ion activities in equation 6.28 that are functions of bicarbonate ion activity and pH. We then solve equation 6.28 to obtain an expression for the activity of the bicarbonate ion as a function of total CO₂ and hydrogen ion activity:

$$a_{HCO_{3}^{-}} = \frac{\Sigma CO_{2}}{(a_{H^{+}}/\mathrm{K}_{1}) + 1 + (\mathrm{K}_{2}/a_{H^{+}})}$$
6.29

Similar equations may be found for carbonic acid and carbonate ion. Carrying out these calculations at various pH, we can construct the graph shown in Figure 6.1. In this figure, we see that carbonic acid is the dominant species at low pH, bicarbonate at intermediate pH, and carbonate at high pH.

6.2.3.1 Equivalence Points

Particularly simple relationships occur when the activities of two species are equal. For example, choosing pH to be 6.35 (at 25°C) we can rearrange equation 6.19 and substitute to obtain:

$$\frac{a_{HCO_3^-}}{a_{H_2CO_3}} = \frac{K_1}{a_{H^+}} = \frac{10^{-6.35}}{10^{-6.35}}$$
6.30

6.26

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

The activity of the carbonate ion will be very low at this pH. If we ignore it, equ. 6.28 becomes:

$$a_{H_2CO_3} + a_{HCO_3^-} \cong 10^{-1}$$

According to equ. 6.30, the activities of both species are equal, so each must be 5×10^{-3} .

Now consider the point where the hydrogen ion concentration equals the bicarbonate ion concentration ($[H^+] = [HCO_3^-]$). At this point, the concentration of the carbonate ion is extremely low, and there is exactly enough H⁺ to convert all HCO₃⁻ to H₂CO₃. From the perspective of the proton balance then, the HCO₃⁻ concentration is equivalent to the same concentration of H₂CO₃. This point, labeled EP on Figure 6.1, is called the *CO*₂ *equivalence point*. In a similar way, the point where the carbonic acid and carbonate ion concentrations are equal is called the *bicarbonate equivalence point*, and that where bicarbonate and hydroxyl concentrations are equal is called the *carbonate equivalence point*. The pH of these equivalence points depends, among other things, on the Σ CO₂ concentration.

The exact concentrations of carbonate species depends on total carbonate concentration as well as the concentration of other ions in solution. Thus the distribution shown in Figure 6.1 is unique to the conditions specified ($\Sigma CO_2 = 10^{-2}$, no other ions present). Nevertheless, the distribution will be qualitatively similar for other conditions.

6.2.4 Conservative and Non-Conservative lons

We can divide dissolved ions into *conservative* and *non-conservative* ones. The conservative ions are those whose concentrations are not affected by changes in pH, temperature, and pressure, *assuming no precipitation or dissolution*. In natural waters, the principal conservative ions are Na⁺, K⁺, Ca²⁺, Mg²⁺, Cl⁻, $SO_4^{2^-}$ and NO_3^{-} . These ions are conservative because they are fully dissociated from their conjugate acids and bases over the normal range of pH of natural waters. Non-conservative ions are those that will undergo association and dissociation reactions in this pH range. These include the proton, hydroxyl, and carbonate species as well as B(OH)⁺₄, H₃SiO⁺₄, HS⁻, NH₄OH, phosphate species, and many organic anions. Virtually all the non-conservative species are anions, the two principle exceptions being H⁺ and NH₄OH (which dissociates to form NH⁺₄ at low pH). Variations in the concentrations of non-conservative ions result from reactions between them, and these reactions can occur in the absence of precipitation or dissolution. For example, reaction of the carbonate and hydrogen ion to form bicarbonate will affect the concentrations of all three ions. Of course, if the system is at equilibrium, this reaction will not occur in the absence of an external disturbance, such as a change in temperature.

Example 6.3. pH of a Solution with Fixed Total Carbonate Concentration

A groundwater moving through soil into a deep aquifer acquires a total dissolved CO_2 concentration of 10^{-2} M. Assuming the water does not exchange with surrounding rock, ideal behavior and no other dissolved solids or gases, what is the pH of the water?

Answer: In this case, our charge and proton balance equations are the same as in Example 6.2, i.e., equation 6.25. Since the solution does not exchange with surrounding rock, it can be considered a closed system and we can write the following mass balance equation:

$$\Sigma CO_2 = [H_2 CO_3] + [HCO_3^-] + [CO_3^{2-}] = 10^{-2}$$
6.31

Simultaneously solving the change balance and mass balance equations, and using equilibrium constant expressions to eliminate carbonate and OH species, we obtain:

 $[H^+]^4 + K_1[H^+]^3 + \{K_2K_1 - K_W - K_1\Sigma CO_2\}[H^{2+}] - \{K_W + 2K_2\Sigma CO_2\}K_1[H^+] - K_2K_1K_W = 0$ We might again guess that the concentration of the carbonate ion will be very low, and that we can therefore neglect all terms in which K_2 occurs. We might also guess that pH will be acidic so that $[H^+] \rightarrow [OH^-]$, and therefore that we can neglect terms containing K_W . Our equation becomes:

$$K_1^{-1}[H^+]^2 + [H^+] = \Sigma CO_2$$

Solving this quadratic, we find that pH = 4.18.

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

6.2.5 Total Alkalinity and Carbonate Alkalinity

Alkalinity is a measure of acid-neutralizing capacity of a solution and is defined as the sum of the concentration (in equivalents) of bases that are titratable with strong acid. Mathematically, we define alkalinity as the negative of TOTH when the components are the principal species of the solution at the CO₂ equivalence *point*. The *acidity* can be defined as the negative of alkalinity, and hence equal to *TOT*H.

As a first example, let's consider a solution containing a fixed total dissolved concentration of CaCO₃. At the CO_2 equivalence point, H_2CO_3 is the principal carbonate species, so we choose our components as H^+ , H_2O , CO_2 , and Ca^{2+} (Since we always choose water as a component, we do not want to choose H_2CO_3 as a component, because it contains the component H_2O and hence is not fully independent. Instead, we choose CO₂ as the carbonate component in this case). Species H₂CO₃, HCO² and CO² are made by combining these components as follows:

$$\begin{split} H_2 CO_3 &= H_2 O + CO_2 \\ HCO_3^- &= H_2 O + CO_2 - H^+ \\ CO_3^{2-} &= H_2 O + CO_2 - 2H^+ \end{split}$$

The proton mole balance equation is then:

$$TOTH = [H^+] - [HCO_3^-] - 2[CO_3^{2-}] - [OH^-]$$
6.32

The alkalinity is then:

 H_2CO_{3} well, of

H₄SiO₄,

$$Alk = -TOTH = -[H^+] + [HCO_3^-] + 2[CO_3^{2-}] + [OH^-]$$
6.33

Example 6.4. The Tableau Method of Morel and Hering
Write an expression for the alkalinity of a solution containing
$$H_3SiO_4^-$$
, H_4SiO_4 , $B(OH)_3$, $B(OH)_4^-$,
 H_2S , HS^- , $H_2PO_4^-$, HPO_4^{2-} ,
 H_2CO_3 , HCO_3^- , and CO_3^{2-} , as
well, of course, as OH^- and H^+ .
Answer: The alkalinity will
be the negative of *TOTH* when
the components are the principal
species of the solution at the CO_2
equivalence point, so the real
problem is just choosing com-
ponents and defining our spe-
cies in terms of these. At the
principal species will be
 H_4SiO_4 , $B(OH)_3$, H_2S , $H_2PO_4^-$, H_4SiO_4
 H_2CO_3 1 1 1
 HCO_3^- -1 1 1
 $H_2CO_3^-$ -2 1 1
 HCO_4^{2-} -1 1
 $H_2PO_4^-$ 1
 HO_4^{2-} -1 1
 $H_2PO_4^-$ 1
 HO_4^{2-} -1 1
 $H_2PO_4^-$ 1
 $H_3SiO_4^-$ -1 1
 $H_3SiO_4^-$ -1 1
 $H_5SO_4^-$ -1 1
 H_5S^- -1
 HS^- -1
 HES^- -1
 HS^- -1
 HES^- -1

leau, a table with *the components listed across the top and the species listed vertically*. Entries in the table are just the stoichiometric coefficients used to define each species in terms of its components. In this case, the tableau will look like that shown above. The first column of tableau shows us what the coefficients will be in our *TOT*H equation. Our expression for alkalinity will thus be:

$$A \ l \ k = -TOTH = -\{[H^+] - [OH^-] - [HCO_3^-] - 2[CO_3^{2-}] - [HPO_4^{2-}] - [H_3SiO_4^-] - [B(OH)_4^-] - [HS^-]\}$$

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

This sum, $-[H^+] + [HCO_3^-] + 2[CO_3^2^-] - [OH^-]$, is called the *carbonate alkalinity*. In this example, carbonate alkalinity and alkalinity are equal since there are no other ions in solution. To avoid confusion with carbonate alkalinity, alkalinity is sometimes called *total alkalinity*.

An analytical definition of alkalinity is that it is the quantity of acid that must be added to the solution to bring the pH to the CO_2 equivalence point.

We can also express alkalinity in terms of conservative and non-conservative ions. The charge balance equation, equation 6.14, could be written as:

$$\Sigma$$
 cations (in equivalents) – Σ anions (in equivalents) = 0 6.35

This can the be expanded to:

 Σ conserv. cations - Σ conserv. anions + Σ non-conserv. cations - Σ non-conserv. anions = 0 (all in units of equivalents)*. Rearranging, we have:

 Σ conserv. cations - Σ conserv. anions = - Σ non-conserv. cations + Σ non-conserv. anions 6.36

The right hand side of equation 6.36 is equal to the *alkalinity*. Hence we may write:

 $Alk=\Sigma$ conserv. cations - Σ conserv. anions = - Σ non-conserv. cations + Σ non-conserv. anions 6.37

This equation emphasizes an important point. The difference of the sum of conservative anions and cations is clearly a conservative property, i.e., they cannot be changed except by the addition or removal of components. Since alkalinity is equal to this difference, alkalinity is also a conservative quantity (i.e., independent of pH, pressure and temperature). *Thus total alkalinity is conservative, even though concentrations of individual species are not.*

6.2.5.1 Alkalinity Determination and Titration Curves

If the concentrations of all major conservative ions in a solution are known, the alkalinity can be simply calculated from equation 6.35. It is often useful, however, to determine this independently. This is done, as the definition of alkalinity suggests, through titration. Titration is the process of progressively adding a strong acid or base to a solution until a specified pH, known as an end-point, is reached. In the case of the determination of alkalinity, this end point is the CO_2 equivalence point.

Consider a solution containing a certain concentration of sodium bicarbonate (Na_2CO_3). Because the carbonate ion can act as a proton acceptor, $NaCO_3$ is a base. We can determine both the alkalinity and the total carbonate concentration of this solution by titrating with a strong acid, such as HCl. Let's examine the chemistry behind this procedure.

For clarity, we make several simplifying assumptions. First, we assume ideal behavior. Second, we assume the system is closed, so that all components are conserved, except for $[H^+]$ and $[Cl^-]$, which we progressively add. Third, we assume that the volume of our Na₂CO₃ solution is sufficiently large and our HCl sufficiently concentrated that there is no significant dilution of the original solution. Finally, we assume both Na₂CO₃ and HCl dissociate completely.

The charge balance equation during the titration is:

$$[Na^{+}] + [H^{+}] = [Cl^{-}] + [HCO_{3}^{-}] + 2[CO_{3}^{2-}] + [OH^{-}]$$
6.38

Since the Cl⁻ concentration is conservative, it will be equal to the total amount of HCl added. Into equation 6.38, we can substitute the following:

$$[HCO_{3}^{-}] = \frac{K_{1}[H_{2}CO_{3}]}{[H^{+}]} \quad 6.39a \quad [CO_{3}^{2-}] = \frac{K_{1}K_{2}[H_{2}CO_{3}]}{[H^{+}]^{2}} \quad 6.39b \quad and \quad [OH^{-}] = \frac{K_{w}}{[H^{+}]} \quad 6.39c$$

^{*} One *equivalent* of a species is defined as the number of moles multiplied by the charge of the species. Thus one equivalent of CO_3^2 is equal to 0.5 moles of CO_3^2 , but one equivalent of CI^- is equal to 1 mole of CI^- . For an acid or base, an equivalent is the number moles of the substance divided by the number of hydrogen or hydroxide ions that can be potentially produced by dissociation of the substance. Thus there are 2 equivalents per mole of H_2CO_3 , but 1 equivalent per mole of Na(OH).

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

Doing so and rearranging yields:

$$[Cl^{-}] = [Na^{+}] + [H^{+}] - \frac{K_{1}[H_{2}CO_{3}]}{[H^{+}]} - \frac{K_{1}K_{2}[H_{2}CO_{3}]}{[H^{+}]^{2}} - \frac{K_{W}}{[H^{+}]}$$
6.40

We may also write a conservation equation for carbonate species, which is the same as equation 6.24 in Example 6.1. Substituting equations 6.39a and 6.39b into 6.24 and rearranging, we have:

$$[H_{2}CO_{3}] = \frac{\Sigma CO_{2}}{1 + \frac{K_{1}}{[H^{+}]} + \frac{K_{1}K_{2}}{[H^{+}]^{2}}}$$
6.41

Substituting this expression into 6.40, we obtain:

$$[Cl^{-}] = [Na^{+}] + [H^{+}] - \frac{\Sigma CO_{2}}{[H^{+}] + K_{1} + K_{1}K_{2}/[H^{+}]} \left\{ K_{1} - \frac{K_{1}K_{2}}{[H^{+}]} \right\} - \frac{K_{w}}{[H^{+}]}$$
6.42

From stoichiometry, we also know that $\Sigma CO_2 = 2[Na^+]$. From this equation we can construct a plot showing how many moles of HCl we must add to achieve a certain pH. We can also use equation 6.39 and similar ones expressing the bicarbonate and carbonate ions as functions of pH to plot the change in the carbonate speciation during the titration. Figure 6.2 shows such a plot for a 0.005 M Na₂CO₃ solution. There are two regions where pH changes rapidly with small additions of HCl. These are the two end-points of the titration. Comparing the titration curve with the speciation curves, we see that the two end-points correspond to the CO₂ and bicarbonate equivalence points.

An analytical definition of alkalinity is its *acid neutralizing capacity when the end-point of the titration is the* CO_2 *equivalence point* (Morel and Hering, 1993). We had previously defined alkalinity as the negative of *TOTH* when the principal components are those at the CO_2 equivalence point. Let's now show that these definitions are equivalent.

Our *TOT*H expression, written in terms of components at the CO_2 equivalence point, is identical to 6.22:

$$TOTH = [H^+] - [HCO_3^-] - 2[CO_3^{2-}] - [OH^-]$$
(6.32)

and the charge balance equation (before any HCl is added) is:

$$[Na^{+}] + [H^{+}] = [HCO_{3}^{-}] + 2[CO_{3}^{2-}] + [OH^{-}]$$
(6.14)

Combining the two we have: $TOTH = -[Na^+]$

Since the alkalinity is the negative of *TOTH*, it follows that (before the addition of HCl):

$$Alk = [Na^+] \tag{6.43}$$

We obtain exactly the same result from equation 6.37. It is easy to show that after titrating to the CO_2 equivalence point, the alkalinity is 0. The change in alkalinity is thus equal to the number of equiva-

Example 6.5. Calculating Alkalinity of Spring Wate	R			
Calculate the alkalinity of spring water from Thonon, France,	Anions	mМ	Cations	mМ
whose analysis is given at right (this is the same analysis as	HCO_{2}^{-}	5.436	Ca ²⁺	2.475
in Problem 3.9).	SO_4^{2-3}	0.146	Mg^{2+}	0.663
Answer: We can use equation 6.37 to calculate alkalinity.	NO_{3}^{-}	0.226	K^+	0.036
All the ions listed here are conservative with the exception	Cl	0.231	$N a^+$	0.223
of HCO_3 . To calculate alkalimity, we first need to convert the malar concentrations to equivalently use do so by multiplying	the come	atuation	at an ala an	
the motar concentrations to equivalents; we do so by multiplying	, the conce		f each sp	ecies by
its charge. We find the sum of conservative anion concentration	is to be 0.74	49 meq (r	nilliequiv	'alents),
and that of the conservative cation concentrations to be 6.535 m	eq. The all	kalinity	is the dif	ference,
5.786 meg.				

CHAPTER O: AQUATIC CHEMISTRY

lents, or moles, of H^+ we have added to the solution. Since at the end point, $[H^+] = [HCO_3^-]$ and the concentrations of CO_3^{2-} and $OH^$ are negligible, our charge balance equation, 6.38, reduces to:

$$[Na^+] = [Cl^-]$$

Comparing this with 6.41, we see that the alkalinity is equal to the amount of HCl added. In the example in Figure 6.2, the equivalence point occurs after the addition of 10 ml of 1M HCl, or a total of 0.01 moles of Cl-. (Notice that since small additions of acid result in large changes in pH at the end points, we do not have to determine pH particularly accurately during the titration for an accurate determination of alkalinity.) So the alkalinity is 0.01 equivalents. This is exactly the answer we obtain from 6.41 for 1 liter of 0.005 M



Figure 6.2. Titration curve (solid red line) for a one liter 0.005 M Na_2CO_3 solution titrated with 1M HCl. Left axis shows the number of ml of HCl to be added to obtain a given pH. Also shown are the concentrations of carbonate species, H+, and OH- (dashed black lines, right axis gives scale). EP1 is the bicarbonate equivalence point, EP2 is the CO₂ equivalence point.

 Na_2CO_3 since there are 2 moles of Na^+ for each mole of Na_2CO_3 .

By assuming that the concentration of H^+ contributes negligibly to charge balance, it is also easily shown (Problem 6.2) that at the bicarbonate equivalence point:

$$\Sigma CO_2 = [Cl^-] + [OH^+]$$
 6.44

Thus total carbonate is obtained by titrating to the bicarbonate equivalence point (knowing the pH of the end-point allows us to determine the ΣCO_2 exactly; however neglecting the [OH⁻] term in 6.42 results in only a 1% error in the example shown). In Figure 6.2, this occurs after the addition of 5 ml 1M HCl.

6.2.6 Buffer Intensity

The carbonate system is a good example of a pH *buffer*. We define the *buffer intensity* of a solution as the inverse of change in pH per amount of strong base (or acid) added:

$$\beta \equiv \frac{dC_B}{dpH} = -\frac{dC_A}{dpH} \qquad 6.45$$

where C_B and C_A are the concentrations, in equivalents, of strong base or acid respectively. The greater the buffer capacity of a solution, the less change in its pH as an acid or base is added. The buffer capacity of a solution can be calculated by differentiation of the



Figure 6.3. Buffer intensity as a function of pH for several ideal natural systems: β_{C_T} fixed total dissolved CO₂, $\beta_{P_{CO_2}}$ water in equilibrium with atmospheric CO₂, $\beta_{CaCO_{3(s)}}$ water in equilibrium with calcite, and $\beta_{An-Kaol}$ water in equilibrium with anorthite and kaolinite. After Stumm and Morgan (1996).

CHAPTER O: AQUATIC CHEMISTRY

Example 6.6. Calculating Buffer Intensity

How will pH change for given addition of a strong base such as NaOH for a solution of pure water in equilibrium with atmospheric CO_2 ? Calculate the buffer intensity of this solution as a function of pH. Assume that NaOH completely dissociates and behavior is ideal.

Answer: We want to begin with the charge balance equation in this case because it relates the two quantities of interest in this case, $[Na^+]$ and $[H^+]$. The charge balance equation is the same as in Example 6.1:

$$[Na^{+}]+[H^{+}]=[OH^{-}]+[HCO_{3}^{-}]+2[CO_{3}^{2-}]$$
(6.23)

Since Na⁺ is a conservative ion, its concentration will depend only on the amount of NaOH added, so that $C_B = [Na^+]$. Substituting this into equation 6.14 and rearranging, we have:

$$C_B == [OH^{-}] + [HCO_3^{-}] + 2[CO_3^{2-}] - [H^{+}]$$
6.46

We can now use the equilibrium constant relations to substitute for the first three terms of the right hand side of 6.46 and obtain:

$$C_{B} = \frac{K_{W} + K_{1}K_{sp}P_{CO_{2}}}{[H^{+}]} + 2\frac{K_{2}K_{1}K_{sp}P_{CO_{2}}}{[H^{+}]^{2}} - [H^{+}]$$

Using the relation $pH = -\log [H^+]$ to replace $[H^+]$ in this equation with pH, we have:

$$C_{B} = \frac{K_{W} + K_{1}K_{sp}P_{CO_{2}}}{10^{-pH}} + 2\frac{K_{2}K_{1}K_{sp}P_{CO_{2}}}{10^{-2pH}} - 10^{-pH}$$

Now differentiating with respect to pH, we obtain:

$$\frac{dC_B}{dpH} = \beta = \ln 10\{(K_w + K_1 K_{sp} P_{CO_2}) 10^{pH} + 4K_2 K_1 K_{sp} P_{CO_2} 10^{2pH} + 10^{-pH}\}$$



Figure 6.4 shows a plot of this equation using the values $atmospheric CO_2$. in Table 6.1. Buffer intensity is negligible in neutral to slightly acidic conditions, but increases rapidly with pH.

equation relating base (or acid) concentration to pH, as is illustrated in Example 6.5.

A pH buffer acts to control pH within a narrow range as H^+ ions are added or removed from solution by other reactions. To understand how this works, imagine a solution containing carbonic acid, $CO_3^{2^-}$, HCO₃, and H⁺ ions in equilibrium concentrations. Now imagine that additional H⁺ ions are added (for example, by addition of rain water containing HNO₃). In order for the right hand side of equation 6.19 to remain equal to K₁ despite an increase in the activity of H⁺ (which it must at constant temperature and pressure), the bicarbonate activity must decrease and the carbonic acid activity increase. It is apparent then that reaction 6.16 must be driven to the left, taking free hydrogen ions from solution, hence driving the pH back toward its original value. Similarly, reaction 6.20, the dissociation of bicarbonate, will also be driven to the left, increasing the bicarbonate concentration and decreasing the hydrogen and carbonate ion concentrations.

The buffer capacity of the carbonate system depends strongly on pH and also on the concentration of the carbonate species and the concentration of other ions in solution. In pure water containing no other ions and only carbonate in amounts in equilibrium with the atmosphere, the buffering capacity is negligible near neutral pH, as is shown in Figure 6.4. Natural solutions, however, can have substantial buffering capacity. Figure 6.3 illustrates three other examples of natural pH buffers. "Hard water" is an example of water with a substantial buffering capacity due to the presence of dissolved carbonates. As

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

we shall see, how adversely lakes and streams are impacted by "acid rain" depends on their buffering intensity.

6.3 Complexation

Ions in solution often associate with other ions, forming new species called complexes. Complex formation is important because it affects the solubility and reactivity of ions, as we will see in the following section. In some cases, complex formation is an intermediate step in the precipitation process. In other cases, ions form stable, soluble complexes that greatly enhance the solubility of the one or both of the ions.

Complexation is usually described in terms of a central ion, generally a metal, and ions or molecules that surround, or *coordinate*, it, referred to as *ligands*. Perhaps the simplest and most common complexes are those formed between metals and water or its dissociation products. We learned in Section 3.7 that a solvation shell surrounds ions in aqueous solutions. The solvation shell (Figure 3.10) consists of water molecules, typically 6, though fewer in some cases, loosely bound to the ion through electrostatic forces. This solvation shell is referred to as an *aquo complex*. Water molecules are the ligands in aquo-complexes. Aquo-complexes are ubiquitous: all charged species have a solvation shell. Truly "free ions" do not exist: ions not otherwise complexed ("free ions") are in reality associated with surrounding water molecules and hence actually aquo-complexes. However, the existence of this type of complex is implicitly accounted for through the activity coefficient and not usually explicitly considered. Nevertheless, it is important to bear in mind that since all ions are complexed in some way to begin with, every complexation reaction in aqueous solution is essentially a ligand exchange reaction.

Beyond aquo complexes, we can distinguish two types of complexes:

• *Ion pairs*, where ions of opposite charge associate with one and other through electrostatic attraction, yet each ion retains part or all if its solvation sphere. Figure 6.5 illustrates two possibilities: one where the two solvation spheres are merely in contact, the other where the water molecules are shared between the two solvation spheres. Ion pairs are also called *outer sphere complexes*.

• *Complexes (senso stricto)*, where the two ions are in contact and a bond forms between them that is at least partly covalent in nature (Figure 6.5c). These are often called *inner sphere complexes*.

6.3.1 STADILITY CONSTANTS

In its simplest form, the reaction for the formation of an ion pair or complex between a metal cation M and an anion or ligand L may be written as:

$$mM^+ + \ell L^- \rightleftharpoons M_m L_\ell$$

As with any other reaction, we may define an equilibrium constant as:

$$\mathbf{K} = \frac{a_{M_m L_\ell}}{a_M^m a_L^\ell} \qquad \qquad 6.48$$

For example, the equilibrium constant for the reaction:

$$\operatorname{Zn}(\operatorname{H}_2\operatorname{O})_6^{2+} + \operatorname{OH}^- \rightleftharpoons \operatorname{Zn}(\operatorname{H}_2\operatorname{O})_5\operatorname{OH}^+ + \operatorname{H}_2\operatorname{O}$$

Ion Pairs

a. Solvation Shell Contact

b. Shared Solvation Shell





c. Ion Contact

Figure 6.5. Illustration of ion pair and complex formation. Two types of ion pairs can be envisioned: (a)solvation shell contact and (b) solvation shell sharing. Ion pairs are sometimes referred to as outer sphere complexes. In formation of true complexes, ions are in contact (c) and there is some degree of covalent bonding between them. These are some times referred to as inner sphere complexes.

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

is:

$$\mathbf{K}_{1} = \frac{a_{Zn(H_{2}O)_{5}(OH)^{+}}}{a_{Zn(H_{2}O)_{6}^{2+}}a_{OH^{-}}}$$

. .

(we omit the activity of water because we assume, as usual, that it is 1). As we noted, however, the aquo-complex is generally not explicitly expressed, so this same reaction would more often be written as:

$$Zn^{2+} + OH^{-} \rightleftharpoons ZnOH^{+}$$

and the equilibrium constant as:

$$\beta_1 = \mathbf{K}_1 = \frac{a_{Zn(OH)^+}}{a_{Zn^{2+}}a_{(OH)^-}}$$

Equilibrium constants for complex formation reactions are often referred to as *stability constants*, since their value is an indication of the stability of the complex, and often denoted by β . Thus for the reaction above, β_1 and K_1 are synonymous. *By convention, stability constants are written so as the complex appears in the numerator (i.e., as a product of the reaction)*.

The zinc ion might associate with a second hydroxyl:

$$ZnOH^- + OH^- \rightleftharpoons Zn(OH)_2$$

The equilibrium constant is:

$$\mathbf{K}_{2} = \frac{a_{Zn(OH)_{2}}}{a_{Zn(OH)^{+}}a_{(OH)^{-}}}$$

Here, however, the notation for the stability constant and the equilibrium constant differs. Whereas K_2 refers to the reaction above, β_2 refers to the reaction:

$$Zn^{2+} + 2OH^{-} \rightleftharpoons Zn(OH)_{2}$$
$$\beta_{2} = \frac{a_{Zn(OH)_{2}}}{a_{Zn^{2+}}a_{OH^{-}}^{2}} = K_{1}K_{2}$$

Hence:

Finally, the notation *K and * β are sometimes used for reactions in which the complexation reactions are written so as the hydrogen ion occurs as a product, for example:

and:

 ${}^{*}\mathbf{K}_{1} = \frac{a_{Zn(OH)^{+}}a_{H^{+}}}{a_{Zn^{2+}}}$

 $Zn^{2+} + H_2O \rightleftharpoons ZnOH^+ + H^+$

K₁ is then related to K₁ as: ${}^{}K_{1} = K_{1}K_{w} = {}^{*}\beta_{1}$ where K_w is the water dissociation constant (10⁻¹⁴).

We can define *apparent* equilibrium and stability constants, where the molar concentrations are used in place of activity. Indeed, as in other aspects of geochemistry, apparent equilibrium constants are more commonly encountered than true ones.

The equilibrium constant may in turn be related to the Gibbs Free Energy of the reaction, as in equation 3.86. Interestingly, the free energy changes involved in complexation reactions result largely from entropy changes. Indeed, the enthalpy changes of many complexation reactions are unfavorable, and the reaction proceeds only because of large positive entropy



Figure 6.6. Predominant aquo, hydroxo, and oxo complexes as a function of pH and valence state. After Stumm and Morgan (1996).

October 27, 2005

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

changes. These entropy changes result from the displacement of water molecules from the solvation shell.

The link between the equilibrium constant and the free energy change is particularly important and useful in complexation reactions because it is in most instances difficult to determine the concentrations of individual complexes analytically. Thus our knowledge of chemical speciation in natural waters derived largely from predictions based on equilibrium thermodynamics.

6.3.2 WATER-RELATED COMPLEXES

Let's further consider the types of complexes typically found in aqueous solution. Ferric iron, for example, can form a $Fe(H_2O)_6^{3+}$ complex. The positive charge of the central ion tends to repel hydrogens in the water molecules, so that water molecules in these aquo-complexes are more readily hydrolyzed than otherwise. Thus these aquo-complexes can act as weak acids. For example:

$$\operatorname{Fe}(\operatorname{H}_{2}\operatorname{O})_{6}^{3+} \rightleftharpoons \operatorname{Fe}(\operatorname{H}_{2}\operatorname{O})_{5}(\operatorname{OH})^{2+} + \operatorname{H}^{+} \rightleftharpoons \operatorname{Fe}(\operatorname{H}_{2}\operatorname{O})_{4}(\operatorname{OH})_{2}^{+} + 2\operatorname{H}^{+} \rightleftharpoons \operatorname{Fe}(\operatorname{H}_{2}\operatorname{O})_{3}(\operatorname{OH})_{3}^{0} + 3\operatorname{H}^{+} \rightleftharpoons \operatorname{Fe}(\operatorname{H}_{2}\operatorname{O})_{2}(\operatorname{OH})_{4}^{-} + 4\operatorname{H}^{+} \qquad 6.49$$

As reaction 6.49 suggests, equilibrium between these hydroxo -complexes depends strongly on pH.

The repulsion between the central metal ion and protons in water molecules of the solvation shell will increase with decreasing diameter of the central ion (decreasing distance between the protons and the metal) and with increasing charge of the central ion. For highly charged species, the repulsion of the central ion is sufficiently strong that all hydrogens are repelled and it is surrounded only by oxygens. Such complexes, for example, MnO_4^- and $CrO_4^{2^-}$, are known as *oxo-complexes*. Intermediate types in which the central ion is surrounded, or *coordinated*, by both oxygens and hydroxyls are also possible, for example $MnO_3(OH)$ and $CrO_3(OH)^-$, and are known as hydroxo-oxo complexes. Figure 6.6 summarizes the predominance of aquo, hydroxo, hydro-oxo, and oxo complexes as a function of pH and valence state. For most natural waters, metals in valence states I and II will be present as "free ions", i.e., aquo complexes, valence III metals will be present as aquo and hydroxo complexes, those with higher charge will present as oxo-complexes.

Polynuclear hydroxo- and oxo-complexes, containing two or more metal ions, are also possible, for example:

Example 6.7. Complexation of Pb

Assuming a Pb concentration of 10⁻⁸ M and an equilibrium constant for the reaction:

$$Pb^{2+} + H_2O \rightleftharpoons PbOH^+ + H^+$$

of $10^{-7.7}$, calculate the fraction of Pb that will be present as PbOH⁺ from pH 6 to 9.

Answer: The equilibrium constant expression is:

$$K = \frac{[PbOH^+][H^+]}{[Pb^{2+}]}$$

In addition to the equilibrium constant expression, we also need the conservation equation for Pb:

$$\Sigma Pb = Pb^{2+} + PbOH^{+}$$

Solving the c onservation equation for Pb^{2+} and substituting into the equilibrium constant expression, we obtain:

$$(\Sigma Pb - [PbOH^+])K = [PbOH^+][H]$$

With some rearranging, we eve ntually obtain the following expression:

$$[PbOH^+]/\Sigma Pb = \frac{K}{K + [H^+]}$$

The result is illustrated in Figure 6.7. Below pH 6, virtually all Pb is present as Pb^{2+} , above pH 9, virtually all Pb is present as $PbOH^+$.



Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

$$Mn^{2+} - OH - Mn^{2+}$$
 Fe^{3+} OH^{-1}

As one might expect, the extent to which such *polymeric* species form depends on the metal ion concentration: they become increasingly common as concentration increases. Most highly-charged metal ions (3+ and higher oxidation states) are highly insoluble in aqueous solution. This is due in part to the readiness with which they form hydroxo-complexes, which can in turn be related to the dissociation of surrounding water molecules as a result of their high charge. When such ions are present at high concentration, formation of polymeric species such as those above quickly follows formation of the hydroxo complex. At sufficient concentration, formation of these polymeric species leads to the formation of colloids and ultimately to precipitation. In this sense, these polymeric species can be viewed as intermediate products of precipitation reactions.



Figure 6.8. pH and $-\log \alpha$, as a function of total copper concentration in aqueous solution. α is the fraction of copper present as the hydroxo-complex.

Interestingly enough, however, the tendency of metal ions to hydrolyze decreases with concentration. The reason for this is the effect of the dissociation reaction on pH. For example, increasing the concentration of dissolved copper decreases the pH, which in turn tends to drive the hydrolysis reaction to the left. To understand this, consider the following reaction:

for which the apparent equilibrium constant is
$$K^{app} = 10^{-8}$$
. We can express the fraction of copper pre-

$$\alpha = \frac{[\text{CuOH}^+]}{Cu_T} = \frac{\text{K}}{[\text{H}^+] + \text{K}}$$
 6.50

where Cu_T is the total dissolved copper. At constant pH, the fraction of Cu complexed is constant. However, for a solution with a fixed amount of Cu ion dissolved, we can also write a proton balance equation:

$$[H^+] = [CuOH^+] + [OH^-]$$

and a mass balance equation. Combining these with the equilibrium constant expression, we can calculate both α and pH as a function of Cu_T (Problem 6.11). When we do this, we see that as Cu_T increases, both pH and α decrease, as is demonstrated in Figure 6.8.

6.3.3 Other Complexes

sent as CuOH⁺, α_{CuOH^+} as:

When non-metals are present in solution, as they would inevitably be in natural waters, then other complexes are possible. In this respect, we can divide the elements into four classes (Table 6.2, Figure 6.9). The first is the non-metals, which form anions or anion groups. The second group is the "A-type" or "hard" metals. These metals, listed in Table 6.2, have spherically symmetric, inert-gas type outer electron configurations. Their electron shells are not readily deformed by electric fields and can be viewed as "hard spheres". Metals in this group preferentially forms complexes with fluorine and ligands having oxygen as the donor atoms (e.g., OH⁻, CO₃³⁻, PO₄³⁻, SO₄²⁻). Stability of the complexes

Geochemistry

W. M. White

CHAPTER O: AQUATIC CHEMISTRY

Н																	Не
Lí	Ве											В	С	N	0	F	Ne
Na	Мg											Al	Sí	Р	S	Cl	Ar
K	Са	Sc	Tí	ν	Cr	Mn	Fe	Со	Ní	Си	Źn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Υ	Zr	Nb	Мо	Τc	Ru	Rh	Pd	Ag	Cđ	In	sn	Sb	Те	1	Xe
Cs	Ва	La	₽f	Та	W	Re	Os	Ir	Pt	Au	Hg	τl	Pb	Bi	Ро	At	Rd
Fr	Ra	Ac	ι N														
			La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Тв	Dy	Но	Er	Tm	Yb	Lu
		١	Ac	Th	Pa	U	Np	Pu									
A-Metals Transition B Metals Ligand Formers																	

Figure 6.9. Classification of the elements with respect to complex formation in aqueous solution.

formed by these metals increases with charge to radius ratio. Thus the alkalis form only weak, unstable complexes, while elements such as Zr^{4+} form very strong, stable complexes (e.g., with fluorine). In complexes formed by A-type metals, anions and cations are bound primarily by electrostatic forces, i.e., ionic-type bonds. The A-type elements correspond approximately to the lithophile elements of Goldschmidt's classification presented in Chapter 7.

The third group is the B-type, or "soft", metal ions. Their electron sheaths are not spherically symmetric and are readily deformed by the electrical fields of other ions (hence the term soft). They preferentially form complexes with bases having S, I, Br, Cl, or N (such as ammonia; not nitrate) as the donor atom. Bonding between the metal and ligand(s) is primarily covalent and is comparatively strong. Thus Pb form strong complexes with Cl⁻ and S²⁻. Many of the complexes formed by these elements are quite insoluble. The B-type elements consist primarily of the "chalcophile elements", a term we will define in the next chapter.

The first series transition metals form the fourth group, and correspond largely to the siderophile elements (see Chapter 7). Their electron sheaths are not spherically symmetric, but they are not so readily polarizable as the B-type metals. On the whole, however, their complex-forming behavior is similar to that of the B-type metals.

Among the transition metals, the sequence of complex stability is $Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+}$ <Cu²⁺>Zn²⁺, a sequence known as the *Irving*-Williams Series. This is illustrated in Figure 6.10. In that figure, all the sulfate complexes have approximately the same stability, a reflection of the predominantly electrostatic bonding between sul-From Stumm and Morgan (1996). fate and metal. Pronounced differences are ob-

Table 6.2. Classification of Metal Ions

A-Type Metals
Li ⁺ , Na ⁺ , K ⁺ , Rb ⁺ , Cs ⁺ , Be ²⁺ , Mg ²⁺ , Ca ²⁺ , Sr ²⁺ ,
Ba^{2+} , Al^{3+} , Sc^{3+} , Y^{3+} , REE, Ti^{4+} , Si^{4+} , Zr^{4+} ,
$Hf^{4+}, Th^{4+}, Nb^{5+}, Ta^{5+}, U^{6+}$
B-Type Metals
Cu ²⁺ , Ag ⁺ , Au ⁺ , Tl ⁺ , Ga ⁺ , Zn ²⁺ , Cd ²⁺ ,
Hg ²⁺ , Pb ²⁺ , Sn ²⁺ , Tl ³⁺ , Au ³⁺ , In ³⁺ , Bi ³⁺
Transition Metal Ions
V ²⁺ , Cr ²⁺ , Mn ²⁺ , Fe ²⁺ , Co ²⁺ , Ni ²⁺ ,
Cu ²⁺ , Ti ³⁺ , V ³⁺ , Cr ⁺ , Mn ³⁺ , Fe ³⁺ , Co ³⁺

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY



Figure 6.10. Stability constants for transition metal sulfate and organic complexes and their sulfide solubility constants, illustrating the Irving-Williams series. From Stumm and Morgan (1996).

served for organic ligands. The figure demonstrates an interesting feature of organic ligands: although the absolute value of stability complexes varies from ligand to ligand, the relative affinity of ligands having the same donor atom for these metals is always similar.

Organic molecules can often have more than one functional group and hence can coordinate a metal at several positions, a process called chelation. Such ligands are called *multidentate* and organic compounds having these properties are referred to as chelators or chelating agents. We will explore this topic in greater detail in the Chapter 14.

The kinetics of complex formation is quite fast in most cases, so that equilibrium can be assumed. There are exceptions, however. As we noted earlier in this section, all complexation reactions are ligand exchange reactions: water playing the role of ligand in "free ions". The rate at which complexes form is thus gov-

erned to a fair degree by the rate at which water molecules are replaced in the hydration sphere.

6.3.4 Complexation in Fresh Waters

Where only one metal is involved, the complexation calculations are straightforward, as exemplified in Example 6.7. Natural waters, however, contain many ions. The most abundant of these are Na⁺, K⁺, Mg²⁺, Ca²⁺, Cl⁻, SO²⁻, HCO⁻, CO²⁻, and there are many possible complexes between them as well as with H⁺ and OH⁻. To calculate the speciation state of such solutions an iterative approach is required. The calculation would be done as follows. First, we need the concentrations, activity coefficients, and stability constants (or apparent stability constants) for all species. Then we assume all ions are present as free ions and calculate the concentrations of the various possible complexes on this basis. In this pass, we need only consider the major ions (we can easily understand why with an example: formation of PbCl⁺ when the concentration of Pb is 10⁻⁸ or less and the abundance of Cl⁻ is 10⁻⁴ or more will have an insignificant affect on the free ion Cl concentration). We then iterate the calculation, starting with the free ion concentrations corrected for abundances of complexes we calculated in the previous iteration. This process is repeated until two successive iterations produce the same result. Although it sounds difficult, such calculations typically converge within 2 to 4 iterations. Example 6.8 shows how this is done. Once free ion concentrations of the major ligands are known, the speciation of trace metals may be calculated.

As Example 6.8 demonstrates, the major metals in fresh waters are present mainly as free ions (aquo complexes), as are the three most common anions, chloride, sulfate, and bicarbonate. The alkali and alkaline earth trace elements are also largely uncomplexed. Co^{2+} , Ni^{2+} , Zn^{2+} , Ag^+ , and Cd^{2+} are roughly 50% complexed. The remaining metals are present as primarily as complexes. B, V, Cr, Mo, and Re, as well as As and Se are present as anionic oxo-complexes. Other metals are usually present as hydroxide, carbonate, chloride, or organic complexes. Under reducing conditions, HS^- and S^{2-} complexes are important. In organic-rich waters such as swamps, organic complexes can be predominant. We will discuss organic complexes in more detail in Chapter 14, and speciation in seawater in Chapter 15.

CHAPTER O: AQUATIC CHEMISTRY

Example 6.7. Speciation in Fresh Water

Using the water and stability constants given in th adjacent tables, calculate the activities of the major specie in this water.

Answer: The first two problems we need to address an the nature of the carbonate species and activit coefficients. At this pH, we can see from Figure 6.1 that bicarbonate will be the dominant carbonate species. Making the initial assumption that all carbonate is bicarbonate, we can calculate ionic strength and activity coefficients using the

Debye-Hückel Law (equation 3.74). These are shown in the table below. Having the activity coefficients, we can calculate the approximate abundance of the car bonate ion by assuming all carbonate is bicarbonate:

ne	Analys	is of Stream	WATER (MM)	
es	Na ⁺	0.32 Cl ⁻	0. 22	
	K^+	0. 06 SO_4^{2-}	0. 12	
re	Mg^{2+}	0. 18 ΣCO ₂	1.0	
ty	Ca ²⁺	0.36 pH	8.0	

LOG STADILITY CONSTANTS

	OH⁻	HCO_3^-	CO_{3}^{2-}	SO_4^{2-}	Cl⁻	
H^+	14	6.35	10.33	1.99	_	
Na ⁺	-	-0.25	1.27	1.06	-	
K^+	-	-	-	0.96	_	
Mg ²⁺	2.56	1.16	3.4	2.36	_	
Ca ²⁺	1.15	1.26	3.2	2.31	-	

 HCO_3^-

9.17×10⁻⁰⁴

 CO_{3}^{2-}

2.12×10⁻⁰⁵ 9.46×10⁻⁰⁴ 1.75×10⁻¹⁰

1.62×10⁻⁰⁷ 2.51×10⁻⁰⁸ 6.27×10⁻⁰⁷

2.0310⁻⁰⁶ 1.65×10⁻⁰⁶ 6.10×10⁻⁰⁶

5.08×10⁻⁰⁶ 2.07×10⁻⁰⁶ 1.08×10⁻⁰⁵

9.12×10⁻⁰⁴ 1.12×10⁻⁰⁶ 1.65×10⁻⁰⁴

2.06×10⁻⁰⁵ 9.12×10⁻⁰⁴ 1.58×10⁻¹⁰

1.57×10⁻⁰⁷ 2.35×10⁻⁰⁸ 5.64×10⁻⁰⁷

$$[CO_{3}^{2-}] = \frac{\gamma_{HCO_{3}^{-}}\Sigma CO_{2}}{\beta a_{H^{+}}\gamma_{CO_{3}^{2-}}}$$
6.5

1×10⁻⁰⁶

5.09×10⁻⁰⁸

4.29×10⁻⁰⁹

 1×10^{-06}

 Ca^{2+}

ION ACTIVITIES: ITERATION 1

free ion OH-

 1×10^{-08}

 3.03×10^{-04}

5.69×10⁻⁰⁵

 1.39×10^{-04}

 2.77×10^{-04}

free ion OH-

 1×10^{-08}

3.03×10⁻⁰⁴

5.69×10⁻⁰⁵

 1.40×10^{-04}

 2.80×10^{-04}

ION ACTIVITIES: ITERATION 2

where β is the stability constant for the complexation reaction:

free ion

 H^{+}

 Na^+

 K^+

 Mg^{2+}

Ca²⁺

free ion

 H^+

 Na^+

 K^+

 Mg^{2+}

 Ca^{2+}

$$\mathrm{CO}_3^{2-} + \mathrm{H}^+ \rightleftharpoons \mathrm{HCO}_3^-$$

The "corrected" bicarbonate ion is then calculated as:

$$[\text{HCO}_{3}^{-}] = \Sigma CO_{2} - [\text{CO}_{3}^{2-}]$$

The result confirms our initial first order assumption that all carbonate is present as bicarbonate.

Using the concentrations and stability constants given as well as the activity coefficients we calculated, we can the m make a first pass at calculating the concentrations of the complexes. For example, MgCO₃ is calculated as

$$a_{MgCO_3} = \beta_{MgCO_3} a_{Mg^{2+}} a_{CO_3^{2-}}$$

The results of this first iteration are shown in a matrix. Chlorine does not complex with any of the major ions to any significant degree, so we can neglect it in our calculations.

We then correct the free ion activities by subtracting activities of the complexes they form. Thus for example, corrected free ion activity of Mg²⁺ is calculated as:

$$a_{Mg^{2+}}^{corr} = a_{Mg^{2+}}^{ini} - a_{MgOH^{-}} - a_{MgHCO_3^{-}} - a_{MgCO_3} - a_{MgSO_4}$$

We then repeat the calculation of the activities of the complexes using these corrected free ion activities. A second matrix shows

the results of this second iteration. A third table shows the percent of each ion present as a free ion (aquo complex). In fresh waters such as this one, most of the metals are present as free ions, the alkaline earths being 5% complexed by sulfate and carbonate.

the	% Free Ion								
une	Na ⁺	99.76% Cl⁻	100%						
	K^+	99.85% SO ₄ ²⁻	91.7%						
	Mg^{2+}	94.29% HCO 3	99.3%						

94.71% CO₂²⁻

 5.03×10^{-08} 1.84×10^{-06} 1.45×10^{-06} 5.14×10^{-06}

 3.92×10^{-09} 4.64×10^{-06} 1.83×10^{-06} 9.16×10^{-06}

 $HCO_3^ CO_3^2^-$

1

 SO_4^{2-}

9.33×10⁻⁰⁸

 SO_4^{2-}

 8.40×10^{-08}

25.3%

6.79×10⁻⁰⁷ 1.62×10⁻⁰⁴

October 27, 2005

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

6.4 Dissolution and Precipitation Reactions

6.4.1 Calcium Carbonate in Ground and Surface Waters

Calcium carbonate is an extremely common component of sedimentary rocks and is present in weathered igneous and metamorphic rocks. It is also a common constituent of many soils. Water passing through such soils and rocks will precipitate or dissolve calcite until equilibrium is achieved. This process has a strong influence on carbonate concentrations, hardness, and pH as well as dissolved calcium ion concentrations. Let's examine calcite solubility in more detail.

The solubility product of calcite is:

$$K_{sp-cal} = a_{Ca^{2+}} a_{CO_3^{2-}}$$
 6.52

This can be combined with equations 6.18–20 to obtain the calcium concentration water in equilibrium with calcite as a function of P_{CO_2} :

$$[Ca^{2+}] = P_{CO_2} \frac{K_1 K_{sp-cal} K_{sp-CO_2}}{K_2 \gamma_{Ca^{2+}} \gamma_{HCO_3}^2 [HCO_3^-]^2}$$
6.53

In a solution in equilibrium with calcite and a CO_2 gas phase and containing no other dissolved species, it is easy to modify equation 6.53 so that the calcium ion concentration is a function of P_{CO_2} only. A glance at Figure 6.1 shows that we can neglect OH⁻, H⁺, and CO_3^{2-} if the final pH is less than about 9. The charge balance equation in this case reduces to:

$$2[Ca^{2+}] = [HCO_3^{-}]$$
 6.54

Substituting this into 6.53, we obtain:

$$[Ca^{2+}] = P_{CO_2} \frac{K_1 K_{sp-cal} K_{sp-CO_2}}{4K_2 \gamma_{Ca^{2+}} \gamma_{HCO_3}^2 [Ca^{2+}]^2}$$
6.55

$$[Ca^{2+}] = \left\{ P_{CO_2} \frac{K_1 K_{sp-cal} K_{sp-CO_2}}{4K_2 \gamma_{Ca^{2+}} \gamma_{HCO_3^-}^2} \right\}^{1/3}$$
6.56

There are two interesting aspects to this equation. First, the calcium ion concentration, and therefore calcite solubility, increases with increasing P_{CO_2} . This might seem counter-intuitive at first, as one might think that that increasing P_{CO_2} should produce an increase the carbonate ion concentration and therefore drive the reaction toward precipitation. However, increasing P_{CO_2} decreases pH, which de-



Figure 6.11. Concentration of calcium ion in equilibrium with calcite at 25°C and 1 atm as a function of P_{CO_2} . From Drever (1988).

or

CHAPTER O: AQUATIC CHEMISTRY

creases CO₃²⁻ concentration, and therefore drives the reaction towards dissolution. Second, calcium ion concentration varies with the one-third power of P_{CO_2} (Figure 6.11). Because of this non-linearity, mixing of two solutions, both of which are saturated in Ca²⁺ with respect to calcite, can result in the mixture being undersaturated with respect to Ca^{2+} . For example, consider the mixing of stream and ground water. Stream water is in equilibrium with the atmosphere for which P_{CO_2} is 10^{-3.5}. On the other hand, P_{CO_2} in soils is often as high as 10⁻². So mixing between calcite-saturated groundwater and calcitesaturated surface water would produce a solution that is undersaturated with calcite.

Equation 6.56 describes calcite solubility for a system open to exchange with gaseous CO₂. For a P_{CO_2} of 10^{-3.5} (i.e., the atmosphere), this equation yields a calcium concentration of 1.39 mM. Water in pores and fractures in rocks does not exchange with a gas phase. Example 6.8 shows that under those circumstances, less calcite will dissolve; in the case of $P_{CO_2 initial} = 10^{-2}$, calcite saturation is reached at only 0.33 mM, or about a fourth as much. The difference is illustrated in Figure 6.12,



Figure 6.12. Comparison of the evolution of systems with constant P_{CO_2} (open systems) with those closed to gas exchange. After Stumm and Morgan (1996) and Deines et al. (1974).

which is a plot of log [HCO_3^-] vs. pH. Systems in equilibrium with constant P_{CO_2} (open systems) evolve along straight lines on this plot and ultimately reach calcite saturation at higher pH and lower [HCO_3^-] (and [Ca^{2+}]) than closed systems that initially equilibrate with the same P_{CO_2} .

6.4.2 Solubility of MG

There are a number of compounds that can precipitate from Mg-bearing aqueous solutions, including brucite (Mg((OH)₂), magnesite (MgCO₃), dolomite (CaMg(CO₃)₂), as well as hydrated carbonates such as hydromagnesite (MgCO₃(OH)₂· $3H_2O$). The stability of these compounds may be described by the following reactions:

$$Mg(OH)_2 \rightleftharpoons Mg^{2+} + 2OH^ K_{bru} = 10^{-11.6}$$
 6.57

$$MgCO_3 \rightleftharpoons Mg^{2+} + CO_3^{2-}$$
 $K_{mag} = 10^{-7.5}$ 6.58

$$CaMg(CO_3)_2 Mg^{2+} + Ca^{2+} + 2CO_3^{2-} K_{dol} = 10^{-17}$$
 6.59

(The solubility of dolomite is poorly known; values for this equilibrium constant vary between $10^{-16.5}$ and 10^{-20} .)

We can use these reactions and their equilibrium constants, together with the reactions for the carbonate system (equ. 6.15-6.17) to construct predominance diagrams for Mg-bearing solutions in equilibrium with these phases. For example, for reaction 6.57, we may derive the following relationship (assuming the solid is pure):

$$\log a_{Mg_{aq}^{2+}} = -pK_{bru} + 2pK_{W} - 2pH = 16.4 - 2pH$$
6.60

where we use the notation $pK = -\log(K)$. For reaction 6.58, the equilibrium constant relationship may be written as:

CHAPTER O: AQUATIC CHEMISTRY

Example 6.8. Calcite Solubility in a Closed System

Suppose ground water initially equilibrates with a P_{CO_2} of 10^{-2} and thereafter is closed to gas exchange, so that there is a fixed $\Sigma CO_{2 \text{ initial}}$. The water then equilibrates with calcite until saturation is reached. What will be the final concentration of calcium in the water. Assume ideal behavior and an initial calcium concentration of 0.

Answer: Since the system is closed, a conservation equation is a good place to start. We can write the following conservation equation for total carbonate:

$$\Sigma CO_2 = \Sigma CO_{2initial} + \Sigma CO_{2 \text{ from calcite}}$$

Since dissolution of one mole of calcite adds one mole of ΣCO_2 for each mole of Ca²⁺, this equation may be rewritten as:

$$\Sigma CO_2 = \Sigma CO_{2 \text{ initial}} + [Ca^{2+}]$$

Neglecting the contribution of the carbonate ion to total carbonate, this equation becomes: $[H_2CO_3] + [HCO_3^-] = ([H_2CO_3]_{initial}) + [Ca^{2+}]$ 6.61

where
$$([H_2CO_3]_{initial})$$
 denotes that amount of H_2CO_3 calculated from equation 6.21 for equilibrium with CO₂ gas; in this case a partial pressure of 10^{-2} . This can be rearranged to obtain:

$$([H_2CO_3]_{initial}) = [H_2CO_3] + [HCO_3^-] - [Ca^{2+}]$$
 6.62

Further constraints are provided by the three carbonate equilibrium product expressions (6.21–6.23) as well as the solubility product for calcite (6.50), and the charge balance equation. We assume a final pH less than 9 and no other ions present, so the charge balance equation reduces to equation 6.54. From equation 6.18 and the value of K_{CO_2} in Table 6.1, $[H_2CO_3]_{initial} = 10^{-2} \times 10^{-1.47}$ M. Dividing equation 6.19 by 6.20 yields:

$$\frac{K_1}{K_2} = \frac{[HCO_3^{-}]^2}{[H_2CO_3][CO_3^{2-}]}$$

Then substituting equations 6.52, 6.54, and 6.62 gives:

$$\frac{K_1}{K_2} = \frac{4[Ca^{2+}]^3}{K_{sp-cal} \{[H_2CO_3]_{initial} - [Ca^{2+}]\}}$$

Into this equation we substitute P_{CO_2} = [H₂CO₃] and rearrange to obtain:

$$[Ca^{2+}]^{3} + \frac{K_{1}K_{cal}}{4K_{2}}[Ca^{2+}] - \frac{K_{1}K_{cal}K_{CO_{2}}}{4K_{2}}(P_{CO_{2}initial}) = 0$$
6.63

This is a cubic that is readily solved for $[Ca^{2+}]$. For an initial P_{CO_2} of 10^{-2} , we calculate a calcium concentration of 0.334 mM.

$$\log \frac{a_{Mg^{2+}_{aq}}}{a_{MgCO_{3(s)}}} = -\log a_{CO_{3}^{2-}} - pK_{mag}$$
6.64

However, the carbonate concentration will depend on both total carbonate (or the partial pressure of CO_2) and pH. To simplify things, let's specify that the solid is pure, the solution ideal, and $\Sigma CO_2 = 10^{-2.5}$ M. Then we can think of three limiting cases: where carbonic acid, bicarbonate ion, and carbonate ion predominate. In the latter case, $[CO_3^2] \approx \Sigma CO_2 = 10^{-2.5}$ M, so we have:

$$\log[Mg^{2+}]_{aq} = 2.5 - pK_{mag} = -5.0$$
6.64a

When bicarbonate ion predominates, [HCO₃] $\approx \Sigma CO_2 = 10^{-2.5}$ M, and the carbonate ion concentration is:

CHAPTER O: AQUATIC CHEMISTRY



Figure 6.13. Predominance diagrams for Mg-bearing phases in equilibrium with aqueous solution. Total CO_2 is fixed at $10^{-2.5}$ M in 6.14a. The concentration of Mg²⁺ is fixed at 10^{-4} M in 6.14b. After Stumm and Morgan (1996).

$$\log[CO_3^{2-}] = -pK_2 + \log[HCO_3^{-}] + pH = -12.88 + pH$$

Substituting this into 6.64 we have:

 $\log[Mg^{2+}] = 5.33 - pH$ 6.64b

Finally, when carbonic acid predominates, $[H_2CO_3] \approx \Sigma CO_2 = 10^{-2.5}$ M. The carbonate ion concentration as a function of $[H_2CO_3]$ is given by equation 6.39b. Taking the log and substituting into equation 6.64, we have:

$$\log[Mg_{aq}^{2+}] = -pK_{mag} + pK_1 + pK_2 - \log[H_2CO_3] - 2pH = 26.68 - 2pH$$
 6.64c

We can use these equations to construct *stability, or predominance diagrams* in a manner similar to that used to construct pɛ-pH predominance diagrams (Chapter 3). Equations 6.63-6.64c express the Mg ion concentration as a function of pH and hence represent lines on on a plot of log $[Mg^{2+}]$ vs. pH. The lines divide the diagram (Figure 6.13) into 3 regions: (1) where only an Mg-bearing aqueous solution is stable (2) where magnesite is stable and (3) where brucite is stable. For example, on a plot of log $[Mg^{2+}]$ vs. pH, the predominance boundary between Mg^{2+}_{aq} and brucite plots as a line with a slope of -2 and an intercept of +16.4. Figure 6.13b shows a predominance diagram for this system, but where the Mg^{2+} concentration is fixed and P_{CO_2} and pH are the variables.

Virtually all natural solutions will contain dissolved calcium as well as magnesium. This being the case, we must also consider the stability of dolomite. We can construct similar predominance diagrams for these systems but we must add an additional variable, namely the Ca^{2+} concentration. To describe the relative stability of dolomite and calcite, it is more convenient to express the solubility of dolomite as:

$$CaMg(CO_3)_2 + Ca^+ \rightleftharpoons Mg^{2+} + 2CaCO_3$$

because the reaction contains calcite as well as dolomite. Since this reaction can be constructed by subtracting 2 times the calcite dissolution (equ. 6.50) from the dolomite dissolution (6.59), the equilibrium constant for this reaction can be calculated from:

$$\mathbf{K} = \frac{\mathbf{K}_{dol}}{\mathbf{K}_{cal}^2}$$

Figure 6.14 illustrates the stability of magnesite, dolomite, brucite and calcite as a function of P_{CO2} and the Ca^{2+}/Mg^{2+} concen-



Figure 6.14. Stability of magnesite, dolomite, calcite, and brucite in equilibrium with a Mg- and Ca-bearing aqueous solution.

CHAPTER O: AQUATIC CHEMISTRY

tration ratio. Whether any of these phases are stable relative to a Mg^{2+} -bearing solution depends on the Mg^{2+} concentration, which is not specified in the graph.

Example 6.9. Constructing Stability Diagrams Using the equilibrium constant data below as well as from Table 6.1, construct a stability diagram showing the stability of FeS, siderite (FeCO_{3(S)}), and Fe(OH)₂ as a function of total sulfide concentration and pH assuming $\Sigma CO_2 = 5 \times 10^{-2}$ M, Fe²⁺ = 10⁻⁶ M, and ideal behavior. Neglect any S²⁻. $K_{FeS} = 10^{-4.2} 6.65$ $K_{FeCO_3} = 10^{-0.1} 6.66$ $F \in S_{(S)} + H^+ \rightleftharpoons Fe^{2+} + HS^ F \in CO_{3(S)} + H^+ \rightleftharpoons Fe^{2+} + HCO_3^ K_{Fe(OH)_2} = 10^{12.9} \qquad 6.67$ $F e (O H)_{2(S)} + 2H^+ \rightleftharpoons Fe^{2+} + 2H_2O$ $K_{s} = 10^{-7}$ $H_2S_{(aq)} \rightleftharpoons H^+ + HS^-$ 6.68 Answer: Let's first consider reaction 6.65 above. Our first step is to set up an equation that describes the concentration of HS⁻ as a function of pH. From the conservation of sulfur, we have: $\Sigma S = [H_2S] + [HS^-]$ From the equilibrium constant expression for the dissociation of H_2S , we may substitute: $[H_2S] = \frac{[H^+][HS^-]}{K_s}$ and obtain: $\Sigma S = \frac{[H^+][HS^-]}{K_s} + [HS^-]$ 6.69

Solving for [HS⁻] we hav e : $[HS^-] = \Sigma S \frac{K_s}{K_s + [H^+]}$

We substitute this into the FeS (pyrrhotite) solubility product and solving for Σ S we have:

$$\Sigma S = \frac{K_{FeS}[H^+](K_{S} + [H^+])}{[Fe^{2+}]K_{S}}$$

or in log form: $\log (\Sigma S) = \log(K_S + 10^{-pH}) - pH - \log([Fe^{2+}]) - pK_{FeS} + pK_S = 6.70$ This plots as line ① on our ΣS vs. pH stability diagram (Figure 6.15). The area above the line is the region where FeS is stable.

Next, let's consider reaction 6.66, the solubility of siderite. We need an equation describing the concentration of HCO $\frac{1}{3}$ as a function of pH and Σ CO₂, which is equation 6.29. Substituting this into the siderite solubility product we have:

$$\mathbf{K}_{FeCO_3} = \frac{[Fe^{2+}]\Sigma CO_2}{\left\{\frac{[H^+]}{K_1} + 1 + \frac{K_2}{[H^+]}\right\}[H^+]}$$

or in log form:

$$pH + \log[Fe^{2^{+}}] + \log[\Sigma CO_2] - \log\left\{\frac{10^{-pH}}{K_1} + 1 + \frac{K_2}{10^{-pH}}\right\} + pK_{FeCO_3} = 0 \qquad 6.71$$

An approximate solution may be found by assuming $\text{HCO}_3^- = \Sigma \text{CO}_2$, which yields pH = 7.20. An exact solution requires an indirect method. Using the Solver in Microsoft ExcelTM, we obtain pH = 7.25, very close to our approximate solution (Solver uses a succession of "intelligent" guesses to find solutions to equations, such as 6.71 that have no direct solution). Thus siderite will precipitate when the pH is greater than 8.21 and, not surprisingly, this is independent of Σ S. The boundary between the Fe²⁺ and FeCO₃ field is then a vertical line (line ⁽²⁾) at pH = 7.25.

Now let's consider the solubility of fer rous iron hydroxide, The condition for precipitation of $Fe(OH)_2$ is described by the equation:

$$pH = \frac{pK_{Fe(OH)_2} - \log[Fe^{2+}]}{2} = 9.45$$
6.72

October 27, 2005

From

-pK

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

Thus $FeOH_2$ will not precipitate until pH reaches 9.5, i.e., above the point where $FeCO_3$ precipitates, so there is no boundary between a $Fe(OH)_2$ phase and a Fe^{2+} -bearing solution.

Next we need equations that describe the reactions between s olid phases. A reaction between $FeCO_3$ and $Fe(OH)_2$ can be obtained by subtracting reaction 6.67 from reaction 6.66. The corresponding equilibrium constant is obtained by dividing 6.67 by 6.66:

$$FeCO_3 + 2H_2O \rightleftharpoons (FeOH)_2 + H^+ + HCO_3^ K_{FeCO_3}/K_{Fe(OH)_2} = 10^{-13.0}$$

this we derive: $p H = \log [HCO_3^-] + 13.0$ 6.73

We can obtain an approximate solution, by simply assuming all carbonate is bicarbonate, in which case, we obtain pH = 11.69. Or we can substitute equation 6.29 for HCO₃⁻ and use the Solver in ExcelTM, which yields an exact solution of pH =11.02. Not surprisingly, our approximate solution is less accurate than in the previous case because at this high pH, the carbonate ion makes up a significant fraction of the total carbonate (Figure 6.1). The boundary between the FeCO₃ and Fe(OH)₂ fields is thus a vertical line at pH = 10.02 (line ^③).

The boundary between FeS and $FeCO_3$ is found by subtracting reaction 6.66 from 6.65 and dividing the corresponding equilibrium constants:

$$FeS + HCO_3^{-} \rightleftharpoons FeCO_3 + HS^{-} \qquad K_{FeS}/K_{FeCO_3} = 10^{-4.1} \qquad 6.74$$

Substituting 6.68 and 6.29 into the corresponding equilibrium constant expression, we have:

$$= \log[HS] - \log[HCO_{3}]$$

$$= \log \Sigma S - pK_{s} - \log(K_{s} + [H^{+}])$$

$$-\log \Sigma CO_{2} + \log\left\{\frac{[H^{+}]}{K_{1}} + 1 + \frac{K_{2}}{[H^{+}]}\right\}$$

Solving for Σ S and substituting 10^{-pH} for [H⁺], we have:

$$\log \Sigma S = -pK_{pyr-sid} + pK_{s} + \log(K_{s} + 10^{-pH}) + \log \Sigma CO_{2} - \log\left\{\frac{10^{-pH}}{K_{1}} + 1 + \frac{K_{2}}{10^{-pH}}\right\}$$

This plots as line 0 on our diagram. Finally, the reaction between FeS and Fe (OH)₂ is obtained by subtracting reaction 6.67 from reaction 6.65:

FeS +
$$2H_2O \rightleftharpoons Fe(OH)_2 + H^+ + HS^-$$

 $K_{FeS}/K_{Fe(OH)_2} = 10^{-17.1}$

From the equilibrium constant equation we obtain an expression of ΣS as a function of pH:

$$\log \Sigma S = pH - pK_{FeS - Fe(OH)_2} + pK_s + \log(K_s + 10^{-pH})$$

This plots as line ^⑤ on our diagram, which is now complete.

6.4.3 Solubility of SiO₂

Silicon is the most common element on the Earth's surface after oxygen. Its concentration in solution plays an important role in determining how weathering will proceed.

The dissolution of silica may be represented by the reaction:

$$SiO_{2(qtz)} + 2H_2O \rightleftharpoons H_4SiO_{4((aq))}$$
 6.75

The equilibrium constant expression is simply:



Figure 6.15. Stabil ity diagram showing the stable solid Fe-bearing phases in equilibrium with a solution containing 10^{-6} M Fe²⁺ and 5 $\times 10^{-3}$ M ΣCO_2 .

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

$$K_{qtz} = a_{H_4SiO_4} = 10^{-4}$$
 at 25°C 6.76

This is to say, water is saturated with respect to quartz when the concentration of H_2SiO_4 is 10^4 moles per kilogram, or about 7.8 ppm by weight SiO₂.

However, there are some complicating factors. First, precipitation of quartz seems to be strongly kinetically inhibited. Equilibrium is more likely to be achieved with amorphous silica, the equilibrium constant for which is 2×10^{-3} (~115 ppm). Second, H₄SiO₄ is a weak acid and undergoes successive dissociation with increasing pH:

$$H_4 SiO_4 \rightleftharpoons H_3 SiO_4^- + H^+$$
$$a_{H_5 SiO_4^-} a_{H^+} + a_{H^+}$$

$$K_1 = \frac{H_{3}SiO_4 + H_{4}}{a_{H_4SiO_4}} = 10^{-9.9}$$
$$H_3SiO_4 \rightleftharpoons H_2SiO_4^{-1} \leftrightarrow H_2SiO_4^{-1} + H_{4}$$

and

$$\mathbf{K}_{2} = \frac{a_{H_{2}SiO_{4}^{2}} - a_{H^{+}}}{a_{H_{3}SiO_{4}^{-}}} = 10^{-11.7}$$



Figure 6.16. Log activity of dissolved silica in equilibrium with quartz and amorphous silica (dashed line) as a function of pH. After Drever (1988).

The total dissolved silica concentration will be the sum of H_4SiO_4 , $H_3SiO_4^-$, and $H_2SiO_4^{-2}$. Assuming activity coefficients of unity, the concentration of dissolved silica is then:

$$[SiO_2]_T = [H_4SiO_4] \left\{ 1 + \frac{K_1}{a_{H^+}} + \frac{K_1K_2}{a_{H^+}^2} \right\}$$
6.77

From equation 6.77, we would expect silica solubility to be pH dependent. This dependence is illustrated in Figure 6.16.

We could have defined the second dissociation reaction as:

$$H_4SiO_4 \rightleftharpoons H_2SiO_4^2 + 2H^+$$

In which case, the equilibrium constant would be:

$$K_{2}^{*} = \frac{a_{H_{2}SiO_{4}^{2-}}a_{H^{+}}^{2}}{a_{H_{4}SiO_{4}^{-}}} = 10^{-9.9} \times 10^{-11.7} = 10^{-21.6}$$

and equation 6.67 would have been:

$$[SiO_2]_T = [H_4SiO_4] \left\{ 1 + \frac{K_1}{a_{H^+}} + \frac{K_2^*}{a_{H^+}^2} \right\}$$
 6.77a

The concentration of SiO_{2T} we calculate in this way would, of course, be identical. The point is, reactions and their corresponding equilibrium constants can be expressed in various ways, and we need to be alert to this.

6.4.4 Solubility of Al(OH)3 AND OTHER Hydroxides

The hydroxide is the least soluble salt of many metals. Therefore, it is the solubility of their hydroxides that controls the solubility of these metals in natural waters. These are shown in Figure 6.17. Since these dissolution reactions involve OH^- , they are pH dependent, and the slope of the solubility curve depends on the valence of the metal (e.g., -3 for Fe³⁺, -2 for Fe²⁺, -1 for Ag⁺). Let's consider in more detail the solubility of gibbsite, the hydroxide of aluminum.

CHAPTER O: AQUATIC CHEMISTRY

Dissolution of gibbsite $(Al(OH)_3)$ can be described by the reaction:

Al(OH)_{3(s)} + 3H⁺
$$\rightleftharpoons$$
 Al³⁺ + 3H₂O
K_{gib} = $\frac{a_{Al^{3+}}}{a_{H^{+}}^{3}}$ = 10^{-8.1} 6.78

However, a complication arises from hydrolyzation of the aluminum, which occurs in solutions that are not highly acidic (hydrolyzation is typical of many of the highly charged, \geq 3, metal ions):

$$Al^{3+} + H_2O \rightleftharpoons Al(OH)^{2+} + H^+$$

$$\mathbf{K}_{1} = \frac{a_{Al(OH)^{2^{+}}}a_{H^{+}}}{a_{Al^{3^{+}}}} = 10^{-5}$$
 6.79

Al³⁺ + 2H₂O ≈ Al(OH)₂⁺ + 2H⁺
K₂ =
$$\frac{a_{Al(OH)_2^+}a_{H^+}^2}{a_{Al^{3+}}^2} = 10^{-9.3}$$
 6.80

Al³⁺ + 3H₂O ≈ Al(OH)₃⁰ + 3H⁺
K₃ =
$$\frac{a_{Al(OH)_3^0} a_{H^+}^3}{a_{Al^{3+}}^3} = 10^{-16}$$
 6.81



Figure 6.18. Log activity of dissolved aluminum species and total Al (solid red line) in equilibrium with gibbsite as a function of pH.

$$Al^{3+} + 4H_2O \rightleftharpoons Al(OH)_3^{1-} + 4H^+ \qquad K_4 = \frac{a_{Al(OH)_3^{1-}}a_{H^+}^4}{a_{Al^{3+}}^3} = 10^{-22}$$
 6.82

Neglecting activity coefficients, the total dissolved aluminum concentration is given by:

$$a_{Al^{3+}T} = a_{Al^{3+}} \left\{ 1 + \frac{\mathbf{K}_1}{a_{H^+}} + \frac{\mathbf{K}_2}{a_{H^+}^2} + \frac{\mathbf{K}_3}{a_{H^+}^3} + \frac{\mathbf{K}_4}{a_{H^+}^4} \right\}$$
 6.83

Figure 6.18 shows the concentrations of the various aluminum species and total aluminum as a function of pH. The solubility of Al is low except at low and high pH, and that as pH increases Al³⁺ be-



Figure 6.17. Solubility of metal hydroxides as a function of pH. After Stumm and Morgan (1981).

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

comes increasingly hydrolyzed. Also note that where positively charged species dominate, e.g., Al^{3+} , solubility increases with decreasing pH; where negatively charged species dominate, solubility increases with increasing pH. Minimum solubility occurs in the range of pH of natural waters (which is just as well because it has some toxicity).

Equation 6.83 is a general result that can be applied to other metals that undergo hydrolyzation reactions, *when the reactions are expressed in the same form as those above*. A general form of this equation would thus be:

$$[\mathbf{M}]_{T} = [\mathbf{M}^{z}] \left\{ 1 + \frac{\mathbf{K}_{1}}{a_{H^{+}}} + \frac{\mathbf{K}_{2}}{a_{H^{+}}^{2}} + \frac{\mathbf{K}_{3}}{a_{H^{+}}^{3}} + \dots \right\}$$
 6.84

Thus iron and other metals show a pH dependence similar to Al. For Fe, the relevant equilibrium constants are:

$$Fe(OH)^{2^{+}} + H^{+} \rightleftharpoons Fe^{3^{+}} + H_{2}O$$
 $\log K_{1} = 2.2$ 6.85a $Fe(OH)_{2}^{+} + 2H^{+} \rightleftharpoons Fe^{3^{+}} + 2H_{2}O$ $\log K_{2} = 5.7$ 6.85b $Fe(OH)_{4}^{-} + 4H^{+} \rightleftharpoons Fe^{3^{+}} + 4H_{2}O$ $\log K_{3} = 21.6$ 6.85c $Fe(OH)_{3(x)} + 3H^{+} \rightleftharpoons Fe^{3^{+}} + 3H_{2}O$ $\log K = 3.2$ 6.85d

Using these equilibrium constants, the solubility of amorphous goethite as a function of pH is readily calculated and is shown in Figure 6.19.

6.4.5 Dissolution of Silicates and Related Minerals

The concentrations of Al and Si will usually not be controlled by equilibrium with quartz and gibbsite, rather by equilibrium with other silicates. An example of this is shown in Figure 6.20, which shows the concentration of dissolved Al in equilibrium with gibbsite, kaolinite and pyrophyllite at four different activities of dissolved silica. Only at the lowest dissolved silica concentrations will gibbsite precipitate before one of the aluminosilicates.

For the most part, silicates do not dissolve in the conventional sense, rather they react with water to release some ions to solution and form new minerals in place of the original ones. This phenomenon is known as *incongruent solution*. In considering such reactions, we can usually assume all Al remains in the solid phase. If we



Figure 6.19. Solubility of goethite as a function of pH. Solubility of individual Fe-hydroxide species shown as dashed lines.

consider only Al and Si, a simple reaction might be the breakdown of kaolinite to form gibbsite plus dissolved silica:

$$\frac{1}{2}\operatorname{Al}_{2}\operatorname{Si}_{2}\operatorname{O}_{5}(\operatorname{OH})_{4(s)} + \frac{5}{2}\operatorname{H}_{2}\operatorname{O} \rightleftharpoons \operatorname{Al}(\operatorname{OH})_{3(s)} + \operatorname{H}_{4}\operatorname{SiO}_{4}$$

$$6.86$$

Since the solid phases can be viewed as pure, the equilibrium constant for this reaction is simply:

$$\mathbf{K} = a_{H_4 SiO_4} = 10^{-4.4}$$
 6.87

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

which tells us that at H_2SiO_4 activities greater than $10^{-4.4}$, kaolinite is more stable than gibbsite. Similar reactions can, of course, be written between other phases, such as kaolinite and pyrophyllite. Introducing other ions into the system allows the possibility of other phases and other reactions, e.g.:

$$KAlSi_{3}O_{8} + H^{+} + 7H_{2}O \rightleftharpoons Al(OH)_{3} + K^{+} + 3H_{4}SiO_{4} \qquad K = \frac{a_{K^{+}}a_{H_{4}SiO_{4}}}{a_{H^{+}}}$$
 6.88

and:
$$\text{KAl}_3\text{Si}_3\text{O}_{10}(\text{OH})_2 + \text{H}^+ + \frac{3}{2} \text{H}_2\text{O} \rightleftharpoons \frac{3}{2} \text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4 + \text{K}^+ \qquad \text{K}_{mus-kao} = \frac{a_{K^+}}{a_{\mu^+}}$$
 6.89

From a series of such reactions and their corresponding equilibrium constant expressions, we can construct stability diagrams such as the one in Figure 6.21. The procedure for constructing this and similar diagrams is essentially similar to that used in construction of p ϵ -pH diagrams and is illustrated in Example 6.9. In the case of Figure 6.21, we seek an equilibrium constant expression containing the a_{K^+}/a_{H^+} ratio and the activity of H_4SiO_4 . From this expression we determine the slope and intercept, which allows us to plot the predominance boundary. For example, the boundary between the kaolinite and muscovite fields is given by equation 6.89. The equilibrium constant for the reaction is 10^4 , so the boundary is a line with slope 0 at log $a_{K^+}/a_{H^+} = 4$. The boundary between gibbsite and kaolinite is reaction 6.86, and equation 6.87, written in log form, defines the line dividing the two regions. This boundary thus plots as a vertical line at log $a_{H_4SiO_4} = -4.4$. The kaolinite – K-feldspar boundary is the reaction:



Figure 6.20. Total dissolved Al activity in equilibrium with gibbsite, pyrophyllite, and kaolinite as a function of pH at different dissolved silica activities. After Drever (1988).

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY



Figure 6.21. Stability diagram for the system $K_2O-Al_2O_3$ -SiO₂-H₂O at 25° C. After Drever (1988).

6.5 Clays and Their Properties

Clays are ubiquitous on the surface of the Earth. Unfortunately, the term *clay* has two meanings in geology: in a mineralogical sense it refers to a group of sheet silicates, in another sense it refers to the finest fraction of a sediment or soil. In addition to clay minerals, fine-grained oxides and hydroxides are present in the 'clay fraction'. Clays, in both senses, exert important controls on the composition of aqueous fluids, both because of chemical reactions that form them and because of their sorptive and ion exchange capacities. Generally, only clays in the mineralogical sense have true ion-exchange capacity, where ions in the clay can be exchanged for ions in the surrounding solution, but oxides and hydroxides can adsorb and desorb ions on their surfaces. We will first consider the mineralogy of the true clays, then consider their interaction with solution.

6.5.1 Clay Mineralogy

Clay minerals (*sensu stricto*) are sheet silicates. We can think of each sheet in a clay mineral as consisting of layers of silica tetrahedra bound to a hydroxide layer in which the cation (most comThe boundary thus plots as a line with a slope of - 2 and an intercept equal to the negative of the log of the equilibrium constant. Boundaries for the remaining fields can be derived similarly.

The fields in Figure 6.21 show the phase that is the most stable of those we considered in constructing the diagram. (Strictly speaking, it does not tell us whether the phase can be expected to precipitate or not, as this depends on the Al activity; however, because of the low solubility of Al, an aluminum-bearing phase can be expected to be stable in most instances.) By considering sodium and calcium as well as potassium, we can construct a 3-dimensional stability diagram, such as that in Figure 6.22, in a similar manner.

Because many low-temperature reactions involving silicates are so sluggish, equilibrium constants are generally calculated from thermodynamic data rather than measured.



Figure 6.22. Stability diagram for the system $K_2O-Na_2O-CaO-Al_2O_3-SiO_2-H_2O$ at 25° C. After Garrels and Christ (1965).

monly Al, Mg, or Fe) is in octahedral coordination much as in pure hydroxide minerals such as gibbsite $(Al(OH)_3)$ or brucite $(Mg(OH)_2)$. Brucite and gibbsite are structurally similar to each other except that in gibbsite every third octahedral site is left empty to maintain charge balance. This is illustrated in Figure 6.23. Because only 2 out of 3 octahedral sites are occupied, gibbsite is said to have a *dioctahedral*

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

structure, while brucite is said to be *trioc-tahedral*. This terminology also applies in clay minerals in exactly the same sense.

6.5.1.1 Kaolinite Group (1:1 Clays)

The simplest clays consist of a tetrahedral silicate layer and an octahedral hydroxide layer, hence the term 1:1 clays. Kaolinite, Al₄Si₄O₁₀(OH)₈, is a good example. The structure of kaolinite is shown in Figure 6.24. Its unit cell consists of a layer of silica tetrahedra bound to an octahedral alumina layer whose structure is very similar to that of gibbsite except that some hydroxyls are replaced by oxygens. Note that individual sheets are not bound together; they are held together only by van der Waals interactions, which are quite weak. The structure of serpentine, Mg₆Si₄O₁₀(OH)₈, is similar to that of kaolinite, with Mg replacing Al, and every octahedral site is occupied. In both minerals, a mismatch in the spacing of octahedra and tetrahedra results in a curvature of the lattice. Also successive layers of kaolinite are generally stacked in a random manner.

6.5.1.2 Pyrophyllite Group (2:1 Clays)

This is a large group of clay minerals consisting of a "hydroxide layer" sandwiched between two layers of silica tetrahedra, hence the term 2:1 clays. More of the hydroxyls in the hydroxide layer are replaced by oxygen than in 1:1 clays. The simplest two such clays are pyrophyllite, Al₂Si₄O₁₀(OH)₂, and talc, $Mg_3Si_4O_{10}(OH)_2$. In pyrophyllite, every third octahedral site is vacant, so that it, like kaolinite, is said to be dioctahedral, while serpentine and talc are trioctahedral. The structure of pyrophyllite is shown in Figure 6.25. Other, more complex clays, including those of the smectite group, the biotite group, the vermiculite group, saponite, and muscovite, have structures similar to those of pyrophyllite and talc. These are derived in the following ways:

 Substitution of Al⁺³ for Si⁺⁴ in the tetrahedral sites resulting in a charge deficiency that is balanced by the



Figure 6.23. Structure of gibbsite and brucite. (a) plan (vertical) view; (b) expanded x-sectional view.



Figure 6.24. Structure of kaolinite. (a) Plan view of the tetrahedral layer. (b) Cross-sectional view.

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

presence of a cation between the layers (interlayer positions). When the number of interlayer cations is small, they are generally exchangeable; when the number is large, K is typically the cation and it is not very exchangeable.

Substitution of Mg²⁺, Fe²⁺, Fe³⁺ or a similar cation for Al³⁺ or substitution of Fe²⁺, Fe³⁺, or Al³⁺ for Mg²⁺ in the octahedral layer. Where charge deficiency results, it is balanced by the presence of exchangeable cations in the interlayer sites, or by vacancies in the octahedral sites.

Smectites are distinguished by their expansion to a unit cell thickness of 14 Å upon treatment with ethylene glycol. This expansion results from entry of the ethylene glycol molecule into the interlayer position of the clays. Smectites generally also have water present in the interlayer space, the amount of water being determined by the cation Figure 6.25. Structure of pyrophyllite.

present. Generally there is little water present when the interlayer cation is divalent Mg or Ca, but can be very considerable when the cation is Na. The amount of water also depends on the humidity; as a result smectites, and sodium bearing ones in particular will swell on contact with water, affecting permeability. The most common smectite is montmorillonite, $X_{1/3}(Mg_{1/3}Al_{5/3})Si_4O_{10}$ (OH)₂·*n*H₂O. The interlayer cation of smectites is exchangeable: in a NaCl solution it will be Na, in a CaCl₂ solution it will be Ca, etc.

Vermiculites have a higher net charge on the 2:1 layer (that is, there is greater cation deficiency). As a result, the electrostatic forces holding the layers together are greater so that the interlayer space is less expandable and the interlayer cations less exchangeable.

Micas (biotite and muscovite) are related to pyrophyllite and talc by substitution of an Al³⁺ for a Si⁴⁺ in a tetrahedral site. The structure of muscovite is illustrated in Fig-





Figure 6.26. Structure of muscovite $(KAl_3Si_3O_{10}(OH)_2)$. The structure of the clay illite is similar, but illite typically has less K and Al and more Si than muscovite.

ure 6.26. The result is that the silicate layers are relatively strongly bound to the interlayer K, which is not normally exchangeable. *Illite* is a name applied to clay-sized micas, though it is sometimes restricted to the dioctahedral mica (muscovite). Generally, illite has less K and Al and more Si than igne-

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY





ideal formula of chlorite is $(Mg, Fe, Al)_6 (SiAl)_4O_{10}(OH)_8$, where the elements in parentheses can be in any proportions. The structure of chlorite is shown in Figure 6.27. Chlorites with unit cells consisting of a single tetrahedral and single octahedral layer also occur and are called septechlorite (because they have a 7Å spacing; true chlorites have 14Å spacing), but

are less stable and uncommon.

6.5.2 ION-Exchange Properties of Clays

One of the most important properties of clays is their capacity for ion exchange. In soil science, the term *ion exchange* refers specifically to replacement of an ion adsorbed to the surface by one in solution. However, we shall use the term in a more general sense here, and include as well exchange reactions between ions in solution and ions bound within the solid. The ability of a substance to exchange ions is Table 6.3. Ion Exchange Capacity of Clays

Clay	Exchange Capacity (meq/100g)
Smectite	80-150
Vermiculite	120-200
Illite	10-40
Kaolinite	1-10
Chlorite	<10

called the *ion exchange capacity* and is generally measured in equivalents or milliequivalents (meq). Ion exchange capacities of clays are listed in Table 6.3.

The exchange reaction of two monovalent ions between clay and solution may be written:

$$X_{clay} + Y^+ \rightleftharpoons Y_{clay} + X^+$$

The corresponding equilibrium constant expression is written as:

$$K = \frac{a_{Y-clay}a_{X^{+}}}{a_{X-clay}a_{Y^{+}}} \qquad \text{or:} \qquad \qquad \frac{a_{Y-clay}}{a_{X-clay}} = K \frac{a_{Y^{+}}}{a_{Y^{+}}} \qquad \qquad 6.90$$

If we express this with molar concentrations in solution and mole fractions (X) in the solid rather than activities:

$$\frac{X_{Y-clay}}{X_{X-caly}} = K' \frac{[Y^+]}{[X^+]}$$

ous or metamorphic muscovite; in this sense it can be viewed as a solid solution of muscovite and pyrophyllite.

Because of the structural similarity of various 2:1 clays, they can form crystals that consist of layers of more than one type, for example illite-smectite. In addition, layers of gibbsite or brucite may occur in smectite. Different layers may be distributed randomly or may be ordered. These are called, for example, mixed-layer chlorite-smectite or hydroxy-interlayer smectite.

6.5.1.3 Chlorite Group (2:2 clays)

This group is characterized by having a unit cell consisting of two tetrahedral layers and two (hydroxide) octahedral layers. The

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

K' is called the *selectivity constant*. It expresses the selectivity of the clay for the Y ion over the X ion. Because we have expressed it in mole fraction rather than activity, we can expect its value to depend on the composition of both the clay and the solution. We may also define a *distribution coefficient* K_d as:

$$\mathbf{K}_{d} = \frac{[X_{clay}]}{[X^{+}]} \tag{6.91}$$

Where Henry's Law holds, K_d should be independent of the concentration of the ion in solution or in the clay, but it will nevertheless depend on the overall composition of the solution and the clay (in other words, the Henry's law constant will be different for different clays and solutions). A more general expression for the equilibrium constant is:

$$\frac{a_{Y-clay}^{v_Y}}{a_{X-clay}^{v_X}} = \mathbf{K} \frac{a_Y^{v_Y}}{a_X^{v_X}}$$

$$6.92$$

where v is the stoichiometric coefficient.

The power term is important. Consider the case of exchange between Na⁺ and Ca²⁺. The reaction is:

$$2\mathrm{Na}_{\mathrm{clay}}^{+} + \mathrm{Ca}_{\mathrm{aq}}^{2+} \rightleftharpoons \mathrm{Ca}_{\mathrm{clay}}^{2+} + 2\mathrm{Na}_{\mathrm{aq}}^{+}$$

$$6.93$$

The K' expression is:

$$\frac{X_{Ca^{2+}}}{X_{Na^{+}}^{2}} = \mathbf{K}' \frac{[\mathbf{Ca}^{2+}]}{[\mathbf{Na}^{+}]^{2}}$$
6.94

If we assume that (1) the mole fractions of Na⁺ and Ca²⁺ in the clay must sum to one (i.e., they are the only ions in the exchanging site), (2) molar concentrations of 1 for Na and Ca in solution, and (3) K' = 1, solving equation 6.94 yields $X_{Na} = 0.62$ and $X_{Ca} = 0.38$. If we kept the ratio of Ca and Na in solution constant, but dilute their concentrations 1000 fold, we obtain $X_{Na} = 0.03$ and $X_{Ca} = 0.97$. Thus by diluting the solution, the divalent cation has almost entirely replaced the monovalent ion. The composition of the clay will thus depend on the ionic strength of the solution. The dominant exchangeable cation is Ca²⁺ in fresh water, but Na⁺ in seawater.

6.6 Mineral Surfaces and Their Interaction With Solutions

Reactions between solutions and solids necessarily involves the interface between these phases. The details of interface processes thus govern equilibria between solids and solutions. Because clays and other sedimentary particles are typically very small, their surface area to volume ratio is high. This adds to the importance of surface chemistry. For example, the concentrations of many trace elements, particularly the transition metals, dissolved in streams, rivers, and the oceans are controlled not by precipitation and dissolution, but rather by adsorption on and desorption from mineral and organic surfaces. These surface reactions maintain the concentrations of these elements in seawater well below saturation levels. Surface processes also play an important role in soil fertility. Soils have concentrations of many elements above levels one would predict from equilibrium dissolution and precipitation because of adsorption onto particles surfaces. Surface adsorption will also strongly affect the dispersion of pollutants in soils, ground and surface waters. We discussed some aspects of surface chemistry in Chapter 5 within the context of kinetic fundamentals. We return to it in this chapter in a broader context.

6.6.1 Adsorption

We can define adsorption as attachment of an ion in solution to a preexisting solid surface, for example a clay particle. Adsorption involves one or more of the following:

- *Surface complex formation:* The formation of coordinative bonds between metals and ligands at the surface. Considered in isolation, this process is very similar to the formation of complexes between dissolved components.
- *Electrostatic interactions:* As we shall see, solid surfaces are typically electrically charged. This electrostatic force, which is effective over greater distances than purely chemical forces, affects surface

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

complex formation and loosely binds other ions to the surface. For solutions, we were able to make the simplifying assumption of electrical neutrality. We cannot make this assumption about surfaces.

• *Hydrophobic adsorption:* Many organic substances, most notably lipids, are highly insoluble in water due to their non-polar nature. These substances become adsorbed to surfaces, not because they are attracted to the surface, but rather because they are repelled by water.

The interaction of the three effects make quantitative prediction of adsorption behavior more difficult than prediction of complexation in solution. The functional groups of mineral and organic surfaces have properties similar to those of their dissolved counterparts. In this sense, surface complexation reactions are similar to complexation reactions in solution. However, reactions between these surface groups and dissolved species is complicated by the proximity of surface groups to each other, making them subject to long range electrostatic forces from neighboring groups. For example, the surface charge will change systematically as the adsorbed surface concentration of a positive species such as H⁺ increases. This change in surface charge will decrease the attraction between H⁺ ions and the surface. As a result, the equilibrium constant for the surface protonation reaction will change as the surface concentration of H⁺ increases.

We found in Chapter 5 that adsorption is usually described in terms of adsorption isotherms. We introduced two such isotherms, the *Langmuir* isotherm:

$$\Theta_M = \frac{\mathbf{K}_{ad}[M]}{1 + \mathbf{K}_{ad}[M]} \tag{5.130}$$

(where Θ_M is the fraction of surface sites occupied by species M, [M] is the dissolved concentration of M, and K_{ad} is the adsorption equilibrium constant), and the *Freundlich* isotherm:

$$\Theta_{\rm M} = \mathbf{K}_{\rm ad}[M]^n \tag{5.134}$$

where *n* is an empirical constant. We derived the Langmuir isotherm from kinetic fundamentals, but we could have also derived it from thermodynamics. Inherent in its derivation are the assumptions that (1) the free energy of adsorption is independent of the number of sites available, and therefore that (2) the law of mass action applies, and that (3) only a monolayer of adsorbate can form. The Langmuir isotherm thus shows a decrease in the fraction of M adsorbed when the concentration of M in solution is high, reflecting saturation of the surface. In contrast, the Freundlich isotherm, which is merely empirical, shows no saturation. We also found that at low relative saturation of the surface, the Freundlich isotherm with n=1 approximates the Langmuir isotherm.

Adsorption phenomena are can be treated with the *surface complexation model*, which is a generalization of the Langmuir isotherm (Stumm and Morgan, 1996; Morel and Hering, 1993). The model incorporates both *chemical bonding of solute species to surface atoms* and *electrostatic interactions between the surface and solute ions*. The model assumes that these two effects can be treated separately. Thus the free energy of adsorption is the sum of a complexation, or intrinsic, term and an electrostatic, or coulombic term:

$$\Delta G_{ad} = \Delta G_{intr} + \Delta G_{coul}$$

$$6.95$$

From this it follows that the adsorption equilibrium constant can be written as:

$$\mathbf{K}_{\mathrm{ad}} = \mathbf{K}_{\mathrm{intr}} \, \mathbf{K}_{\mathrm{coul}} \tag{6.96}$$

Letting =S denote the surface site and M denote a solute species, we may write the adsorption reaction as:

$$\equiv S + M \rightleftharpoons \equiv SM$$

Let's begin by considering comparatively simple surfaces: those of metal oxides. Although silicates are likely to be more abundant that simple oxides, the properties of silicate surfaces approximate those of mixtures of their constituent oxides' surfaces. Hence what we learn from consideration of oxides can be applied to silicates as well. We will initially focus just on the intrinsic terms in equations 6.95 and 6.96. We will return to the coulombic term at the end of this section.

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

Oxygen and metal atoms at an oxide surface are incompletely coordinated; i.e., they are not surrounded by oppositely charged ions as they would be in the interior of a crystal (Figure 6.28a). Consequently, mineral surfaces immersed in water attract and bind water molecules (Figure 6.28b). These water molecules can then dissociate, leaving a hydroxyl group bound to the surface metal ion. We may write this reaction as:

$$\equiv M^{+} + H_2 O \rightleftharpoons \equiv MOH + H^{+}$$

where \equiv M denotes a surface metal ion.

a

C

$$\equiv O^- + H_2 O \rightleftharpoons \equiv OH + OH^-$$

Thus the surface on an oxide immersed in water very quickly becomes covered with hydroxyl groups (Figure 6.28c), which we can write as \equiv SOH and which are considered to constitute part of the surface rather than the solution. These hydroxyl groups can then act as either proton acceptors or proton donors through further association or dissociation reactions, e.g.:

 $\equiv SOH + H^+ \rightleftharpoons \equiv SOH_2^+$ $\equiv SOH \rightleftharpoons \equiv SO^- + H^+$

or

We should not be surprised to find that these kinds of reactions
$$b$$
 are strongly pH dependent.

Adsorption of metals to the surface may occur through replacement of a surface proton, as is illustrated in Figure 6.29a, while ligands may be absorbed by replacement of a surface OH group (Figure 6.29b). The adsorbed metal may bind an additional ligand (Fig. 6.29c), and the adsorbed ligand may bind an additional metal (Fig. 6.29d).

An additional possibility is multidentate adsorption, where a metal or ligand is bound to more than one surface site (Figures 6.29e and 6.29f). This raises an interesting dilemma for the Langmuir isotherm. Where x sites are involved, we could write the reaction as:

$x \equiv S + M \rightleftharpoons \equiv S_x M$

and the corresponding equilibrium constant expression as:

$$\mathbf{K}_{ad} = \frac{\left[\equiv S_x M\right]}{\left[\equiv S\right]^x \left[M\right]} \tag{6.97}$$

where x is the number of sites involved and M is the species being adsorbed. This assumes, however, that the probability of finding x sites together is proportional to the xth power of concentration, which is not the case. A better approach is to assume that the reaction occurs with a multidentate surface species, $\equiv S_x$ and that its concentration is $[\equiv S]/x$. The equilibrium constant is then:

$$K_{ad} = \frac{[\equiv S_x M]}{[M][\equiv S]/x}$$

Alternatively, the 1/x can be contained within the equilibrium constant.



Figure 6.28. (a) Metal ions (small red spheres) and oxygens (large gray spheres) on a mineral surface are incompletely coordinated, leading to a partial charge on the surface (indicated by δ^+ and δ^-). (b) When the mineral surface is immersed in water, water molecules coordinate metal ions on the surface. (c) Water molecules will dissociate leaving hydroxyl groups coordinating metal ions. Protons (small dark spheres) will associate with surface oxygens, forming additional hydroxyl groups.

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

 $(a) = x - OH + M^{z+} \implies x - O - M^{(z-1)} + H^{+}$ $(b) = x - OH + L^{-} \implies x - L + OH^{-}$ $(c) = x - O - M^{(z-1)} + L^{-} \implies x - O - M - L^{(z-2)}$ $(d) = x - L + M^{z+} \implies x - L - M^{(z+1)}$ $(e) = x - O - H + M^{z+} \implies x - O \longrightarrow M^{(z-2)+} + 2H^{+}$ $(f) = x - O - H + HPO_{4} \implies x - O \longrightarrow P \longrightarrow P \longrightarrow P + 2OH^{-}$

Figure 6.29. Complex formation of solid surfaces may occur when (a) a metal replaces a surface proton, or (b) a ligand replaces a surface OH group. The adsorbed metal (c) may bind an additional ligand, and the ligand (d) may bind an additional metal. Multidentate adsorption involves more than one surface site (e, f).

Since surface bound protons and OH groups are almost inevitably involved in adsorption, we would expect that adsorption of metals and ligands will be strongly pH dependent. This is indeed the case, as may be seen in Figure 6.30 and 6.32: adsorption of cations increases with increasing pH while adsorption of anions decreases with increasing pH. Figure 6.32 shows that adsorption of metals on SiO₂ goes from insignificant to nearly complete over a very narrow range of pH. This strong dependence on pH certainly reflects protonation of the surface as we have discussed above, but it also reflects the extent of hydrolysis of the ion in solution.

As is the case with soluble complexes, surface complexes may be divided into inner sphere and outer sphere complexes (Figure 6.33). Inner sphere complexes involve some degree of covalent bonding between the adsorbed species and atoms on the surface. In an outer-sphere complex, one or more water molecules separate the adsorbed ion and the surface; in this case adsorption in-

volves only electrostatic forces. The third possibility is that an ion may be held within the diffuse layer (see following section) by long range electrostatic forces.

6.6.2 Development of Surface Charge and the Electric Double Layer

Mineral surfaces develop electrical charge for three reasons:

1. Complexation reactions between the surface and dissolved species, such as those we discussed in the pre-

vious section. Most important among these are protonation and deprotonation. Because these reactions depend on pH, this aspect of surface charge is pH dependent. This pH dependence is illustrated in Figure 6.34.

2. *Lattice imperfections* at the solid surface as well as substitutions within the crystal lattice (e.g., Al³⁺ for Si⁴⁺). Because the ions in interlayer sites of clays are readily exchangeable, this mechanism is particularly important in the development of surface charge in clays.

3. *Hydrophobic adsorption*, primarily of organic compounds, and "surfactants" in particular. We will discuss this effect in Chapter 14.



Figure 6.30. Binding of ligands (anions) on the surface of hydrous ferric oxide ($\Sigma Fe = 10^{-3} \text{ M}$) from dilute solution ($5 \times 10^{-7} \text{ M}$; I=0.1) as a function of pH. From Stumm and Morgan (1996).
CHAPTER O: AQUATIC CHEMISTRY

Example 6.10. Adsorption of Pb^{2*} on H	Iydrous Ferric Oxide as a Fun	стіом of pH					
Using the following apparent equilibrium cor	nstants:						
$\equiv \text{FeOH}_2^+ \rightleftharpoons \equiv \text{FeOH} + \text{H}^+$	$pK_{a1} = 7.29$	6.98					
\equiv FeOH \rightleftharpoons \equiv FeO ⁻ +H ⁺	$pK_{a2} = 8.93$	6.99					
$\equiv F e O P b^{+} \rightleftharpoons \equiv F e O^{-} + P b^{2+} p K_{ad} = 8.15$	$P b^{2+} + H_2 O \rightleftharpoons PbOH^+ + H^+$	$p K_{OH^{-}} = 7.7$					
calculate the fraction of surface adsorbed Pb as a	a function of pH from pH 5 to pH 8	for concentrations					
of surface sites of 10^{-3} M, 10^{-4} M, and 10^{-5} M assuming a total Pb concentration of 10^{-9} M.							
Answer: The quantity we wish to calculate is	$[=FeOPb^+]/\Sigma Pb$, so we want to find	l an expression for					
$[=FeOPb^+]$ as a function of pH. We chose our c	components to be H^+ , Pb^{2+} , and $\equiv FeO$	OH^+_2 and begin by					

writing the two relevant conservation equations:

$$\Sigma Pb = [Pb^{2+}] + [PbOH^{+}] + [\equiv FeOPb^{+}]$$
 6.100

$$\Sigma = Fe = [=FeOH_{2}^{+}] + [=FeOH] + [=FeO^{-}] + [=FeOPb^{+}]$$
 6.101

From the equilibrium constant expressions, we have the following:

$$[\equiv \text{FeOH}] = \frac{[\equiv \text{FeOH}_2^+]}{[\text{H}^+]} K_{\text{al}}$$

$$6.102$$

$$[\equiv \text{FeO}^{-}] = \frac{[\equiv \text{FeOH}]}{[\text{H}^{+}]} K_{a2} = \frac{[\equiv \text{FeOH}_{2}^{+}]}{[\text{H}^{+}]^{2}} K_{a1} K_{a2}$$
 6.103

$$[\equiv \text{FeOPb}^+] = [\equiv \text{FeO}^-][\text{Pb}^{2+}]K_{ad}^{-1} = \frac{[\equiv \text{FeOH}_2^+]}{[\text{H}^+]^2}[\text{Pb}^{2+}]\frac{K_{a1}K_{a2}}{K_{ad}} \qquad 6.104$$

Substituting equations 6.102-6.104 into 6.101 we have:

$$\Sigma \equiv \text{Fe} = \left[\equiv \text{FeOH}_{2}^{+} \right] \left\{ 1 + \frac{K_{a1}}{[\text{H}^{+}]} + \frac{K_{a1}K_{a2}}{[\text{H}^{+}]^{2}} + \frac{[\text{Pb}^{2+}]}{[\text{H}^{+}]^{2}} \frac{K_{a1}K_{a2}}{K_{ad}} \right\} \qquad 6.105$$

Since the $[Pb^{2+}]$ is small, the last term on the right can be neglected so we have:

$$\Sigma \equiv \text{Fe} \cong [\equiv \text{FeOH}_2^+] \left\{ 1 + \frac{K_{a1}}{[\text{H}^+]} + \frac{K_{a1}K_{a2}}{[\text{H}^+]^2} \right\}$$

In a similar way, we obtain:

$$\Sigma Pb = [Pb^{2+}] \left\{ 1 + \frac{[\equiv FeOH_2^+]}{[H^+]^2} \frac{K_{a1}K_{a2}}{[H^+]^2} + \frac{K_{OH}}{[H^+]} \right\}$$

(

Solving this pair of equations for $[=FeOH_2^+]$ and $[Pb^{2+}]$, and substituting these into 6.102, we obtain:

$$[\equiv \text{FeOPb}^{+}] = \frac{[\equiv \text{FeOH}_{2}^{+}]}{[\text{H}^{+}]^{2} + [\text{H}^{+}]\text{K}_{a1} + \text{K}_{a1}\text{K}_{a2}} \left\{ \frac{\Sigma \equiv \text{Fe } \text{K}_{a1}\text{K}_{a2}/\text{K}_{ad}}{[\text{H}^{+}]^{2} + [\text{H}^{+}]\text{K}_{a1} + \text{K}_{a1}\text{K}_{a2}} + \frac{\text{K}_{\text{OH}}}{[\text{H}^{+}]} \right\} \left\{ \left(\frac{\text{K}_{a1}\text{K}_{a2}}{\text{K}_{ad}} \right) \right\}$$

Dividing by Σ Pb and simplifying, we have:

$$\frac{[\equiv \text{FeOPb}^+]}{\Sigma \text{Pb}} = \frac{\Sigma \equiv \text{FeK}_{a2}}{K_{ad}([\text{H}^+]^2/\text{K}_{a1} + [\text{H}^+] + \text{K}_{a2}) + ([\text{H}^+]/\text{K}_{a1} + 1 + \text{K}_{a2}/[\text{H}^+])\text{K}_{OH}\text{K}_{ad} + \Sigma \equiv \text{FeK}_{a2}}$$

The result is shown in Figure 6.31. For the highest concentration of surface sites, Pb goes from virtually completely in solution to virtually completely adsorbed within 2 pH units.

)

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

Thus there are several contributions to surface charge density. We define σ_{net} as the *net density of electric charge on the solid surface,* and express it as:

$$\sigma_{\rm net} = \sigma_0 + \sigma_{\rm H} + \sigma_{\rm SC} \qquad 6.106$$

where σ_0 is the *intrinsic* surface charge due to lattice imperfections and substitutions, σ_H is the net proton charge, i.e., the charge due to binding H⁺ and OH⁻, σ_{SC} is the charge due to other surface complexes. σ is usually measured in coulombs per square meter (C/m²). σ_H is given by:

$$\sigma_{\rm H} = \mathcal{F}(\Gamma_{\rm H} - \Gamma_{\rm OH}) \qquad 6.107$$

where \mathcal{F} is the Faraday constant and Γ_{H} and Γ_{OH} are the adsorption densities (mol/m²) of H⁺ and OH⁻ respectively. In a similar way, the charge due to other surface complexes is given by

$$\sigma_{\rm SC} = \mathcal{F}(z_{\rm M}\Gamma_{\rm M} + z_{\rm A}\Gamma_{\rm A}) \qquad 6.108$$

where the subscripts M and A refer to metals and anions respectively, Γ is again adsorption density and z is the charge of the ion. The surface complex term may also be broken into an inner sphere and outer sphere component:

$$\sigma_{\rm SC} = \sigma_{\rm IS} + \sigma_{\rm OS} \qquad 6.109$$

Thus net charge on the mineral surface is:

$$\sigma_{\text{net}} = \sigma_0 + \mathcal{F}(\Gamma_H - \Gamma_{OH} + Z_M \Gamma_M + Z_A \Gamma_A) \quad 6.110$$

Figure 6.34 shows that at some value of pH the surface charge, σ_{nev} will be zero. The pH at which this occurs is known as the isoelectric point, or zero point of charge (ZPC). The ZPC is the pH at which the



Figure 6.33. Inner sphere surface complexes involve some degree of covalent bonding between the surface and the ion; outer sphere complexes form when one or more water molecules intervenes between the surface and the ion. Ions may also be held in the diffuse layer by electrostatic forces.



Figure 6.31. Calculated adsorption of Pb^{2+} on hydrous ferric oxide for three different concentrations of surface sites: 10^{-3} M, 10^{-4} M, and 10^{-5} M.



Figure 6.32. Adsorption of metals on SiO₂. Symbols show experimentally determined values; curves show theoretical values (from James and Healy, 1972).

charge on the surface of the solid caused by binding of all ions is 0, which occurs when the charge due to adsorption of cations is balanced by charge due to adsorption of anions. A related concept is the point of zero net proton charge (pznpc), which is the point of zero charge when the charge due to the binding of H⁺ and OH⁻ is 0; i.e., pH where $\sigma_{\rm H} = 0$. Table 6.4 lists values of the point of zero net proton condition for some important solids. Surface charge depends on the nature of the surface, the nature of the solution, and the ionic strength of the latter. An important feature of the point of zero charge, however, is that it is independent of ionic strength, as is illustrated in Figure 6.35.

CHAPTER O: AQUATIC CHEMISTRY

6.6.2.1 DETERMINATION OF SURFACE CHARGE

The surface charge due to binding of protons and hydroxyls is readily determined by titrating a solu-

tion containing a suspension of the material of interest with strong acid or base. The idea is that any deficit in H⁺ or OH⁻ in the solution is due to binding with the surface. For example, consider a simple hydroxide surface with surface species \equiv SOH⁺₂ and \equiv SO⁻ (as well as \equiv SOH⁰). Charge balance requires that:

$$C_A - C_B + [OH^-] - [H^+] = [\equiv SOH_2^+] - [\equiv SO^-]$$

where C_A and C_B are the concentrations of conjugate of the acid or base added (e.g., Na⁺ is the conjugate of the base NaOH) and [\equiv SOH₂⁺] and [\equiv SO⁻] are the concentrations (in moles per liter) of the surface species. The surface charge, Q (in units of moles of charge per liter), is simply:

$$Q = [\equiv SOH_2^+] - [\equiv SO^-]$$

So that the surface charge is determined from:

$$Q = C_A - C_B - [H^+] + \frac{10^{-14}}{[H^+]}$$
 6.111

The surface charge *density*, σ , is calculated from Q as:

$$\sigma = \frac{Q\mathcal{F}}{A[\equiv S]} \qquad 6.112$$

where A is the specific surface area (m^2/mol) and $[\equiv S]$ is the concentration of solid (in mols/l)^{*}.

We can write equilibrium constant expressions for the surface protonation and deprotonation reactions. For example, for surface protonation:

 \equiv SOH + H⁺ $\rightleftharpoons \equiv$ SOH⁺₂

the equilibrium constant is:

$$\mathbf{K} = \frac{[\equiv \text{SOH}_2^+]}{[\equiv \text{SOH}][\text{H}^+]} \quad 6.113$$

We may write a conservation equation for the surface as:

$$\Sigma \equiv S = [\equiv SOH_2^+] + [\equiv SO^-] +$$

 $[\equiv SOH^0]$ 6.114

At pH below the pzpc, we can consider the entire surface charge as due to $[\equiv SOH_2^+]$, so that $Q \approx [\equiv SOH_2^+]$, and $[\equiv SO^-] \approx 0$. Combining equations 6.109, 6.111 and 6.112, we have:

Table 6.4. Point of Zero NetProton Charge of CommonSedimentary Particles

Material	pН
SiO ₂ (quartz)	2.0
SiO ₂ (gel)	1.0-2.5
α -Al ₂ O ₃	9.1
Al(OH) ₂ (gibbsite)	8.2
TiO_2 (anatase)	7.2
Fe_3O_4 (magnetite)	6.5
α -Fe ₂ O ₃ (hematite)	8.5
FeO(OH) (goethite)	7.8
$Fe_2O_3 \cdot nH_2O$	8.5
δ-MnO	2.8
β-MnO	7.2
Kaolinite	4.6
Montmorillonite	2.5
(100 0)	





Figure 6.34. (a) Surface charge of some common sedimentary materials as a function of pH. (b) Electrophoretic mobility, which is related to surface charge, of representative organic substances as a function of pH. The pH dependence of surface charge reflects the predominance of attached protons at low pH and the predominance of attached hydroxyls at higher pH. From Stumm (1992).

^{*} If the concentration of solid is expressed in kg/l, as it commonly is, then the specific surface area should be in units of m^2/kg .

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY



FeOOH as a function of pH for differ-

ent ionic strengths of a 1:1 electrolyte $(10^{-3} \text{ M FeOOH})$. From Dzombak and

Morel (1990).

$$\mathbf{K} = \frac{Q}{(\Sigma \equiv \mathbf{S} - Q)[\mathbf{H}^+]} \tag{6.115}$$

In equation 6.111, we see that if the amount of acid (or base) added is known, the surface charge can be determined by measuring pH (from which $[OH^-]$ may also be calculated). This is illustrated in Figure 6.36.

Thus the value of the protonation reaction equilibrium constant may be calculated from the surface charge and pH. The equilibrium constant for the deprotonation reaction may be obtained in a similar way. These equilibrium constants are also known as *surface acidity constants*, and sometimes denoted (as in Example 6.8) as K_{a1} and K_{a2} for the protonation and deprotonation reaction respectively.

6.6.2.2 Surface Potential and the Double Layer

The charge on a surface exerts a force on ions in the adja-

cent solution and gives rise to an electric potential, Ψ (measured in volts), which will in turn depend on the nature and distribution of ions in solution, as well as intervening water molecules. The surface charge results in an excess concentration of oppositely charged ions (and a deficit of liked charged ions) in the immediately adjacent solution.

The surface charge, σ , and potential, Ψ_0 , can be related by *Gouy-Chapman Theory*[‡], which is conceptually and formally similar to Debye–Hückel Theory (Chapter 3). The relationship between surface charge and the electric potential is:

$$\sigma = (8RT\varepsilon\varepsilon_0 I)^{1/2} \sinh\left(\frac{z\Psi_0 \mathcal{F}}{2RT}\right) \qquad 6.116$$

where *z* is the valence of a symmetrical background electrolyte (e.g., 1 for NaCl), Ψ_0 is the potential at the surface, F is the Faraday constant, *T* is temperature, R is the gas constant, *I* is ionic strength, ε_r is the dielectric constant of water and ε_0 is the permittivity of a vacuum (see Chapter 3). At 25°C, equ. 6.116 may be written as:

$$\sigma = \alpha I^{1/2} \sinh(\beta z \Psi_0)$$
 6.117

where α and β are constants with value of 0.1174 and 19.5 respectively.

Where the potential is small, the potential as a function of distance from the surface is:

$$\Psi(x) = \Psi_0 e^{-\kappa x} \qquad 6.118$$

where κ has units of inverse length and is called the Debye parameter or Debye length and is given by:



Figure 6.36. (a). Titration of a suspension of α -FeOOH (goethite) (6 g/liter) by HClO₄ and NaOH in the presence of 0.1 M NaClO₄. (b). Charge calculated by charge balance (equation 6.109) from the titration. From Stumm (1992).

[‡] Gouy-Chapman Theory assumes an infinite flat charge plane in one dimension. The electrostatic interaction between the surface and a cloud of charged particles is described by the Poisson-Boltzmann equation, as in Debye-Hückel Theory. Unlike Debye-Hückel, the Poisson-Boltzmann equation has an exact case. The theory was developed by Gouy and Chapman around 1910, a decade before Debye and Hückel developed their theory. See Morel and Herring (1993) for the details of the derivation.

CHAPTER O: AQUATIC CHEMISTRY

$$\kappa = \sqrt{\frac{2\mathcal{F}^2 I}{\varepsilon \varepsilon_0 RT}} \qquad 6.119$$

From equation 6.118, we see that the inverse of κ is the distance at which the electrostatic potential will decrease by 1/e. The variation in potential, the Debye length, and the excess concentration of counter-ions with distance from the surface is illustrated in Figure 6.37.

An addition simplification occurs where the potential is small, namely that equation 6.116 reduces to:

$$\sigma = \varepsilon \varepsilon_0 \kappa I^{2/3} \Psi_0$$

As is illustrated in Figure 6.37, an excess concentration of oppositely charged ions develops adjacent to the surface. Thus an electric double layer develops adjacent to the mineral surface. The inner layer, or Stern Layer, consists of charges fixed to the surface, the outer diffuse



Distance from Surface Figure 6.38. The double layer surrounding clay particles.



Figure 6.37. Variation in electrical potential and ions with distance from a negatively charged surface based on the Gouy-Chapman model. Electrical potential varies exponentially with distance, as do ion concentrations. $1/\kappa$ is the distance where the potential has decreased by 1/e. From Morel and Hering (1993).

layer, or Gouy Layer, consists of dissolved ions that retain some freedom of thermal movement. This is illustrated in Figure 6.38. The Stern Layer is sometimes further subdivided into an inner layer of specifically adsorbed ions (inner sphere complexes) and an outer layer of ions that retain their solvation shell (outer sphere complexes), called the inner and outer Helmholz planes respectively. Hydrogens adsorbed to the surface are generally considered to be part of the solid rather than the Stern Layer. The thickness of the Gouy (outer) Layer is considered to be the Debye length, $1/\kappa$. As is apparent in equation 6.119, this thickness will vary inversely with the square root of ionic strength. Thus the Gouy Layer will collapse in high ionic strength solutions and expand in low ionic strength ones.

> When clays are strongly compacted, the Gouy layers of individual particles overlap and ions are virtually excluded from the pore space. This results in retardation of diffusion of ions, but not of water. As a result, clays can act as semi-permeable membranes. Because some ions will diffuse easier than others, a chemical fractionation of the diffusing fluid can result.

> At low ionic strength, the charged layer surrounding a particle can be strong enough to repel similar particles with their associated Gouy layers. This will prevent particles from approaching closely and hence prevent coagulation. Instead, the particles form a relatively stable colloidal suspension. As the ionic strength of the solution increases, the Gouy layer is compressed and the repulsion between particles decreases. This allows particles to approach closely enough that they are bound

M. White

CHAPTER O: AQUATIC CHEMISTRY

together by attractive van der Waals forces between them. When this happens, they form larger aggregates and settle out of the solution. For this reason, clay particles suspended in river water will flocculate and settle out when river water mixes with seawater in an estuary.

6.6.2.3 Effect of the Surface Potential on Adsorption

The electrostatic forces also affect complexation reactions at the surface, as we noted at the beginning of this section.

before it can participate in with no surface potential. surface reactions. We can ac-



An ion must overcome the Figure 6.39. Surface speciation of hydrous ferric oxide for I = 0.1 Melectrostatic forces associated calculated in Example 6.11. Solid lines show speciation when surface with the electric double layer potential term is included, dashed lines show the calculated speciation

count for this effect by including it in the Gibbs Free Energy of reaction term, as in equation 6.95:

$$\Delta G_{ads} = \Delta G_{intr} + \Delta G_{coul} \tag{6.95}$$

where ΔG_{ads} is the total free energy of the adsorption reaction, ΔG_{int} is the *intrinsic* free energy of the reaction (i.e., the value the reaction would have in the absence of electrostatic forces; in general this will be similar to the free energy of the same reaction taking place in solution), and ΔG_{coul} is the free energy due to the electrostatic forces and is given by:

$$\Delta G_{\text{coul}} = \mathcal{F} \Delta Z \Psi_0 \tag{6.120}$$

where ΔZ is the change in molar charge of the surface species due to the adsorption reaction. For example, in the reaction:

 \equiv SOH+Pb²⁺ $\rightleftharpoons \equiv$ SOPb⁺ +H⁺

The value of ΔZ is +1 and $\Delta G_{coul} = \mathcal{F}\Psi$.

Thus if we can calculate ΔG_{coult} this raction Pb adsorbed term can be added to the intrinsic ΔG for the adsorption reaction (ΔG_{intr}) to obtain the effective value of ΔG (ΔG_{ads}). From ΔG_{ads} it is a simple and straightforward matter to calculate K_{ads}. From equation 3.86 we have:

$$\mathbf{K} = e^{-\Delta G_{ads} / RT}$$

Substituting equation 6.96, we have:

$$\mathbf{K} = e^{-\Delta G_{ads} / RT} e^{\Delta G_{coul} / RT} \quad 6.121$$

Since $K_{intr} = e^{\Delta G_{intr}/RT}$ and $\Delta G_{coul} = \mathcal{F} \Delta Z \Psi_{0}$ we have:

$$\mathbf{K} = K_{intr} e^{-\mathcal{F}\Delta Z \Psi_0 / RT} \qquad 6.122$$

Thus we need only find the value of Ψ_{0}



Figure 6.40. Comparison of calculated adsorption of Pb on hydrous ferric oxide with and without including the effect of surface potential.

Geochemistry

6.123

CHAPTER O: AQUATIC CHEMISTRY

Example 6.11. Effect of Surface Potential on Surface Speciation of Ferric Oxide

Using the surface acidity constants given in Example 6.10, calculate the surface speciation of hydrous ferric oxide as a function of pH in a solution with a background electrolyte concentration of I = 0.1 M. Assume the concentration of solid is 10^{-3} mol/l, the specific surface area is 5.4×10^4 m²/mol and that there are 0.2 mol of active sites per mole of solid.

Answer: The concentration of surface sites, Σ =Fe, is 0.2 mol sites/mol solid × 10⁻³ mol solid/l = 2 × 10⁻⁴ mol sites/l. Our conservation equation is:

$$\Sigma \equiv \text{Fe} = [\equiv \text{FeOH}_2^+] + [\equiv \text{FeOH}] + [\equiv \text{FeO}^-] = 2 \times 10^{-4}$$

$$P = e^{-\mathcal{F}\Delta Z \Psi_0 / RT}$$

so that our equilibrium constant expressions (6.98 and 6.99) become:

$$[\equiv \text{FeOH}] = \frac{[\equiv \text{FeOH}_2^+]}{[\text{H}^+]} \text{K}_{a1} P^{-1}$$
6.124

and

$$[\equiv \text{FeO}^{-}] = \frac{[\equiv \text{FeOH}]}{[\text{H}^{+}]} \text{K}_{a2} P^{-1} = \frac{[\equiv \text{FeOH}_{2}^{+}]}{[\text{H}^{+}]^{2}} \text{K}_{a1} \text{K}_{a2} P^{-2}$$
 6.125

Substituting into our conservation equation, and solving for $[=FeOH_2^+]$ we have:

$$[\equiv \text{FeOH}_{2}^{+}] = \Sigma \equiv Fe \left\{ 1 + \frac{K_{a1}P^{-1}}{[H^{+}]} + \frac{K_{a1}K_{a2}P^{-2}}{[H^{+}]^{2}} \right\}$$
 6.126

The concentration of surface charge, Q, is simply: $Q = [\equiv \text{FeOH}_2^+] - [\equiv \text{FeO}^-]$

and the surface charge density is: $\sigma = \frac{\mathcal{F}}{A[S]}([\equiv FeOH_2^+] - [\equiv FeO^-])$

(*P* enters the equations as the inverse because we have defined the equilibrium constants in Example 6.10 for the *desorption* reactions.) Substituting into the surface charge density equation, we have:

$$\sigma = \frac{\mathcal{F}}{A[S]} [= \text{FeOH}_{2}^{+}] \left\{ 1 - \frac{K_{a1}K_{a2}P^{-2}}{[H^{+}]^{2}} \right\}$$
6.127

Substitution equation 6.126 into 6.127, we have:

$$\sigma = \frac{\mathcal{F}}{A[S]} \Sigma \equiv Fe \left\{ 1 + \frac{K_{a1}P^{-1}}{[H^+]} + \frac{K_{a1}K_{a2}P^{-2}}{[H^+]^2} \right\}^{-1} \left\{ 1 - \frac{K_{a1}K_{a2}P^{-2}}{[H^+]^2} \right\}$$
6.128

Finally, substituting equation 6.117 for σ , and 6.123 for *P*, we have:

$$\sinh(\beta z \Psi_0) = \frac{\mathcal{F}}{A[S] \alpha I^{1/2}} \Sigma \equiv Fe \quad \frac{\left\{ 1 - \frac{K_{a1} K_{a2} e^{2T \Delta \Sigma \Psi_0 / RI}}{[H^+]^2} \right\}}{\left\{ 1 + \frac{K_{a1} e^{T \Delta \Sigma \Psi_0 / RI}}{[H^+]} + \frac{K_{a1} K_{a2} e^{2T \Delta \Sigma \Psi_0 / RI}}{[H^+]^2} \right\}}$$

$$6.129$$

A pretty intimidating equation, and one with no direct solution. It can, however, be solved by indirect methods (i.e, iteratively) on a computer. A quick an easy way is to use the Solver feature in Microsoft ExcelTM. Figure 6.39 shows the results and compares them to the surface speciation when surface potential is not considered. The effect of including the surface potential term is reduce the surface concentration of \equiv FeO⁻ and \equiv FeOH⁺₂ and broaden the pH region where \equiv FeO dominates.

which we can calculate from σ using equation 6.114. Example 6.10 illustrates the procedure.

The effect of surface potential on a given adsorbate will be to shift the adsorption curves to higher pH for cations and to lower pH for anions. Figure 6.40 illustrates the example of adsorption of Pb on hydrous ferric oxide. When surface potential is considered, adsorption of a given fraction of Pb occurs at

CHAPTER O: AQUATIC CHEMISTRY

roughly 1 pH unit higher than in the case where surface potential is not considered. In addition, the adsorption curves become steeper.

References and Suggestions for Further Reading

- Chou, L. and R. Wollast. 1985. Steady-state kinetics and dissolution mechanisms of albite. *Amer. J. Sci.* 285: 965-993.
- Drever, J. I., 1988. The Geochemistry of Natural Waters, Prentice Hall, Englewood Cliffs, 437 p.
- Garrels, R. M. and C. L. Christ. 1965. Solutions, Minerals and Equilibria. New York: Harper and Row.
- James, R. O. and T. W. Healy. 1972. Adsorption of hydrolyzable metal ions at the oxide-water interface. *J. Colloid Interface Sci.*, 40: 42-52.
- Morel, F. M. M. and J. G. Hering. 1993. *Principles and Applications of Aquatic Chemistry*. New York: John Wiley and Sons.
- Pankow, J. F. 1991. Aquatic Chemistry Concepts. Chelsea, MI: Lewis Publishers, 673 p.
- Richardson, S. M. and H. Y. McSween. 1988. *Geochemistry: Pathways and Processes*, New York: Prentice Hall.

Schlesinger, W. H. 1991. Biogeochemistry. San Diego: Academic Press.

- Sposito, G. 1989. The Chemistry of Soils. New York: Oxford University Press.
- Stumm, W. 1992. Chemistry of the Solid-Water Interface,. New York: Wiley Interscience 428 p.
- Stumm, W. and J. J. Morgan. 1995. Aquatic Chemistry, New York: Wiley and Sons.
- White, A. F. and S. L. Brantley (ed.), 1995, *Chemical Weathering Rates in Silicate Minerals*, Washington: Mineral. Soc. Am.

Problems

1. Make a plot similar to Figure 6.1, but for water in equilibrium with atmospheric CO_2 ($P_{CO2} = 10^{-3.43}$). Assume ideality and that there are no other species present in solution but those shown on the graph. What is the pH of the CO_2 equivalence point in this case? What is the pH of the bicarbonate equivalence point?

- 2. Using the composition given in Problem 3.4, calculate the alkalinity of seawater at 25°C.
- 3. For a sodium carbonate solution titrated with HCl to the bicarbonate equivalence point, show that:

$$\Sigma CO_2 \cong [Cl^-] + [OH^-]$$

4. Calculate the pH of a solution containing $\Sigma CO_2 = 10^{-2}$ at 25°C at the bicarbonate and carbonate equivalence points. Assume ideality and use the equilibrium constants in Table 6.1.

- 5. Consider a 0.005M solution of Na₂CO₃ at 25°C. Assuming ideality and that the system is closed:
 - a. What is the pH of this solution?
 - b. What is the pH of this solution when titrated to the bicarbonate equivalence point?
 - c. What is the pH of this solution when titrated to the CO₂ equivalence point?
- 6. Consider a 0.01M solution of NaHCO₃ (sodium bicarbonate) at 20°C. Assuming ideality:
 - a. What is the pH of this solution?
 - b. Plot the titration curve for this solution (i.e., moles of HCl added vs. pH).
 - c. What is the pH of the CO₂ equivalence point of this solution?

7. Explain why pH changes rapidly near the bicarbonate and CO₂ equivalence points during titration.

8. Mars probably once had a more substantial atmosphere and water on its surface. Suppose that it had a surface atmospheric pressure of 1 bar (0.1 MPa) and that the partial pressure of CO_2 was the same as it is today, 6×10^3 . Further suppose the surface temperature was 5° C. Assume ideal behavior and use the equilibrium constant is Table 6.1 for this problem. Under these conditions, at what concentra-

CHAPTER O: AQUATIC CHEMISTRY

tion of Ca^{2+} ion would an ancient Martian stream become saturated with $CaCO_3$? What would the pH of that stream be? Assume that calcium, carbonate species, and the dissociate products of water are the only ions present.

9. Calculate the buffer capacity of a solution initially in equilibrium with calcite for pH between 6 and 9 at 25°C.

10. Calculate the calcium ion concentration for a solution in equilibrium with calcite and fixed ΣCO_2 of 10^{-2} M at 25°C.

11. Show that the α , fraction of copper complexed as CuOH⁺, as defined in equation 6.50, will decrease with increasing concentration of total copper in solution. Assume ideal behavior and that H⁺, OH⁻, Cu²⁺ and CuOH⁺ are the only ions present in solution and that the stability constant of CuOH⁺ is 10⁻⁸.

12. Using the following equilibrium constants and reactions, make a plot of Zn^{2+} , $ZnOH^+$, $Zn(OH)_2$, $Zn(OH)_3^-$, $Zn(OH)_4^{2-}$, and total zinc concentration as a function of pH from pH 1 to pH 14. Assume ideal behavior and that H⁺, OH⁻, and various species of Zn are the only ions in solution. *Hint: Use a log scale for the Zn concentrations.*

$ZnO + 2H^+ \rightleftharpoons Zn^{2+} + H_2O$	$\log K_{zn} = 11.2$
$Zn^{2+} + H_2O \rightleftharpoons ZnOH^+ + H^+$	$\log K_1 = -9$
$Zn^{2+} + 2H_2O \rightleftharpoons Zn(OH)_2 + 2H^+$	$\log K_1 = -16.9$
$\operatorname{Zn}^{2+} + 3\operatorname{H}_2\operatorname{O} \rightleftharpoons \operatorname{Zn}(\operatorname{OH})_3^- + 3\operatorname{H}^+$	$\log K_1 = -28.1$
$\operatorname{Zn}^{2+} + 4\operatorname{H}_2\operatorname{O} \rightleftharpoons \operatorname{Zn}(\operatorname{OH})^{2-}_4 + 4\operatorname{H}^+$	$\log K_1 = -40.2$

13. Using the reactions and stability constants given by equations 6.60-6.62, derive the equations used to construct the stability diagram shown in Figure 6.13b. Assume a fixed Mg^{2+} concentration of 10^{-7} M, that the solution is ideal, and that the solids are pure phases.

14. For the adsorption of Zn^{2+} on hydrous ferric oxide:

$$\equiv FeOH^0 + Zn^{2+} \rightleftharpoons \equiv FeOZn^+ + H^+$$

the apparent equilibrium constant is $10^{0.99}$. For this problem, use the surface acidity constants (i.e., equilibrium constants for adsorption and desorption of H⁺) given in equations 6.98 and 6.99 (Example 6.10).

- a. Make a plot of Θ_{Zv} (fraction of sites occupied by Zn) vs. the aqueous concentration of Zn^{2+} (use log [Zn²⁺]) at pH 7 and a total concentration of surface sites of 10⁻³ M. Assume that Zn²⁺ forms no complexes in solution.
- b. Ignoring electrostatic effects and any aqueous complexation of Zn²⁺, make a plot of the fraction of Zn²⁺ adsorbed as a function of pH (from pH 2 to pH 6), assuming a total Zn²⁺ concentration of 10⁻⁸M.
- c. Do the same calculation as in b, but take into consideration the aqueous complexation reactions and equilibrium constants in Problem 11.

15. For the adsorption of Pb^{2+} on aluminum oxide:

$$\equiv$$
SOPb⁺ $\rightleftharpoons \equiv$ SO⁻ + Pb²⁺

the apparent equilibrium constant, K_{ad} , is $10^{-6.1}$. In addition, consider the reaction:

$$\equiv SOH_{2}^{+} \rightleftharpoons \equiv SOH + H^{+} \qquad \qquad K_{a1} = 10^{-6}$$
$$\equiv SOH \rightleftharpoons \equiv SO^{-} + H^{+} \qquad \qquad K_{a2} = 10^{-7.7}$$

Geochemistry

CHAPTER O: AQUATIC CHEMISTRY

Make a plot of the fraction of Pb adsorbed as a function of pH from pH=4 to pH = 7, assuming a total concentration of alumina of 10^{-2} M, a surface site density of 1×10^{-2} moles/mole Al₂O₃, and total Pb concentration of 10^{-9} M, ignoring electrostatic effects and any complexation in solution. (*Hint* Θ_{Pb} , the fraction of sites occupied by Pb will be negligible).

16. Consider a 10^{-3} M suspension of aluminum oxide (Al₂O₃) in a 1:1 electrolyte having a specific surface area of 500 m²/g with 4 x 10^{-6} moles of active surface sites per m². At pH 7, the surface charge, Q, is found to be 7.93 moles/l. What is the surface charge density, σ ? If the temperature is 25°C and ionic strength, I, is 10^{-3} M, what is the surface potential, Ψ_0 ? What is the potential at a distance of 1 Debye length from the surface? Make a plot of how Ψ_0 and the Debye length change as I varies from 10^{-3} to 1 (equivalent to going from river water to seawater).

CHAPTER 7: TRACE ELEMENTS

CHAPTER 7: TRACE ELEMENTS IN IGNEOUS PROCESSES

7.1 INTRODUCTION

rn this chapter we will consider the behavior of trace elements, particularly in magmas, and introduce methods to model this behavior. Though trace elements, by definition, constitute only a Lesmall fraction of a system of interest, they provide geochemical and geological information out of proportion to their abundance. There are several reasons for this. First, variations in the concentrations of many trace elements are much larger than variations in the concentrations of major components, often by many orders of magnitude. Second, in any system there are far more trace elements than major elements. In most geochemical systems, there are 10 or fewer major components that together account for 99% or more of the system. This leaves 80 trace elements. Each element has chemical properties that are to some degree unique, hence there is unique geochemical information contained in the variation of concentration for each element. Thus the 80 trace elements always contain information not available from the variations in the concentrations of major elements. Third, the range in behavior of trace elements is large and collectively they are sensitive to processes to which major elements are insensitive. One example is the depth at which partial melting occurs in the mantle. When the mantle melts, it produces melts whose composition is only weakly dependent on pressure, i.e., it always produces basalt. Certain trace elements, however, are highly sensitive to the depth of melting (because the phase assemblages are functions of pressure). Furthermore, on a large scale, the composition of the Earth's mantle appears to be relatively uniform, or at least those parts of it that give rise to basaltic magmas. Indeed, it has proved very difficult to demonstrate any heterogeneity in the mantle on the basis of the major element chemistry of the magmas it has produced. In contrast, it has been amply demonstrated that trace element concentrations of the mantle are quite variable. Trace elements, particularly when combined with isotope ratios, which we shall discuss the in next chapter, thus provide a chemical fingerprint of different mantle reservoirs. Finally, the behavior of trace elements is almost always simpler than that of major elements because trace elements obey Henry's Law (Chapter 3).

Trace element geochemistry has been of enormous use in understanding the evolution of the Earth. As we shall see in subsequent chapters, a fair amount of what we know about the evolution of the core, the mantle, and the crust has come from the study of trace element abundances. For example, the abundance of certain "siderophile" (a term we shall define shortly) trace elements in the mantle and mantle-derived rocks, provides reason to believe that segregation of the Earth's iron-nickel core must have been largely complete before the Earth had entirely accreted from the cloud of gas and dust surrounding the early Sun (the solar nebula). We also know, for example, that much of the upper mantle has undergone partial melting at some point in the past. These partial melts of the upper mantle have, through time, created the continental crust. From the abundances of trace gases in the mantle and their isotopic composition, we conclude that the solid portion of the Earth must have undergone extensive outgassing within the first few hundred million years of Earth history. As we shall see in subsequent chapters, magmas from a given tectonic setting tend to share patterns of trace element abundances. This allows the tectonic setting of anciently erupted magmas to be deduced.

Though our focus here will be on igneous processes, trace elements are equally useful in other geologic problems as well. For example, trace elements can provide useful clues as to the origin of sulfide ore deposits. The concentrations of trace elements such as cadmium in the fossil shells of micro-organisms provide information about the biological productivity and circulation patterns of ancient oceans, and the concentration of Sr in corals provides a measure of temperature of ancient seas. Indeed, throughout the earth sciences, trace element geochemistry has become a powerful tool.

Our purpose in this chapter is to add this tool to our geochemical toolbox. We will begin by considering the chemical properties of the various groups of trace elements, with particular emphasis on how they behave in nature. We then introduce quantitative means of describing trace element distribution. Our primary tool here will be the distribution, or partition, coefficient, which we first introduced in

Geochemistry

CHAPTER 7: TRACE ELEMENTS

Chapter 3 (as K_D). We will examine how the distribution coefficient depends on temperature, pressure, composition, and the fundamental chemical properties of the element of interest. Finally, we develop equations to predict the behavior of trace elements during melting and crystallization. The knowledge of trace element behavior that we gain in this chapter will be useful in the following one where we discuss radiogenic isotope geochemistry, because all the radioactive elements and their daughter products, with a single exception, are trace elements. We will apply the tools we acquire here to understanding the evolution of the core, the mantle, and the crust in subsequent chapters.

7.1.1 WHAT IS A TRACE ELEMENT?

The term trace element is a bit hard to define. For igneous and metamorphic systems (and sedimentary rocks for that matter), an operational definition might be as follows: trace elements are *those elements that are not stoichiometric constituents of phases in the system of interest*. Clearly this definition is a bit fuzzy: a trace element in one system is not one in another. For example, potassium never forms its own phase in mid-ocean ridge basalts (MORB), its concentration rarely exceeding 1500 ppm; but K is certainly not a trace element in granites. For most silicate rocks, O, Si, Al, Na, Mg, Ca, and Fe are 'major elements'. H, C, S, K, P, Ti, Cr, and Mn are sometimes 'major elements' in the sense that they can be stoichiometric constituents of phases. These are often referred to as 'minor elements'. All the remaining elements are always trace elements, with the exception of a few rare, but important, circumstances such as pegmatites and ore deposits.

The above definition breaks down entirely for fluids and natural waters since there is only one phase, namely the fluid, and it's not stoichiometric. In seawater, anything other than Cl^{-} , SO_{4}^{2-} , CO_{3}^{2-} ,



Figure 7.1. Three-dimensional histogram illustrating the abundance of the elements (as the log of mole fraction) in the silicate portion of the Earth (the "Bulk Silicate Earth"; BSE). Just 6 elements, oxygen, magnesium, silicon, iron, aluminum, and calcium make up 99.1% of the silicate Earth. If we include the core and consider the composition of the entire Earth, then only nickel, and perhaps sulfur, need be added to this list. The remaining elements, though sometimes locally concentrated (e.g., in the crust, in the hydrosphere, in ores) can be considered *trace elements*.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

 HCO_3^- , Mg^{2+} , Ca^{2+} , K^+ and Na^+ (and H_2O , of course) can be considered a trace constituent, though Sr^{2+} , HBO_3^- , and Br^- are sometimes considered major constituents also (constituents or species is a better term here than elements). These, including the last three, constitute over 99.99% of the total dissolved solids in seawater. Trace elements in seawater and in rocks do have one thing in common: *neither affect the chemical or physical properties of the system as a whole to a significant extent*. This might serve as a definition. However, trace (or at least minor) elements can determine the color of a mineral (e.g., the green color of chrome diopside), so even this definition has problems. And CO_2 , with a concentration in the atmosphere of only 360 ppm, profoundly affects the transparency of the atmosphere to infrared radiation, and, as a result, Earth's climate. At even lower concentrations, ozone in the upper atmosphere controls the atmospheric transparency to ultraviolet radiation. So this definition is not satisfactory either.

Yet another possible definition of a trace element is: *an element whose activity obeys Henry's Law in the system of interest.* This implies sufficiently dilute concentrations that for trace element A and major component B, A-A interactions are not significant compared to A-B interactions.

There is perhaps no satisfactory quantitative definition of a trace element that will work in every situation. For our present purposes, any of these definitions might do, but bear in mind that a trace element in one system need not be a trace element in another.

7.2 BEHAVIOR OF THE ELEMENTS

7.2.1 Goldschmidt's Classification

No matter how we define the term "trace element", most elements will fall into this category, as is illustrated in Figure 7.1. That being the case, this is a good place to consider the geochemical characteristics of the elements. Goldschmidt* recognized four broad categories: atmophile, lithophile, chalcophile, and siderophile (Figure 7.2, Table 7.1). *Atmophile* elements are generally extremely volatile (i.e., they form gases or liquids at the surface of the Earth) and are concentrated in the atmosphere and hydrosphere. Lithophile, siderophile and chalcophile refer to the tendency of the element to partition into a silicate, metal, or sulfide *liquid* respectively. *Lithophile* elements are those showing an affinity for silicate phases and are concentrated in the silicate portion (crust and mantle) of the earth. *Siderophile* elements have an affinity for a metallic liquid phase. They are depleted in the silicate portion of the earth and presumably concentrated in the core. *Chalcophile* elements have an affinity for a sulfide liquid phase. They are also depleted in the silicate earth and may be concentrated in the core. Many sulfide ore deposits originated from aqueous fluids rather than sulfide liquid. A chalcophile element need not necessarily be concentrated in such deposits. As it works out, however, they generally are. Most elements that are siderophile are usually also somewhat chalcophile and visa versa.

There is some basis for Goldschmidt's classification in the chemistry of the elements. Figure 7.2 shows that the lithophile elements occur mainly at either end of the periodic table, siderophile elements are mainly group 8, 9 & 10 elements (and their neighbors), chalcophile elements are mainly group 11, 12 and the heavier group 13-16 elements, while the atmophile elements are mainly the noble gases. The distribution of the electropositive elements (those that give up an electron more readily than they accept one) among metal, sulfide, and silicate phases is controlled by the free energies of formation of the corresponding sulfides and silicates. By comparing the free energies of formation with those of ferrous

^{*} Victor Goldschmidt (1888-1947) is often considered the 'father of geochemistry'. Goldschmidt earned a Ph.D. from the University of Oslo in 1911 and remained there until 1929, when he assumed the directorship of the Geochemisches Institut at the University of Göttingen. Because of the worsening political situation in Germany, he returned to Oslo in 1935. He was for a time imprisoned in a concentration camp after Germany invaded Norway in 1940. In 1942 he fled to Sweden and eventually to England. He returned to Oslo in 1946 but never fully recovered from the effects of imprisonment and died a year later. The Geochemical Society has named its most prestigious medal after him and co-sponsors, along with the European Association of Geochemistry, annual Goldschmidt Conferences.

CHAPTER 7: TRACE ELEMENTS

Table 7.1. Goldschmidt's Classification of the Elements							
Siderophile	Chalcophile	Lithophile	Atmophile				
Fe*, Co*, Ni*	(Cu), Ag	Li, Na, K, Rb, Cs	(H), N, (O)				
Ru, Rh, PdZn, Cd, Hg	Be, Mg, Ca, Sr, Ba	He, Ne, Ar, Kr, Xe					
Os, Ir, Pt	Ga, In, Tl	B, Al, Sc, Y, REE					
Au, Re [†] , Mo [†]	(Ge), (Sn), Pb	Si, Ti, Zr, Hf, Th					
Ge*, Sn*, W [‡]	(As), (Sb), Bi	P, V, Nb, Ta					
C [‡] , Cu*, Ga*	S, Se, Te	O, Cr, U					
Ge*, As [†] , Sb [†]	(Fe), Mo, (Os)	H, F, Cl, Br, I					
	(Ru), (Rh), (Pd)	(Fe), Mn, (Zn), (Ga)					

*Chalcophile and lithophile in the earth's crust

[†]Chalcophile in the earth's crust

[‡]Lithophile in the earth's crust

sulfide and ferrous silicate, it is possible to deduce which elements are siderophile, those which are chalcophile and which are lithophile. For historical reasons, namely lack of ΔG_f^o data on silicates, the point is generally illustrated using the enthalpy of formation, ΔH_{tr} of the oxide. Since 'oxyphile' could arguably be a better term than lithophile, this is not such a bad thing. Table 7.2 gives some examples. Elements whose oxides have high $-\Delta G_f$ are lithophile. Why this is the case should be clear from our understanding of thermodynamics. States with the lowest free energy are the most stable: a high $-\Delta G_f$ indicates the oxide is much more stable than the metal. Elements whose oxides have $-\Delta G_f$ similar to that FeO combine with oxygen only about as readily as Fe, and are generally siderophile. Those elements whose oxides have low $|\Delta G^{\circ}|$ are generally chalcophile.

Lithophile elements also have either very low electronegativities or very high ones and tend to form ionic bonds (although the basic silicate bond, the Si—O bond, is only about 50% ionic, metal-oxygen bonds in silicates are dominantly ionic). Siderophile and chalcophile elements have intermediate electronegativities and tend to form covalent or metallic bonds.

7.2.2 The Geochemical Periodic Table

Goldschmidt's classification is relevant mainly to distribution of elements in meteorites and to how elements distribute themselves between the Earth's major geochemical reservoirs: the core, the mantle and crust, and the hydrosphere and atmosphere. Since there is an overabundance of O in the outer part of the Earth, metallic liquids do not form, and siderophile elements have little opportunity to behave as such. Similarly, sufficient S is rarely available to form more than trace amount of sulfides. As a result, siderophile elements such as Ni and chalcophile elements such as Pb occur mainly in silicate phases in the crust and mantle.

We can, however, group the elements based on how they behave in the silicate portion of the Earth, the mantle and crust. Figure 7.3 illustrates this grouping. We first note that we have added sodium to those 6 elements whose molar abundance exceeds 1 percent (Figure 1), to form the group called major elements, and which we will not discuss in this chapter. Although the elements K, Ti, Mn, and P are often reported in rock analyses as major elements, we will include them in our discussion of trace elements. Let's now briefly examine the characteristics of the remaining groups.

CHAPTER 7: TRACE ELEMENTS

	GRO	UP																
	IA	-															VI	IIA
1	<i>H H H H H H H H H H</i>								1	IIA I	VA V	YA V	IA VI	IA J	He			
2	Lí	Ве											В	C 1	N ()	F 1	Ne
3	Na	Мg	111B	IVB	VB	VIB	VIIB	— <i>v</i> 1	111B -		1B	11B -	Al S	5 <i>i</i> 1			Cl A	٩r
4	K	Са	Sc	Τí	ν	Cr	Mn	Fe	Co	Ní	Cu	Zn	Ga (ie 🖌	As .	Se	Br I	Kr
5	Rb	Sr	Υ	Zr	Nb	Мо	Τc	Ru	Rh	Pd	Ag	Cđ	In	ŝn S	ib I	te	1)	хe
6	Cs	Ва	La	Ъf	Ta	W	Re	Os -	Ir	Pt	Аи	Hg	π.	Pb I	Bi 1	20 1	At R	Rn
7	Εr	Ra	Ac		N,													
	_	•	``	L	a C	e P	r N	d Pm	Sm	Еи	Gd	тв	Dy	Но	Er	Тт	Yb	Lu
			``		c Tl	l P	a L	I Nı	l Pu									
			. 1	1.4	Г		. 1	1.4	Z				. Г				a	
		1	ithop	hile			Siden	ophile			Cha	lcophi	le		Atr	noph	ile	

Figure 7.2. Goldschmidt's classification of the elements.

7.2.2.1 The Volatile Elements

The defining feature of the noble gases is their filled outer electron shell, making them chemically inert as well as volatile. Hence, they are never chemically bound in rocks and minerals. Furthermore, except for He, they are have rather large radii and cannot easily be accommodated in either cationic or anionic lattice sites of many minerals. Thus they are typically present at very low concentrations. Their concentrations are usually reported in STP cm³/g at (i.e., cm³/g at standard temperature and pressure: 273 K and 01.MPa; 1 cm³/g =4.46 × 10⁻⁵ moles/g). Concentrations in silicate rocks and minerals are

Oxide	$-\Delta G_{f}^{o}$ (kJ/mole/oxygen)	Oxide $-\Delta G_{f}^{o}$ (kJ/mole/oxygen)
CaO	604.0	In ₂ O ₂ 304.2
ThO ₂	584.6	SnO_2 260.0
MgÓ	569.4	FeO 245.9
Al_2O_3	527.3	WO ₃ 247.3
ZrO_2	521.6	CdŐ 221.9
CeO_2	512.6	NiO 211.6
TiO_2	444.7	MoO ₃ 215.4
SiO_2	428.1	Sb ₂ O ₃ 207.9
Na ₂ O	398.3	PbO 189.3
Ta ₂ O ₃	374.0	As ₂ O ₃ 180.1
MnO	362.8	Bi_2O_3 168.8
Cr_2O_3	353.1	CuO 127.6
ZnO	318.4	Ag_2O_3 10.9

TABLE 7.2. FREE ENERGY OF FORMATION OF SOME OXIDES

CHAPTER 7: TRACE ELEMENTS



Figure 7.3. The Geochemical Periodic Table, in which elements are grouped according to their geochemical behavior.

typically 10^{-4} to 10^{-12} STP cm³/g (10^{-1} to 10^{-9} ppm). Their solubility in silicate melts is a strong function of pressure, as well as both atomic radius and melt composition as is illustrated in Figure 7.4. Although



Figure 7.4. Solubility of the rare gases is melts of varying composition at 1200° to 1400° C. Solubility is a strong function of atomic radius and melt composition, but only a weak function of temperature. From Carroll and Draper (1994).

they cannot form true chemical bonds with other atoms, they can be strongly adsorbed to crystal surfaces through van der Waals forces.

The very strong nature of the N-N bond makes nitrogen relatively unreactive once molecular nitrogen forms; consequently it, like the rare gases, is strongly partitioned into the atmosphere. However, it is quite capable of forming strong covalent bonds with other elements. In silicate minerals, N is probably primarily present as the ammonia ion rather than N₂. As such, it readily substitutes for K⁺. As ammonia, it is highly soluble in aqueous fluids and is therefore readily transported by them. Ammonia, like N₂, is quite volatile, so both species partition readily into the gas phase of magmas. In aqueous solution, nitrogen will be present as nitrate (and trace amounts of ammonia, produced by breakdown of nitrogen-bearing organic compounds), as well as N₂. Nitrogen is a component of proteins and nucleic acids and as such is vital to all organisms. However, most plants can utilize only "fixed" nitrogen, that is nitrate or ammonia. In many natural waters, nitrate concentrations are held at very low concentrations because of biological utilization.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

7.2.2.2 The Semi-Volatiles

The shared characteristic of this group is that they partition readily into a fluid or gas phase (e.g., Cl, Br) or form compounds that are volatile (e.g., SO_2 , CO_2). Not all are volatile in a strict sense (volatile in a strict sense means having a high vapor pressure or low boiling point; indeed, carbon is highly refractory in the elemental form).

The partitioning of sulfur between liquid and gas phases is a strong function of f_{O_2} . At high oxygen fugacities, sulfur is present primarily as SO₂, but at low f_{O_2} it is present primarily as sulfide. The solubility of sulfide in silicate liquids is, however, low. At sufficiently high sulfur concentrations in magmas, sulfide and silicate liquids will exsolve. Sulfide liquids are rich in Fe and Ni and other chalcophile metals and are the source of many economically important ore deposits. Large volumes of sulfide liquid are rare, but microscopic droplets of sulfide liquids commonly occur in mid-ocean ridge magmas.

Similarly, the solubility of CO_2 in silicate magmas is limited and is a strong function of pressure. At low CO_2 concentrations, CO_2 exsolves from magmas to form a CO_2 –H₂O gas phase. However, at higher CO_2/H_2O ratios and total CO_2 concentrations, carbonatite magmas can form in which $CaCO_3$ is the dominant component. On the whole, carbonatites are rare, but over the course of geologic history they have erupted on every continent. In certain localities, such as the modern East Africa Rift, they can be fairly common.

The remaining elements in this group are always present in trace concentrations and never reach saturation in magmas and hence never exsolve as independent gas or fluid phases. Rather, they partition into gas phase formed by exsolution of CO_2 and H_2O .

7.2.2.3 The Alkali and Alkaline Earth Elements

The alkali and alkaline earth elements have electronegativities less than 1.5 and a single valence state (+1 for the alkalis, +2 for the alkaline earths). The difference in electronegativity between these elements and most anions is 2 or greater, so the bonds these elements form are strongly ionic in character (Be is an exception, as it forms bonds with a more covalent character). Ionic bonds are readily disrupted by water due to its polar nature (see Chapter 3). The low ionic potential (ratio of charge to ionic radius) makes these elements relatively soluble in aqueous solution. Because of their solubility, they are quite mobile during metamorphism and weathering.

Because bonding is predominantly ionic, the atoms of these elements behave approximately as hard spheres containing a fixed point charge at their centers (these are among the group A or hard ions discussed in Chapter 6). Thus the factors that most govern their behavior in igneous rocks are *ionic radius and charge*. K, Rb, Cs, Sr, and Ba, are often collectively termed the *large-ion-lithophile* (*LIL*) *elements*. As the name implies, these elements all have large ionic radii, ranging from 118 picometers (pm) for Sr to 167 pm for Cs. The major minerals in basaltic and ultramafic rocks have two kinds of cationic lattice sites: small tetrahedral sites occupied by Si and Al (and less often by Fe³⁺ and Ti⁴⁺) and larger octahedral ones usually occupied by Ca, Mg, or Fe and more rarely by Na. The ionic radii of the heavy alkali and alkaline earth elements are larger than the radii of even the larger octahedral sites. As a result, substitution of these elements thus tend to be concentrated in the melt phase when melting or crystallization occurs. Such elements are called *incompatible elements*. Partial melting of the mantle and eruption or intrusion of these melts on or in the crust over the history of the Earth has enriched the crust in these incompatible elements.

In contrast to the heavy alkaline earths, Be has an ionic radius smaller than most octahedral sites. Substitution of a small ion in a large site is also energetically unfavorable as the bond energy is reduced. Thus Be is also an incompatible element, though only moderately so. While Li has an ionic radius similar to that of Mg and Fe^{2+} , its substitution for one of these elements creates a charge imbalance that requires a coupled substitution. This is also energetically unfavorable, hence Li is also an incompatible element, though again only moderately so.

M. White

Geochemistry

CHAPTER 7: TRACE ELEMENTS

7.2.2.4 The Rare Earth Elements and Y

The rare earths are the two rows of elements commonly shown at the bottom of the periodic table. The first row is the *lanthanide* rare earths, the second is the *actinide* rare earths. However, the term "rare earths" is often used in geochemistry to refer to only to the lanthanide rare earths. We will follow that practice in this book, though we will discuss both the actinide and lanthanides in this section. Only two of the actinides, U and Th, have nuclei stable enough to survive over the history of the Earth. Y shares the same chemical properties, including charge and ionic radius, as the heavier rare earths, and as a result behaves much like them.

As the alkalis and alkaline earths, the rare earths and Y are strongly electropositive; the lathanide have electronegativities of 1.2 or less, the actinides U and Th have slightly higher electronegativities. As a result, they form predominantly ionic bonds, and the hard charged sphere again provides a good model of their behavior. The lanthanide rare earths are in the +3 valence state over a wide range of oxygen fugacities. At the oxygen fugacity of the Earth's surface, however, Ce can be partly or wholly in the +4 state and Eu can be partly in the +2 state at the low oxygen fugacities of the Earth's interior. Th is always in a +4 valence state, but U may be in a +4 or +6 valence state, depending on oxygen fugacity (or $p\epsilon$, if we chose to quantify the redox state that way). Unlike the alkali and alkaline earth elements, they are relatively insoluble in aqueous solutions, a consequence of their higher charge and high ionic potential and resulting need to be coordinated by anions. The one exception is U in its fully oxidized ${
m U}^{
m 6+}$ form, which forms a soluble oxyanion complex, ${
m UO}^{
m 2-}_2$.

The rare earths are transition metals. In the transition metals, the *s* orbital of the outermost shell is filled before filling of lower electron shells is complete. In atoms of the period 6 transition elements, the 6s orbital is filled before the 5d and 4f orbitals. In the lanthanide rare earths, it is the 4f orbitals that are being filled, so the configuration of the valence electrons is similar in all the rare earth, hence all exhibit similar chemical behavior. Ionic radius, which decreases progressively from La³⁺ (115 pm) to Lu³⁺ (93 pm), illustrated in Figure 7.5, is thus the characteristic that governs their relative behavior.

Because of their high charge and large radii, the rare earths are incompatible elements. The degree of incompatibilities varies, however. Highly charged U and Th are highly incompatible elements, as are the lightest rare earths. However, the heavy rare earths have sufficiently small radii that they can be accommodated to some degree in many common minerals. The heaviest rare earths readily substitute for Al^{3+} in garnet, and hence can be concentrated by it. Eu, when in its 2+ state, substitutes for Ca^{2+} in plagioclase feldspar more readily than the other rare earths. Thus plagioclase is often anomalously rich in Eu compared to the other rare earths, and other phases in equilibrium with plagioclase become relatively depleted in Eu as a consequence.

The systematic variation in lathanide rare earth behavior is best illustrated by plotting the log of the relative abundances as a function of atomic number (this sort of plot is sometimes called a Masuda, Masuda-Corvell, or Corvell plot, but most often is simply termed a rare earth plot or diagram). Relative abundances are calculated by dividing the concentration of each rare earth by its concentration in a set of normalizing values, such as the concentrations of rare earths in chondritic meteorites. Why do we use relative abundances? As we shall see in Chapter 10, the abundances of even numbered elements in the solar system (and most likely the cosmos) are greater than those of neighboring odd-



La Ce Pr Nd Pm Sm Eu Gd Tb Dy Ho Er Tm Yb Lu

Figure 7.5. Ionic radii of the lanthanide rare earth elements (3+ state except where noted). Promethium (Pm) has no isotope with a half-life longer than 5 years.

Geochemistry

CHAPTER 7: TRACE ELEMENTS



Figure 7.6. Concentrations of the rare earths in the carbonaceous chondritic meteorite Orgueil.



Figure 7.7. A rare earth plot showing rare earth patterns for average upper continental crust (Rudnick and Fountain, 1995), average shale, average mid-ocean ridge basalt, and the meteorite Orgueil. As a carbonaceous chondrite, Orgueil has lower rare earth concentrations than the average of ordinary chondrites used for normalization.

earth-depleted pattern; upper continental crust is *light rare earth-enriched* with a negative 'Eu anomaly'. The light rare earth depletion of MORB reflects the incompatible element-depleted nature of the upper mantle from which these magmas are derived. This incompatible element depletion of the mantle is generally thought to have resulted from extraction of partial melts, in which the incompatible elements were concentrated. Those partial melts have crystallized to form the continental crust. If this is so, the complimentary nature of the rare earth patterns of MORB and continental crust is not coincidental. There are good reasons, which we will discuss in Chapters 10 and 11, to believe that the relative abundances of rare earths in the Earth as a whole are similar to those of chondrites, i.e., the rare earth pattern

numbered elements. Furthermore, because of the way the elements have been created, abundances generally decrease with increasing atomic number. Thus a simple plot of abundances produces a saw tooth pattern of decreasing abundances. This can be seen in Figure 7.6, which shows rare earth abundances in the CI chondrite Orgueil (CI chondrites are a class of meteorites that are taken to be the best representative of the average concentrations of non-volatile elements in the solar system; see Chapter 10). "Normalizing" the rare earth abundances to those of chondritic meteorites eliminates effects related to nuclear stability and nucleosynthesis and produces a smooth pattern, such as those seen in Figure 7.7.

Though all igneous geochemists normalize rare earth abundances to some set of chondritic values, there is no uniformity in the set chosen. Popular normalization schemes include the CI chondrite Orgueil, an average of 20 ordinary chondrites reported by Nakamura (1974), and the chondritic meteorite Leedy (Masuda et al., 1973). Although the absolute values of the normalizing values vary (for example, the Nakamura values are about 28% higher than those of Orgueil), the relative abundances are essentially the same. Thus the normalized rare earth pattern should be the same regardless of normalizing values. Some sets of normalizing values are listed in Table 7.3. A more complete tabulation can be found in Rollinson (1993).

Rare earth patterns for upper continental crust and average mid-ocean ridge basalt (MORB) are also shown in Figure 7.7. MORB exhibits a *light rare*

M. White

CHAPTER 7: TRACE ELEMENTS

of the Earth should be flat. Table 7.3 NORMALIZING VALUES USED FOR RADE FADTH PLOTS Mass balance therefore requires the sum of all the various rare earth reservoirs in the Earth have a flat rare earth pattern. If we assume the continental crust and the mantle are the only two reservoirs with significant concentrations of rare earth elements, and if the continental crust is light rare earth-enriched, then the mantle should be light rare earth-depleted.

A negative Eu anomaly is typical of many continental rocks, as well as most sediments and seawater. The Eu anomaly probably arises because many crustal rocks of granitic and granodioritic composition were produced by intracrustal partial melting. The residues of those melts were rich in plagioclase, hence retaining somewhat more of the Eu in the lower crust, and creating a complimentary Eudepleted upper crust. Sediments and seawater inherit this Eu anomaly from their source rocks in the upper continental crust.

Many sedimentary rocks, and seawater, have rare earth patterns that are similar to each other, and to that of the continental crust. To accentuate the difference in rare earth patterns between sediments, low temperature geochemists often normalize rare earth abundances to the concentrations in average shale. Again, there are

TADU	\mathbf{E} 1 . 7 . TNOR	MALIZING VA			KE LAKIHI (U	' I
	Ordinary	Orgueil (CI	Leedy	BSE	Ave. Shale	
	Chondrites	Chondrite)				
La	0.329	0.236	0.378	0.648	41	
Ce	0.865	0.619	0.976	1.675	83	
Pr	0.130	0.09		0.254	10.1	
Nd	0.630	0.463	0.716	1.25	38	
Sm	0.203	0.144	0.23	0.406	7.5	
Eu	0.077	0.0547	0.0866	0.154	1.61	
Gd	0.276	0.199	0.311	0.544	6.35	
Tb	0.055	0.0353		0.099	1.23	
Dy	0.343	0.246	0.39	0.674	5.5	
Но	0.077	0.0552		0.149	1.34	
Er	0.225	0.162	0.255	0.438	3.75	
Tm	0.035	0.022		0.068	0.63	
Yb	0.220	0.166	0.249	0.441	3.53	
Lu	0.034	0.0245	0.0387	0.0675	0.61	

"Ordinary chondrites" is modified from Nakamura (1974), Orgueil are the values tabulated by Anders and Grevesse (1989), Leedy, an ordinary chondrite, is from Masuda (1973), BSE is Bulk Silicate Earth from McDonough and Sun (1994), and Average Shale is from Piper (1974).



La Ce Pr Nd Pm Sm Eu Gd Tb Dy Ho Er Tm Yb Lu Figure 7.8. Shale-normalized REE earth patterns of a Pacific pelagic sediment (V21-196; Ben Othman et al., 1989), the Post-Archean shale composite (McLennan, 1989), and typical seawater (Elderfield and Greaves, 1982). Both the pelagic sediment and seawater display a negative Ce anomaly, a consequence of Ce being in the +4 oxidation state.

several sets of normalizing values (one set is given in Table 7.3, others may be found in Rollinson, 1993), but the relative abundances are all similar. Figure 7.8 shows examples of shale-normalized rare earth patterns.

Because the rare earths are highly insoluble and immobile, rare earth patterns often remain unchanged during metamorphism. Hence rare earth patterns can provide information on the premetamorphic history of a rock. Indeed, even during the production of sediment from crystalline rock, the

Geochemistry

CHAPTER 7: TRACE ELEMENTS

rare earth patterns often remain little changed, and rare earth patterns have been used to identify the provenance, i.e., the source, of sedimentary rocks. Rare earth patterns have also become useful tools in chemical oceanography, now that modern analytical techniques allow their accurate determination despite concentrations in the parts per trillion range.

7.2.2.5 THE HFS ELEMENTS

The *high field strength (HFS) elements* are so called because of their high ionic charge: Zr and Hf have +4 valence states and Ta and Nb have +5 valence states. Th and U are sometimes included in this group. As we noted, Th has a +4 valence state and U either a +6 or +4 valence state. Because of their high charge, all are relatively small cations, with ionic radii of 64 pm for Nb⁵⁺ and Ta⁵⁺, and 72 and 76 pm for Zr⁴⁺ and Hf⁴⁺ respectively (U⁴⁺ and Th⁴⁺ are larger, however). Although they are of appropriate size for many cation sites in common minerals, their charge is too great and requires one or more coupled substitutions to maintain charge balance. As we noted earlier, such substitutions are energetically unfavorable. Thus Hf and Zr are moderately incompatible elements while Nb and Ta are highly incompatible elements. These elements are less electropositive than the alkalis, and alkaline and rare earths. That, as well as their high charge and the involvement of *d* orbitals (which are highly directional) in bonding in the case of Ta and Nb, means that there is a greater degree of covalency in the bonds they form. Thus the simple charged sphere is a less satisfactory model of their behavior.

As a consequence of their high ionic potential, or ionic charge to ionic radius ratio, the HFS elements are particularly insoluble. As a result, these elements tend to be very immobile during weathering and metamorphism. They are therefore particularly valuable in the study of ancient igneous rock suites as they can sometimes provide insights into the environment in which those rocks formed. Ta and Nb are present in anomalously low concentrations in magmas associated with subduction zones (indeed, this is considered a diagnostic feature of subduction-related volcanism). Although this depletion is not well understood, it is probably at least in part a consequence of the low solubility of these elements and the consequent failure of aqueous fluids generated by dehydration of the subducting oceanic crust to transport these elements into the magma genesis zone.

7.2.2.6 The First Series Transition Metals

The chemistry of the transition elements is considerably more complex than that of the elements we have discussed thus far. There are several reasons for this. First, many of the transition elements have two or more valence states in nature. Second, the transition metals have higher electronegativity than the alkali and alkaline earths, so that covalent bonding plays a more important role in their behavior. Bonding with oxygen in oxides and silicates is still predominantly ionic, but bonding with other non-metals, such as sulfur, can be largely covalent. A final complicating factor is the geometry of the *d*-orbitals, which are highly directional and thus bestow upon the transition metals specific preferences for the geometry of coordinating anions, or ligands. We will discuss this aspect of their behavior in more detail in a subsequent section.

The solubility of the transition metals, though generally lower than that of the alkalis and alkaline earths, is quite variable and depends upon valence state and the availability of anions with which they can form soluble coordination complexes. Their behavior in magmas is also variable. They range from moderately incompatible (e.g., Ti, Cu, Zn) to very compatible (e.g., Cr, Ni, Co), but their exact behavior is generally a stronger function of composition (of both solid and melt phases) than that of the highly incompatible elements. With the exception of Mn, the first transition series metals are also siderophile and/or chalcophile.

7.2.2.7 The Noble Metals

The platinum group elements (Rh, Ru, Pd, Os, Ir, Pt) plus gold are often collectively called the *noble metals*. These metals are so called for two reasons: first they are rare, second they are unreactive and quite stable in metallic form. Their rarity is in part a consequence of their highly siderophilic character. The concentration of these elements in the silicate Earth is only about 1% of their concentrations in

Geochemistry

CHAPTER 7: TRACE ELEMENTS

chondrites. Presumably, the bulk of the Earth's inventory of these elements is in the core. As a result of their low concentrations, their behavior is still poorly understood.

These elements are all also chalcophile (i.e., all are enriched in sulfide liquids relative to silicate liquids), although to varying degrees. Ir appears to be the most chalcophile, Pt and Au the least. Considering the associations of their ore deposits, the platinum group elements (PGE) may be divided into two subgroups: the Ir group (Ir, Os, and Ru) and the Pd group (Rh, Pd, Pt). The former are often associated with chromites (chromite is (Fe,Mg)Cr₂O₄, part of the spinel solid solution) in ultramafic rocks as native metal alloys or sulfide (e.g., the Stillwater Complex of Montana), while the latter are associated with magmatic sulfides of Fe, Ni, and Cu associated with gabbros (e.g., the Sudbury Complex of Ontario). Besides forming compounds with sulfur, these elements form a variety of chloride and other halide complexes. These complexes may play an important role in the formation of some PGE deposits and certainly do in many gold deposits.

These elements are transition elements and, like the first transition series, can exist in multiple valence states, ranging from +1 to +8, and have bonding behavior influenced by the *d*-orbitals. Thus their chemistry is complex. When in high valence states, some, Os and Ru for example, form oxides that are highly volatile.

The nature of the host phase for these elements in the mantle is unclear. Mitchell and Keays (1981) found that they were concentrated in minerals in the order garnet<olivine<orthopyroxene <clinopyroxene<spinel, but that the total inventory in all minerals was less than that in the bulk rock. They concluded that 60-80% of the PGE's were present in intergranular sulfides. An alternative host is native metal grains, such as osmiridium (an approximately 50-50 Os-Ir alloy), which have been found in peridotites. A study by Brügmann et al. (1987) of komatiites revealed that the concentrations of Au and Pd increased with decreasing MgO while those of Ru, Os, and Ir decreased with decreasing MgO concentrations. This suggested that Au and Pd are moderately incompatible elements while are Ru, Os, and Ir are highly compatible. Subsequent work has confirmed this conclusion and shown that Rh is compatible while Pt is incompatible. Recent studies suggest that the elements. The solubility of these elements in silicate melts, relative to native metal alloys, appears to be quite low, but is probably not exceeded in nature.

In a manner analogous to the rare earths, PGE data are sometimes presented as plots of the chondrite-normalized abundance, as in Figure 7.9. In this case, however, elements are not ordered by atomic number, but rather by decreasing melting point. This, as it turns out, places them in order of decreasing compatibility. The dunite in Figure 7.9 has a relatively flat noble metal pattern, as do many mantle peridotites. The remaining samples show varying degrees of enrichment in the Pd group elements relative to the Ir group elements. The chromites and sulfides in Figure 9 are highly enriched in noble metals relative to the silicate magmatic rocks, consistent with their chalcophile nature and demonstrated affinity for oxide phases such as the spinels.

7.2.2.8 Other Elements

Of the elements that do not fit into any of the above groups, several deserve special comment as they are of particular interest in isotope geochemistry. The first of these is boron. The geochemistry of boron remains incompletely understood, but knowledge of it has grown since the mid-1980's when it was shown that the relative proportions of its two isotopes, ¹⁰B and ¹¹B, vary in nature. Boron is only mildly electropositive, meaning that the bonds it forms generally have a substantial covalent component. In nature, it is most often bound to three oxygens to form the borate complex. Borate is quite soluble, making borate one of the major ions in seawater. Furthermore, borate is mobile and easily leached or added to rocks during weathering and metamorphism. In igneous systems it is mildly incompatible. Boron appears to be quite readily removed from subducting oceanic crust and sediment by fluids and enriched in subduction-related magmas.

Geochemistry

CHAPTER 7: TRACE ELEMENTS



Figure 7.9. Chondrite-normalized abundances of the noble metals in a variety of igneous rocks and associated sulfides. "Sudbury" is the magmatic sulfide from the Little Strobie #1 Mine of the Sudbury intrusion (Naldrett et al., 1979), "Langmuir" is the sulfide of the Langmuir komatiite, Canada (Naldrett et al., 1979), "Merensky Reef" is a composite of the Merensky Reef chromite layer of the Bushveld Complex, South Africa (Heimstra, 1979), "JM Reef" is a chromite layer of the Stillwater Complex, Montana (Barnes et al., 1985), "Dunite" is a dunite from the Urals (Stockman, 1982), "Onias's Flow" is a komatiite flow From the Belingwe Greenstone Belt, Zimbabwe (Zhou, 1994), "Tulkonsky Picrite" is a picrite from the Siberia Traps (Brügmann et al., 1993), and "Ave Leg 115" is the average of basalts drilled from the Indian Ocean on ODP Leg 115 (Fryer and Greenough, 1982).

Rhenium is of interest because it decays radioactively to Os. Though it is not one of the platinum group elements, Re is adjacent to them in the periodic table and shares many of their properties, including being highly refractory under reducing conditions, being both highly siderophile and chalcophile, having a variety of potential valence states, having a large E_{H} , so that the metal is relatively resistant to oxidation, and forming a volatile (though only moderately so) oxide species. In oxidized aqueous solutions at the Earth's surface it is present as the perrhenate ion, ReO_4^- , and is quite soluble. However, it is readily adsorbed from solution under reducing conditions, leading to its concentration in organic-rich materials such as black shales and coals. In igneous systems, it is incompatible, though details of its behavior are not understood.

Lead is of great interest not only because of its economic importance but also because it is the product of radioactive decay of 232Th, 235U, and ²³⁸U and is quite toxic. The latter is of concern because its widespread use, particularly in paint and gasoline, has led to widespread pollution. Pb is chalcophile, though not so much as the PGE's and Re, and perhaps slightly siderophile. It is also quite volatile. Pb is in the +2 state throughout virtually the entire range of natural redox conditions. Pb has an electronegativity of 1.9, indicating a greater degree of covalency in its bonding than for the alkalis, alkaline earths and rare earths. The solubility of Pb is reasonably low under most conditions, but it can form strong complexes with elements such as Cl

and it can be readily transported in metamorphic and hydrothermal solutions. Its ionic radius is 119 pm in octahedral coordination, virtually identical to that of Sr. In igneous systems it is moderately incompatible, as might be expected from its ionic radius and charge.

The concentration of phosphorus is sometimes high enough that it is treated as a major element. With a valence of +5 and being moderately electropositive, it is generally present in nature as the oxyanion PO_4^{3-} , in which it is doubly bound to one of the oxygens and singly bound to the others. In

Geochemistry

CHAPTER 7: TRACE ELEMENTS

rocks, it forms the common mineral apatite $(Ca_3(PO)_4(OH,F,Cl))$ as well as rarer minerals such as monazite. It behaves as a moderately incompatible element in mafic and ultramafic igneous systems.

The elements of the U and Th decay series have no stable nuclei. They exist on the Earth only because they are continually created through decay of U and Th. They are of interest in geochemistry only because of their radioactivity. As we shall see in the next chapter, they can be useful geochronological tools. As radioactive substances, they represent potential environmental hazards and are of interest for this reason as well. Radon is perhaps the element that causes the greatest concern. As a noble gas, it is quite mobile. Relatively high levels of this gas can accumulate in structures built on uranium-rich soils and rocks. In this group we could also include the two "missing" elements, Tc and Pm. Like the U-decay series elements, they have no long-lived nuclei. They are, however, present in the Earth in exceedingly minute amounts, created by natural fission of U as well as by capture of fissionproduced neutrons. Merely detecting them is a considerable challenge.

Ga and Ge can substitute, albeit with difficulty because of their larger radii, for their more abundant neighbors directly above them in the periodic table, Al and Si. Both are thus moderately incompatible elements. Their concentrations in the silicate Earth are somewhat low because of their siderophile nature. Germanium's greatest contribution to geochemistry may, however, be in experimental studies. Germanates created by substituting Ge for Si are used to simulate the properties of silicates at high pressures. This approach works because the oxygen ion is far more compressible than are cations. Thus the ratio of the ionic radius of oxygen to silicon at high pressure is similar to that of oxygen to germanium at low pressure. Such studies of "germanium analogs" considerably advanced the understanding of the Earth's deep interior decades before the technology for reproducing the pressure of the deep Earth in the laboratory existed. Studies of germanium analogs continue in parallel with ultrahigh pressure experiments.

The geochemistry of the remaining elements, particularly in igneous processes, is poorly understood. Because of their low abundances, there have been few analyses published of these elements. Progress is being made on this front, however. For example, Newsom et al. (1986) demonstrated that the behavior of the chalcophile element Mo closely follows that of the rare earth Pr. W appears to be highly incompatible and its behavior mimics that of Ba. In oxidizing solutions, Mo forms a very soluble oxyanion complex, MoO_4^{2-} , so that its concentration in seawater is relatively high. Tin (Sn) and antimony (Sb) behave as moderately incompatible elements, with behaviors in igneous systems that are similar to that of the rare earth Sm (Jochum et al., 1985, 1993). However, these elements appear to form soluble species so, unlike Sm, they can be relatively mobile. It appears, for example, that they are readily stripped from subducting oceanic crust and carried into the magma genesis zones of island arc volcanos.

7.3 DISTRIBUTION OF TRACE ELEMENTS BETWEEN CO-EXISTING PHASES

7.3.1 The Partition Coefficient

Geochemists find it convenient to define a partition or distribution coefficient as:

$$D_i^{\alpha-\beta} = \frac{C_i^{\alpha}}{C_i^{\beta}}$$
 7.1

where C is concentration, *i* refers to an element (or species) and α and β are two phases. By convention, if one phase is a liquid, the concentration ratio is written solid over liquid, i.e.:

$$D_i^{s/\ell} = \frac{C_i^s}{C_i^\ell}$$
 7.2

where *s* refers to some solid phase and ℓ refers to the liquid phase. The distribution coefficient is a convenient concept for relating the concentration of some element in two different phases. It is also readily measured, either experimentally or 'observationally'. In the former, two phases are equilibrated at the temperature and pressure of interest and the concentration of *i* is subsequently measured in both. In

Geochemistry

CHAPTER 7: TRACE ELEMENTS

the latter, the concentration of element i is simply measured in two natural phases thought to be in equilibrium. In both cases, the catch is, of course, equilibrium. As we mentioned Chapter 5, kinetic effects can lead to apparent partition coefficients that differ from equilibrium ones.

Having introduced the partition coefficient, we can now quantitatively define two terms we have already introduced: compatible and incompatible. *Incompatible* elements are those with $D^{s/l} \ll 1$. *compatible* elements are those with $D^{s/l} \ge 1$. The partition coefficient for a given element will vary considerably between phases and can be less than one for one phase and greater than one for another. Hence the terms compatible and incompatible have meaning only when the phases are specified. These terms refer to partitioning *between silicate melts and phases common to mafic or ultramafic (i.e., basaltic or peridotitic) rocks*. It is this phase assemblage that dictates whether lithophile trace elements are concentrated in the Earth's crust, hence the significance of these terms.

7.3.1.1 Thermodynamic Basis

Though trace elements are often treated differently than major elements, we should remember that the principle governing their distribution is the same as that governing the distribution of major elements. We are already familiar with this principle: *at equilibrium, elements will distribute themselves between co-existing phases so that the chemical potential of that element is the same in every phase in the system*. The chemical potential of element *i* in phases α and β is given by equation 3.54:

$$\mu_i^{\alpha} = \mu_i^{o\alpha} + RT \ln X_i^{\alpha} \lambda_i^{\alpha}$$
7.3

and

$$\mu_i^{\beta} = \mu_i^{o\beta} + RT \ln X_i^{\beta} \lambda_i^{\beta}$$
7.4

At equilibrium, $\mu^{\alpha} = \mu^{\beta}$, so: $\mu_{i}^{o\beta} - \mu_{i}^{o\alpha} = RT \ln(X_{i}^{\alpha} \lambda_{i}^{\alpha} / X_{i}^{\beta} \lambda_{i}^{\beta})$ 7.5 While equation 7.5 is expressed in mole fraction, the equation is identical when expressed in ppm because any difference in the molecular weights of the phases can be incorporated in the activity coefficients. A little algebra shows that the distribution coefficient is related to chemical potential as:

$$D_i^{\alpha-\beta} = \frac{C_i^{\alpha}}{C_i^{\beta}} = \frac{\lambda_i^{\beta}}{\lambda_i^{\alpha}} \exp\left(\frac{\mu_i^{o\beta} - \mu_i^{o\alpha}}{RT}\right)$$
7.6

The standard chemical potentials on the right hand side are the chemical potentials of *i* in pure *i* versions of phases α and β . Suppose, for example, we were interested in the partitioning of nickel between olivine and a silicate melt. In that case, $\mu_i^{\alpha_0}$ would be the chemical potential of Ni in pure Ni olivine (i.e., Ni₂SiO₄) and $\mu_i^{\beta_0}$ would be the chemical potential of Ni in Ni-silicate melt. This difference in chemical potential is the standard state Gibbs Free Energy change resulting from transfer of *i* between these two phases, so:

$$D_i^{\alpha-\beta} = \frac{\lambda_i^{\beta}}{\lambda_i^{\alpha}} \exp\left(\frac{-\Delta G_i^{\alpha\alpha-\beta}}{RT}\right)$$
7.7

It is reasonable to expect that species present in trace quantities will obey Henry's Law. In that case, we can rewrite equ. 7.7 as:

$$D_i^{\alpha-\beta} = \frac{h_i^{\beta}}{h_i^{\alpha}} \exp\left(\frac{\mu_i^{o\beta} - \mu_i^{o\alpha}}{RT}\right)$$
7.8

where h is the Henry's Law constant (Chapter 3). Drake and Weill (1975) found that Sr and Ba obeyed Henry's Law in the system plagioclase-liquid up to concentrations of 5% or so of the element in the liquid, i.e., well above most natural concentrations. Many subsequent studies have confirmed that trace elements generally obey Henry's Law over their entire range of natural concentration.

Comparing our derivation of the partition coefficient with that of the equilibrium constant shows that the two are closely related. Indeed, the partition coefficient is a form of apparent equilibrium constant and effectively identical with the K_D as defined in equation 3.88 and 3.89. The terms partition coefficient, distribution coefficient, and K_D are synonymous and often used interchangeably.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

In a system with three phases, α , β , and γ , if α and β are in equilibrium and α and γ are in equilibrium, then β and γ must also be in equilibrium. It follows that:

$$D^{\beta/\gamma} = D^{\alpha/\gamma}/D^{\alpha/\beta}$$
 7.9

This relationship has practical use. For example, if we can determine the partition coefficient for an element between pyroxene and melt and between garnet and pyroxene, we can then calculate the garnet-melt partition coefficient for this element from equation 7.9.

7.4 FACTORS GOVERNING THE VALUE OF PARTITION COEFFICIENTS

7.4.1 TEMPERATURE AND PRESSURE DEPENDENCE OF THE PARTITION COEfficient

In ideal solutions, the temperature dependence of the partition coefficient is the same as that of the equilibrium constant, i.e.,:

$$D_i = \exp\left(\frac{-\Delta G_i^o}{RT}\right)$$
7.10

 Δ G in equation 7.10 can be expanded into entropy and enthalpy terms, as in equation 4.38:

$$\left(\frac{\partial \ln D_i}{\partial T}\right)_P = \frac{\Delta H^o + (P - P^\circ)\Delta V}{RT^2}$$
7.11

In ideal solution, and assuming ΔV is independent of temperature and pressure, the pressure dependence is also the same as that of the equilibrium constant, i.e.:

$$\left(\frac{\partial \ln D_i}{\partial P}\right)_T = \frac{-\Delta V}{RT}$$
7.12

From equation 7.12, we would predict a strong pressure dependence when the ionic radius of an element differs greatly from that of the available crystal lattice site. Thus for example, we would predict the partition coefficient for K between pyroxene and melt would be strongly pressure dependent since the ionic radius of K is 150 pm and is much larger that the size of the M2 site in clinopyroxene, which is normally occupied by Ca, with a radius of about 100 pm. Conversely, where the size difference is small, e.g., Mn (83 pm) substituting for Fe (78 pm), we would expect the pressure dependence to be smaller.

For non-ideal solutions (which will be the usual case), the temperature and pressure dependence will be more complex than given in equations 7.11 and 7.12 because the ratio of Henry's Law constants in equation 7.0 will also be pressure and temperature dependent. The form of that dependence is usually difficult to deduce.

7.4.2 Ionic Size and Charge

Much of the interest in trace elements in igneous processes centers on the elements located in the lower left portion of the periodic table (alkalis K, Rb, Cs; alkaline earths Sr and Ba, the rare earths, Y, Zr, Nb, Hf and Ta). The reason for this focus of attention is in part due to the relative ease with which these elements can be analyzed. These elements are all lithophile and therefore present at *relatively* high abundance in the Earth's crust and mantle. There is another reason, however: their chemical behavior is comparatively simple. All have electronegativities less than 1.5 (Nb is the sole exception, with electronegativity of 1.6), and most have only a single valence state over the range of oxygen fugacity in the Earth. The difference in electronegativity between these elements and oxygen, effectively the only electronegative element in igneous rocks, is 2 or greater, so the bonds these elements will form are predominantly ionic. To a reasonable approximation then, the atoms of these elements behave as hard spheres containing a fixed point charge at their centers. Thus the two factors that most govern their chemical behavior are ionic radius and ionic charge. As we shall see, these elements have partition coefficients less than 1 for most common minerals. The term "incompatible elements" often refers specifically to these elements.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

The other trace elements that receive the most attention from igneous geochemists are the first transition series elements. Though their electronic structures and bonding behavior are considerably more complex (as we shall see in a subsequent section), charge and size are also important in the behavior of these elements. Many of these elements, particularly Ni, Co, and Cr have partition coefficients greater than 1 in many Mg-Fe silicate minerals. Hence the term "compatible elements" often refers to these elements.

The effect of ionic charge and size is illustrated in Figure 7.10, which shows a plot of ionic radius vs. charge contoured for the clinopyroxene/melt partition coefficient. Those elements whose charge and ionic radius most closely matches that of the major elements present in the cation sites, Mg, Fe, and Ca, have partition coefficients close to 1, those whose charge or radius differs significantly have lower partition coefficients. Thus Ba, even though its charge is the same (2+), has an ionic radius of 135 pm and fits only with difficulty into the lattice site normally occupied by Ca (100 pm), Mg (72 pm) or Fe (78 pm). This strains the lattice, and hence additional energy is required for the substitution to occur. On the other hand, Zr, which has an ionic radius identical to that of Mg, is not accepted in this site because its charge, 4+, is too great. Substitution of Zr^{4+} for Mg^{2+} would require either leaving one cation site vacant or one or more coupled substitutions (e.g., Al^{3+} for Si⁴⁺) to maintain charge balance. Both of these are energetically unfavorable. In addition, ions having a radius much smaller than that of the element normally occupying the site also have low partition coefficients because their substitution also induces strain in the lattice.

7.4.2.1 Goldschmidt's Rules of Substitution

The importance of ionic radius and charge have been long known. Indeed, Goldschmidt developed the following rules regarding substitutions of elements into crystal structures:

1. If 2 ions have the same radius and the same charge, they will enter a given lattice site with equal facility.

2. If two ions have similar radii and the same charge, the smaller ion will enter a given site more readily.

3. If two ions have similar radii, the ion with the higher charge will enter a given site more readily.

Ringwood[‡] noted the need also to consider the electronegativity in substitution.



Figure 7.10. Ionic radius (picometers) vs. ionic charge contoured for clinopyroxene/liquid partition coefficients. Cations normally present in clinopyroxene are Ca^{2+} , Mg^{2+} , and Fe^{2+} , shown by ***** symbols. Elements whose charge and ionic radius most closely match that of the major elements have the highest partition coefficients.

[‡] Alfred E. Ringwood (1930-1993). After receiving his BSc and PhD degrees at the University of Melborne, Ringwood did a post-doctoral fellowship at Harvard University. His subsequent career was spent at the Australian National University, from where he exerted a strong influence on the development of earth science in Australia. Ringwood pioneered the post-war development of many aspects of geochemistry, including phase relationships of the mantle at very high pressures, the origin of basaltic magmas, the origin of the Earth and the Moon, as well as safe methods of long-term nuclear waste storage. Among his other contributions, he demonstrated that seismic discontinuities in the mantle were due to phase transitions, the existence of which he predicted using germanate analogs.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

His rule is:

4. Whenever a substitution is possible between two elements with significantly different electronegativities ($\Delta > 0.1$), the one with the lower electronegativity will be preferentially incorporated. However, the bottom-line is: the stronger the bond, the more likely the substitution.

7.4.2.2 QUANTITATIVE TREATMENT OF IONIC SIZE AND CHARGE

Goldschmidt's rules are simple qualitative statements based on empirical observation. Can we turn these into quantitative tools? In the last 10 to 15 years, the answer has become yes. We can take advantage of the thermodynamic framework developed over the past century and incorporate into it the energetic effects of substituting a misfit ion into a specific lattice site to develop quantitative predictions of distribution coefficients. As we have seen, for example with the surface complexation model, thermodynamics generally, and the Gibbs Free Energy in particular, provide a wonderfully flexible framework into which we can incorporate new knowledge and understanding, often developed on the microscopic scale, to make quantitative predictions about the macroscopic behavior of chemical systems.

In brief, the idea is that when a trace element of different ionic radius is substituted for the ion that normally occupies a lattice site the lattice must adjust. This adjustment is accomplished through the expenditure of strain energy, which can be calculated by measurable parameters of the mineral. Since this strain energy contributes to the Gibbs Free Energy of reaction, and since the partition coefficient is related to ΔG , the effect of the lattice strain energy on the partition coefficient can be calculated.

Consider the formation of diopside containing metal ion M^{2+} in the site normally occupied by Ca^{2+} . We could write this reaction as:

$$MO^{\ell} + MgO^{\ell} + 2SiO_{2}^{\ell} \rightleftharpoons MMgSi_{2}O_{6}$$
 7.13

where ℓ denotes the liquid phase. However, the mineral MMgSi₂O₆ may not exist in nature and may not be synthesizable in the laboratory, so it might not be possible to determine its thermodynamic properties. Instead, we can imagine two reactions. The first would be the crystallization of diopside from a melt:

$$\operatorname{CaO}^{\ell} + \operatorname{MgO}^{\ell} + 2\operatorname{SiO}_{2}^{\ell} \rightleftharpoons \operatorname{CaMgSi}_{2}\operatorname{O}_{6}$$
 7.13a

The second would be an exchange reaction such as the replacement of Ca^{2+} in the M2 site of diopside by metal ion M^{2+} :

$$M^{\ell} + CaMgSi_2O_6 \rightleftharpoons Ca^{\ell} + MMgSi_2O_6$$
 7.13b

We can see that reaction 7.13 is simply the difference between reaction 7.13b and 7.13a. Hence, the Gibbs Free Energy change of this reaction can be expressed as:

$$\Delta G_r = \Delta G_{exchange}^{M-Ca} - \Delta G_{melting}^{Di}$$
7.14

The first term is the free energy change in involved in transferring a M^{2+} ion from the melt into the crystal lattice and simultaneously transferring a Ca^{2+} ion from the lattice site to the liquid. The second term is the free energy change associated with melting of diopside and in multi-component systems governs the distribution of Ca between diopside and the liquid. The distribution coefficient for element M then depends on these two components of free energy:

$$D_{M}^{Di/\ell} = \exp\left(\frac{\Delta G_{melting}^{Di} - \Delta G_{exchange}^{M-Ca}}{RT}\right)$$
7.15

According to the lattice strain energy theory of Beattie (1994) and Blundy and Wood (1994, 2003), $\Delta G_{exchnage}$ is dominated by the energy associated with the lattice strain that results from the M²⁺ ion being a different size than the Ca²⁺ ion normally occupying the lattice site. Because the melt (at least at low pressure) has a far less rigid structure and is more compressible than the solid, any strain in the melt is essentially negligible compared the strain in the solid. Hence:

$$\Delta G_{exchange}^{M-Ca} \cong \Delta G_{strain}$$

$$7.16$$

CHAPTER 7: TRACE ELEMENTS

According to Brice (1975), the strain energy, ΔG_{strain} , maybe calculated from:

$$\Delta G_{strain} = 4\pi E N_A \left[\frac{r_0}{2} (r_M - r_o)^2 + \frac{1}{3} (r_M - r_0)^3 \right]$$
7.17

where r_0 is the optimal radius of the lattice site, r_M is the ionic radius of M, N_A is Avogadro's Number, and *E* is Young's Modulus. The Young's Modulus of a substance is the ratio of stress applied to the resulting strain and may be calculated as:

$$E = \frac{L_0 F}{\Delta L A}$$
 7.18

where L_0 is the original length of the object undergoing strain, ΔL is the length change under stress, F is the force applied over area A. E has units of pressure. Young's Modulus is related to the Bulk Modulus, K, (which, as we found in Chapter 2, is the inverse of compressibility) through Possion's ratio. Poisson's ratio is the ratio of the transverse contraction per unit dimension to elongation per unit length of a bar of uniform cross section when subjected to tensile stress and in most silicates is ~0.25. Consequently, $E \cong 1.5 K$.

The $\Delta G_{melting}$ term in equation 7.15 governs the distribution of Ca between liquid and crystal, i.e.:

$$D_{Ca}^{Di/\ell} = \exp\left(\frac{\Delta G_{melting}^{Di}}{RT}\right)$$
7.19

Substituting this and equation 7.17 into equation 7.15 we obtain:

$$D_{M}^{Di/\ell} = D_{Ca}^{Di/\ell} \exp\left(\frac{-\Delta G_{strain}^{M-Ca}}{RT}\right) = D_{Ca}^{Di/\ell} \exp\left(\frac{-4\pi E N_{A} \left\lfloor \frac{r_{0}}{2} (r_{M} - r_{0})^{2} + \frac{1}{3} (r_{M} - r_{0})^{3} \right\rfloor}{RT}\right)$$

More generally, this equation may be written as:

$$D_{i}^{s/\ell} = D^{o} \exp\left(\frac{-4\pi E N_{A} \left[\frac{r_{0}}{2}(r_{M} - r_{0})^{2} + \frac{1}{3}(r_{M} - r_{0})^{3}\right]}{RT}\right)$$
7.20

where D° is the partition coefficient of an ion of radius r_0 which has the same charge as *i* and enters the lattice site without strain (Blundy and Wood, 1994).

Now consider the case where the substituting trace element has a different charge than the ion normally occupying the site, for example:

$$La^{3+\ell} + CaMgSi_2O_6 \rightleftharpoons Ca^{2+\ell} + LaMgSi_2O_6$$
7.21

The charge imbalance must be compensated in one of two ways: either through a coupled substitution, such as:

$$La^{3+\ell} + Na^{+\ell} + 2CaMgSi_2O_6 \rightleftharpoons 2Ca^{2+\ell} + 2La_{0.5}Na_{0.5}MgSi_2O_6$$
7.22a

 $La^{3+\ell} + Al^{3+\ell} + CaMgSi_2O_6 \rightleftharpoons Ca^{2+\ell} + Si^{4+} LaMgAlSiO_6$ or through a creation of a vacancy, e.g.:

$$2La^{3+\ell} + 3CaMgSi_2O_6 \rightleftharpoons 3Ca^{2+\ell} + 2LaMgSi_2O_6 + MgSi_2O_6 \qquad 7.23$$

If the charge balance mechanism is known, then free energy of exchange may be calculated as the sum of free energies associated with melting, strain, and the charge balance mechanism. From this, the distribution coefficient can be calculated. For example, if charge is balanced by the coupled substitution in

7.22b

Geochemistry

CHAPTER 7: TRACE ELEMENTS

7.22a, then the distribution can be calculated as

$$D_{La}^{Di/\ell} = e^{\frac{\Delta G_{melting}^{Di} - \Delta G_{strain}^{La} - \Delta G_{exchange}^{Na-Ca}}{RT}}$$
7.24

The free energy of the Na-Ca exchange reaction will, of course, also have a strain component. Equation 7.24 may be written as:

$$D_{La}^{Di/\ell} = e^{\frac{\Delta G_{melting}^{Di} - \Delta G_{exchange}^{Na-Ca}}{RT}} e^{\frac{-\Delta G_{strain}^{La-Ca}}{RT}} 7.24a$$

The first term on right may be replaced by D°, so that equation 7.20 may also be applied to heterovalent exchanges.

Equations 7.20 suggests that the dependence of the partition coefficient on ionic radius should be highly non-linear, and experiments have proved this to be the case. Figure 7.11 compares clinopyroxene-liquid and plagioclase-liquid partition coefficients for the rare earths determined experimentally at 1225° C and 1.5 GPa with partition coefficients predicted using equation 7.20. Plots of partition coefficient vs. ionic radius are known as Onuma diagrams after Onuma et al. (1968).

The lattice strain model also helps us predict the effects of pressure on partitioning, as pressure will affect the dimension of the lattice site. For example, clinopyroxene incorporates an increasing amount of Na and Al as pressure increases, which results in a decrease in the radius of the M2 site. This shifts the partition coefficient parabola on the Onuma diagram. As illustrated in Figure 7.12, partition coefficients of large ions will decrease with increasing pressure, while those of small ions in will increase.

7.4.3 Compositional Dependency

Equation 7.8 predicts that partition coefficients will depend on the composition of the phases involved since the Henry's Law constant will depend on the concentration of the major species. Experience has substantiated this prediction and shown that partition coefficients are composition-dependent. For example, elements that are incompatible in mafic (SiO₂-poor) systems often have partition coefficients equal to or greater than 1 in silicic (SiO₂-rich) systems. This behavior is illustrated by the solubility of zircon, $ZrSiO_4$. In basaltic liquids, zircon will not precipitate until the Zr concentration reaches 10000 ppm (1% Zr) or more. In



Figure 7.11. Experimentally determined clinopyroxene–liquid and plagioclase–liquid partition coefficients for rare earths as a function of ionic radius compared to predicted partition coefficients using equation 7.20. For clinopyroxene, the model assumes $D^{\circ} = 1.47$, E = 262 GPa, and $r_0 =$ 98.1 pm. For plagioclase, these parameters are D° = 0.26, E = 113 GPa, and $r_0 = 128$ pm. Data and model from Blundy et al. (1998).



Figure 7.12. Because increasing pressure increases the amount of jadeite (NaAlSi₂O₆) in clinopyroxene, the radius of the M2 site decreases and partition coefficients for large ions decrease while those for small ions increase. After Wood et al. (1999).

CHAPTER 7: TRACE ELEMENTS

granitic liquids, however, zircon crystallizes at Zr levels around 100 ppm. In another example, Watson (1976) conducted two-liquid partitioning experiments in which the concentration of various elements was measured in coexisting immiscible mafic and silicic silicate liquids. P, the rare earth elements (REE), Ba, Sr, Mg, Zr, Mn, Ti, Cr, Ca, and Ta were found to be enriched in the mafic melt by factors of 1.5 to 10 (in the above order). Cs was enriched in the silicic melt by a factor of about 3.

Some of this compositional dependence results from the need for coupled substitutions to maintain charge balance. For example, Reactions 7.22a and 7.22b suggest clinopyroxene-liquid rare earth partition coefficients will depend on Na and Al concentrations. However, much of the compositional dependence can be related to structural changes in the melt phase. As we saw



Figure 7.13. Variation of the zinc olivine/liquid partition coefficient as a function of the number of non-bridging oxygens per tetrahedral cation (NBO/T) in experiments by Kohn and Schofield (1994).

in Chapter 4, as the SiO₂ concentration of the melt increases, the polymerization of the melt increases as the Si tetrahedra become increasingly linked. On the other hand, Si-poor melts will be largely depolymerized as more non-bridging oxygens are needed to charge-balance network-modifying cations. The ratio of *non-bridging oxygens* to tetrahedral cations (Si⁴⁺ and Al³⁺), written NBO/T, provides a measure of the degree of melt depolymerization. Kohn and Schofield (1994) carried out a series of experiments demonstrating the effect of melt polymerization on olivine/melt partitioning of Zn and Mn. Figure 7.12 shows that the Zn partition coefficient decreases from about 4.5 in highly polymerized melts (low NBO/T) to about 0.8 in depolymerized melts (high NBO/T). Zn is probably present in octahedral sites (i.e., surrounded by 6 oxygens) in silicates melts, though this remains uncertain. Changes in the partition coefficient most likely reflect the availability of suitably coordinated sites in the melt.

Another effect is a decrease in octahedral or other sites for transition and highly charged cations as melts become more polymerized. High-field strength elements, such as Zr^{4+} and Ta^{5+} need to be coordinated by a large number of anions. For example Zr is coordinated by 8 oxygens in zircon. Cations with large radii, such as Cs, do not easily fit in tetrahedral sites. Lacking suitable sites in silicic melts, they partition preferentially into solid phases.

7.4.3.1 Compositional Dependency of Clinopyroxene Partitioning

Because of its rather large M2 octahedral site and its capacity to accept Al in the tetrahedral site for charge balance, the clinopyroxene–melt partition coefficients for many incompatible elements are larger than for other phases occurring in mafic and ultramafic rocks. Consequently, clinopyroxene exerts a strong control on trace element fractionation during melting and crystallization in these rocks and related magmas. Thus considerable attention has been given to clinopyroxene-melt partition coefficients and there have been a number of attempts to quantify the temperature, pressure, and compositional dependencies of these partition coefficients. As an example, we will briefly consider the work of R. L. Nielsen and his colleagues on this problem.

Gallahan and Nielsen (1992) carried out experiments on the partitioning of Sc, Y, and several of the rare earth elements over a range of compositions with temperature held approximately constant. Based on this and previous work, they found that the alumina concentrations in the melt most strongly affected the partition coefficient (Figure 7.15). For Y and the rare earths, they assumed that the reaction:

$$ReeO_{15}^{\ell} + (Fe, Mg)O + AlO_{15}^{\ell} + SiO_2 \rightleftharpoons Ree(Fe, Mg)AlSiO_6$$

$$7.25$$

best described the formation of rare earth (here abbreviated Ree) pyroxene. The equilibrium constant

Geochemistry

CHAPTER 7: TRACE ELEMENTS

for this reaction is:

$$\mathbf{K} = \left[\frac{a_{Ree}^{cpx}}{a_{Ree}^{\ell}a_{AIO_{1.5}}^{\ell}a_{SiO_{2}}^{\ell}}\right] \left[\frac{a_{Mg,Fe}^{cpx}}{a_{Mg,Fe}^{\ell}}\right]$$
7.26

where ℓ designates the liquid. It turns out the at any given temperature, the ratio of the activity of Fe²⁺ and Mg in the melt to their activity in the clinopyroxene is approximately constant, so that 7.26 simplifies to:

Example 7.1. Parameterizing Partition Coefficients

One approach to accounting for compositional variability when the underlying thermodynamics is not fully understood is to parameterize the partition coefficient, i.e., to express it as a function of other measurable or predictable variables. The Zn partition coefficients determined by Kohn and Schofield are clearly strong functions of the NBO/T parameter, suggesting we might usefully try to express the Zn partition coefficient as a function of NBO/T. From equation 7.7, we can expect that the partition coefficient will be a function of temperature as well. Furthermore, we can use equation 6.8 to predict the form of the equation. Taking the log of both sides of equation 7.7 we have:

$$\ln D = \ln \left(\frac{\lambda_i^{\beta}}{\lambda_i^{\alpha}}\right) - \frac{\Delta G^{\circ}}{RT}$$

The first term on the right expresses compositional dependence and the second term expresses the temperature dependence. Thus we might expect an equation of the form:

$$\ln D = a \ln \left(\frac{NBO}{T}\right) + \frac{b}{T}$$

We can then use a technique called *polynomial regression*, a multivariable extension of linear regression, to determine the constants *a* and *b*. Using this approach, we obtain the following equation:

$$\ln D = -0.405 \ln \left(\frac{NBO}{T} + 6.077\right) + \frac{3594}{T}$$
 7.27

Figure 7.14 compares the observed Zn partition coefficients with those predicted by equation 7.27.

This approach does have limitations and dangers, however. In the experiments of Kohn and Schofield (1994), for example, temperature and NBO/T were highly correlated, thus some of the temperature dependence may actually be hidden in the first term of equation. Second, it is highly un-



likely that all compositional effects on the partition coefficient can be adequately expressed by a single parameter. Furthermore, our equation does not take variations in the composition of ol ivine into account. Kohn an d Schofield's experiments were performed using iron-free compositions, so the olivines were nearly pure forsterite. In real systems the composition of the solid may also affect partitioning behavior (though fosterite-fayalite system is nearly ideal, so the importance of this factor for olivine is less than for other Nevertheless, equation 7.27 is minerals). highly successful and a counts for over 96% of the variation in the partition coefficient observed by Kohn and Schofield (1994).

Geochemistry

CHAPTER 7: TRACE ELEMENTS

$$\mathbf{K} \cong \left[\frac{a_{Ree}^{cpx}}{a_{Ree}^{\ell} a_{AlO_{1,5}}^{\ell} a_{SiO_{2}}^{\ell}} \right] \qquad 7.28$$

In addition to the compositional dependence, there is, of course, an expectation that the partition coefficient will depend on temperature as well. Gallahan and Nielsen addressed this problem by regressing the log of K as defined above against 1/T and determining the slope and intercept of the regression line, i.e., by expressing K as:

$$\ln K = a/T - b \qquad 7.29$$

(Compare equ. 7.29 to equ. 3.95. What should the a term represent? What should the b term represent?)

Gallahan and Nielsen's values for regression coefficients *a* and *b* are listed in Table 7.4. If we define $D^* = a_{REE}^{cpx} / a_{REE}^{l}$, the D can be written as:

$$D_{REE}^* = a_{AlO_{1.5}}^{\ell} a_{SiO_2}^{\ell} \mathbf{K}$$
 7.30

Taking the log of both sides of 7.30 and substituting equation 7.29 for ln K, we have:

$$\ln D_{REE}^* = \ln(a_{AIO_{1.5}}^{\ell} a_{SiO_2}^{\ell}) + \frac{a}{T} - b \quad 7.31$$

For Sc, Gallahan and Nielsen (1992) found it necessary also to include CaO in the reaction:

$$ScO_{1.5}^{\ell} + Cao^{\ell} + AlO_{1.5}^{\ell} + SiO_{2}^{\ell} \rightleftharpoons ScCaAlSiO_{6}$$
7.32

The corresponding equilibrium constant expression is:

$$\mathbf{K} = \frac{a_{Sc}^{cpx}}{a_{Sc}^{\ell} a_{AlO_{1,5}}^{\ell} a_{SiO_{2}}^{\ell} a_{CaO}^{\ell}}$$
 7.33

From this equation, an expression analogous to 7.31 may be derived. Activities for melt components are calculated using the 2-lattice model of Nielsen and Dungan (1983):

$$a_{AIO_{1.5}}^{\ell} = \frac{X_{AIO_{1.5}} - (X_{NaO_{0.5}} + X_{KO_{0.5}})}{\Sigma NM}$$



Figure 7.15. Experimentally determined clinopyroxeneliquid partition coefficient for Sm as a function the mole fraction of Al in the melt and the mole fraction of Ti in cpx. The correlation with Ti suggests Ti and REE partitioning is controlled by similar factors. Cross shows analytical error. From Gallahan and Nielsen (1992).

 Table 7.4. Regression Constants for

 Clinopyroxene-Melt Partitioning

	а	b	
La	119	1.95	
Sm	11371	8.60	
Gd	8959	6.62	
Ho	8651	7.62	
Lu	19709	14.29	
Y	12084	8.81	
Sc	11956	16.83	

From Gallahan and Nielsen (1992).

$$a_{SiO_2}^{\ell} = \frac{X_{SiO_2}}{\Sigma NF} \qquad \qquad a_{CaO}^{\ell} = \frac{X_{CaO}}{\Sigma NM}$$

$$7.34$$

where ΣNM is the sum of the 'network modifiers' and ΣNF is the sum of the 'network formers': $\Sigma NM = X_{AIO_{1.5}} + X_{CaO} + X_{MgO} + X_{FeO} + X_{P_2O_5} + X_{TiO_2} + X_{MnO} + X_{CrO_{1.5}} - X_{NaO_{0.5}} - X_{KO_{0.5}}$ 7.35

Geochemistry

CHAPTER 7: TRACE ELEMENTS

$$\Sigma NF = X_{SiO_2} + X_{NaO_{0.5}} + X_{KO_{0.5}}$$
7.36

The activities of Sc, Y, and the rare earths are assumed equal to:

$$a_i^{\ell} = \frac{X_i^{\tau}}{\Sigma NM} \qquad \qquad a_i^{cpx} = X_i^{cpx} \qquad \qquad 7.37$$

where *i* is Sc, Y, La, etc.

How does the D^{*} we defined above relate to D as defined in equation 7.2? With activities as defined in equation 7.34, D is related to D^{*} as:

$$D_{i} = D_{i}^{*} \frac{\sum moles \ oxides \ cpx \times Mol. \ Wt. \ i}{\sum NM \sum moles \ oxides \ liq. \times Mol. \ Wt. \ i}$$
7.38

The molecular weight terms cancel, of course, and the molar sums for cpx and basaltic liquid will generally be equal to within about 5%, so that 7.38 simplifies to:

$$D_i \approx \frac{D_i^*}{\Sigma NM}$$
 7.39

Hack et al. (1994) extended the work of Gallahan and Nielsen (1994) to include the effects of pressure. They found the equilibrium constant for reaction 7.25 could be expressed as follows:

$$\ln \mathbf{K} = c_1 + c_2 \times \frac{10000}{T} + \frac{c_3}{P} + c_4 X_{Ca}^{cpx} c_5 D_{Ti}^{cpx/\ell} + c_6 (Al + Si - Na - K)$$

where *T* is temperature (in kelvins), *P* is pressure (in kbar), X_{Ca} is the mole fraction of clinopyroxene octahedral sites occupied by Ca, D_{Ti} is the clinopyroxene–melt partition coefficient for Ti, (Al + Si – Na – K) refer to the mole fraction of the corresponding oxides in the melt, and $c_1 \dots c_6$ are regression coefficients. This equation is considerably more complex than equation 7.30, but is only marginally more accurate. It does have the advantage of including the pressure term.

Beattie (1993) used a modification of the two-lattice model to predict the partitioning of Mg, Fe, Mn, Co, and Ni between olivine, orthopyroxene, and melt as a function of temperature and composition. Beattie relaxes the assumption that the network-forming lattice is ideal and introduces an empirical function to calculate the activity of SiO₂. He then calculated values of ΔH_m , ΔV_m , and ΔS_m from experiments on mineral melt pairs. The resulting thermodynamic data were used to calculate the equilibrium constant for the exchange reaction, e.g.,

$$2MO^{\ell} + Mg_2SiO_4^{ol} \rightleftharpoons 2MgO^{\ell} + M_2SiO_4^{ol}$$

where M is Fe, Mn, Co, etc. The equilibrium constant is:

$$\mathbf{K} = \left(\frac{D_{M}^{\alpha/\ell} \lambda_{MO}^{\alpha} \lambda_{MgO}^{\ell}}{D_{Mg}^{\alpha/\ell} \lambda_{MgO}^{\alpha} \lambda_{MO}^{\ell}}\right)^{2} \exp\left(\frac{\Delta S}{R} - \frac{\Delta H + (P - P^{\circ})\Delta V}{RT}\right)$$
7.40

where γ is the activity coefficient, ΔS , ΔH , and ΔV refer to the difference in entropy, enthalpy, and volume changes of fusion between the Mg version and the M version of phase α (e.g., $\Delta S = \Delta S_m^{M-\alpha} - \Delta S_m^{MgO-\alpha}$).

7.4.4 Mineral-Liquid Partition Coefficients for Mafic and Ultramafic Systems

As we have seen, partition coefficients depend on temperature, pressure, and the composition of the phases involved. There are nevertheless circumstances when a general set of partition coefficients is useful. For example, temperature or melt composition may not be known, or great accuracy may not be needed. Table 7.5 is a set of mineral–melt partition coefficients for mafic and ultramafic magmas. They have been assembled from a variety of literature sources and adjusted where necessary to make them self-consistent. The assembly of any set of partition coefficients will always be somewhat subjective. The following principles have guided selection of this partition coefficient set.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

Example 7.2. Calculating Partition Coefficients

Using the composition of the basalts given below, calculate the clinopyroxene-matrix partition coefficients for La and Sm using the Gallahan and Nielsen equations and the coefficients given in Table 7.4. Assume a temperature of 1080° C.

			Answer: Once again this is a problem that is best done in a
	Mauna Loa	Mount Hop	pespreadsheet. Our first task will be to convert the wt. percent
Oxide	t holeiite	alkali basal	It to mole fractions. We calculate the moles of each component
	wt %	wt %	(Gallahan and Nielsen chose their components as single ca-
SiO.	49 56	50 95	tion oxides, except for P_2O_5) by dividing by the molecular
TiO_2	4 28 1 38	50.75	weight. We convert this to mole fractions by dividing by the
ALO.	14 09	17 87	sum of the moles of all components. Then we calcul ate the
FeO	12 47	8 64	sums of the network formers and network modifiers
MnO	0.22	0.01	(equations 7.33 and 7.36), and finally the activities of SiO_2
MaO	4.62	6 1 8	and AlO _{1.5} using equation 7.34. We calculate K using equa-
$C_{2}O$	9.63	11 29	tion 7.28 (temperature in kelvins, as in all thermodynami-
	2 0 2	2 2 2	cally based formulae). The d istribution coefficient, D* can
$K_{2}O$	5.05	2.52	then be calculated using equation 7.31. Finally, we convert
K_2O	1.18	0.95	to Dusing equation 7.39.
P_2O_5	1.02	0.51	Though the temperature is the same and the composi-

Though the temperature is the same and the compositions of these magmas are similar (both are basalts), there is

a large difference, almost a factor of 2, in the partition coefficients we have calculated.

Oxide	Mauna Loa			Mt. Hope			Mole. Wt
	wt %	moles	mol fract.	wt %	moles	mol fract.	
SiO2	49.56	0.826	0.45555	50.95	0.8492	0.4553	60
TiO2	4.28	0.0536	0.02955	1.38	0.0173	0.00926	79.88
AI203	14.09	0.3278	0.1808	17.87	0.4158	0.22293	42.98
FeO	12.47	0.1736	0.09572	8.64	0.1203	0.06448	71.85
MnO	0.22	0.0031	0.00171	0.21	0.003	0.00159	70.94
MgO	4.62	0.1146	0.06321	6.18	0.1533	0.0822	40.31
CaO	9.63	0.1717	0.09471	11.29	0.2013	0.10794	56.08
Na2O	3.03	0.0977	0.05391	2.32	0.0748	0.04013	31
K20	1.18	0.0251	0.01382	0.95	0.0202	0.01081	47.1
P2O5	1.02	0.02	0.01102	0.51	0.01	0.00536	51.03
Total	100.1	1.8132	1	100.3	1.8651	1	
NM			0.409			0.44282	
NF			0.52328			0.50624	
aSiO2			0.87058			0.89937	
aAl01.5			0.09311			0.15838	
In(aSiO2	2*aAlO1.5)		-2.51254			-1.94882	
Т	1353						
	La	Sm	Mauna Loa			Mount Hope	e
а	119	11371	InD La	InD Sm		InD La	InD Sm
b	1.95	8.6	-4.37459	-2.708		-3.81087	-2.1445
ln K	-1.86205	-0.196	D* La	D* Sm		D* La	D* Sm
D *			0.01259	0.0667		0.02213	0.11712
D			0.0308	0.163		0.05	0.2645

Geochemistry

CHAPTER 7: TRACE ELEMENTS

- Errors in measurement and experimental limitations are such that measured partition coefficients will be too large more often than too low. Thus low values are generally preferred.
- In situ microanalytical techniques (e.g., ion probe) of experiments on natural basaltic liquids were preferred over bulk analysis of natural samples. Undoped experiments were preferred over doped experiments.
- Knowing relative relationship of the partition coefficient of one element to that of another is more important than knowing absolute values. Accuracy of relative values is maximized when partition coefficients for all elements are measured in a single experiment or series of experiments. Thus studies where partition coefficients for a large number of elements were determined were preferred over more limited set of elements.
- Partition coefficients I measured experimentally using natural basaltic liquids with temperatures in the range of 1200-1400°C were preferred.
- Elements known to have partition coefficients that depend strongly on composition or temperature (e.g., Ni) have not been included.

Figure 7.16 illustrates the rare earth partition coefficients from this data set. In general, the minerals clinopyroxene, garnet and plagioclase and, when present, amphibole (amphibole is not usually present in basalts because it is not stable at low pressure or above $1100 \degree$ C) will control the patterns of incompatible element parti-

IADLI	e 7.5. Minei	RAL-MELT	[PARTITI	ION CO	DEFFICIENT	s for E	SASALTS
	Olivine	Opx	Срх	Plag	Spinel	Garnet	Amph
Li	0.041	0.11	0.59	0.45			
Be	0.035		0.047	0.36			
В	0.034	0.027	0.117		0.08		
K	0.00017		0.0028	0.18		0.002	0.35
Sc	0.37	0.6	0.8	0.065	0.048	0.688	
V	0.3	2.6	1.81	0.04	38	1.48	
Ga	0.024		0.74	0.86	4.6		
Ge	0.097	0.25	1.4	0.51	0.1		
Rb	0.000044		0.0033	0.025		0.007	0.437
Sr	0.000063	0.0068	0.157	2.7		0.0099	0.184
Y	0.0098	0.014	0.62	0.013		5.42	0.634
Zr	0.00068	0.004	0.195	0.001	0.06	2.12	0.3
Nb	0.00005	0.015	0.0081	0.033	0.08	0.0538	0.197
Cs	0.0015		0.0039	0.026			
Ba	0.0000034	0.0067	0.0022	0.33		0.0007	0.282
La	0.0000088	0.0056	0.052	0.082	0.01	0.0164	0.058
Се	0.000019	0.0058	0.108	0.072		0.065	0.116
Pr	0.000049	0.006		0.056			0.178
Nd		0.007	0.277	0.045		0.363	0.273
Sm	0.000445	0.0085	0.462	0.033	0.0064	1.1	0.425
Eu		0.0078	0.458	0.55	0.0061	2.02	0.387
Gd		0.011		0.034			0.725
Гb	0.00324	0.011			0.0078		0.779
Dy		0.015	0.711	0.03		4.13	0.816
Но	0.00927	0.019					0.783
Er		0.021	0.66	0.02		3.95	0.699
Гm		0.025					
Yb	0.0366	0.032	0.633	0.014	0.0076	3.88	0.509
Lu		0.042	0.623			3.78	0.645
Hf	0.001	0.021	0.223		0.05	1.22	0.638
Га	0.00005	0.015	0.013	0.11	0.06	0.11	
Pb	0.0076		0.01	1.07		0.0001	
Гh	0.000052	0.0056	0.014	0.19		0.0014	0.016
U	0.00002	0.015	0.013	0.34		0.0059	

Data sources: Beattie (1994), Chaussidon and Libourel (1993), Chazot et al. (in press), Dunn and Sen (1994), Hart and Brooks (1974), Hart and Dunn (1993), Hauri et al. (1994), Kennedy et al. (1993), Nagasawa et al. (1980), and compilations by Green (1994), Irving (1978), and Jones (1995).
CHAPTER 7: TRACE ELEMENTS

tioning during melting and crystallization of basaltic magmas because they have the highest partition coefficients. Olivine, though by far the most abundant mineral in the mantle, will produce little fractionation[§] of incompatible elements because its partition coefficients are so low. Spinel, which is usually not pres??ent in more than a few volume percent, will also have little effect on relative trace element abundances. On the other hand, olivine largely controls the fractionation of the compatible transition metals.

7.5 Crystal-Field Effects

We pointed out earlier that the ions of the alkalis, alkaline earths and rare earths can be satisfactorily modeled as hard spheres containing point charges at their centers. This model of ionic behavior is notably less satisfactory for many of the transition metals, because of the complex geometry of the bonding



Figure 7.16. Rare earth mineral–melt partition coefficients for mafic magmas. Data from Table 7.5.

electron orbitals, illustrated in Figure 7.17. A more accurate prediction of bonding and substitution requires consideration of the electrostatic field of surrounding ions on the electron structure of transition elements.

Crystal-Field Theory, which was developed in 1929 by physicist Hans Bethe, describes the effects of electrostatic fields on the energy levels of a transition-metal ion in a crystal structure. These electrostatic forces originate from the surrounding negatively charged anions or dipolar groups, known as ligands. Crystal Field Theory is the simplest of several theories that attempt to describe the interaction and bonding between ligands and transition metals. In crystal field theory, the ligands are regarded simply as negative point charges about the transition metal ion. The electromagnetic field produced by these ligands, the "crystal field", destroys the spherical symmetry possessed by an isolated transition metal. The changes induced depend on the type, position, and symmetry of the coordinating ligands, as well as the nature of the transition metal.

In the usual case, electron orbitals are filled successively from inner to outer as one proceeds 'up' the periodic table to heavier elements. In transition metals, however, filling of the outermost s orbital is begun before the d orbitals are completely filled. Ions are formed when the outermost s and, in some cases, some of the outermost d electrons (outermost will be 4s and 3d for the first transition series, 5s and 4d for the second, etc.) are removed from the metal.

[§] *Fractionation*, in this context, refers to a change in the relative abundances, or ratios, of elements. For example, if the ratio of La to Sm changes during a process, these elements are said to have been fractionated by that process.

M. White

Geochemistry

CHAPTER 7: TRACE ELEMENTS



Figure 7.17. Geometry of the *d* orbitals.

In an isolated first series transition metal, the 5 3d orbitals (each containing up to 2 electrons: a total of 10 electrons are possible in the *d* orbitals) are energetically equivalent and have equal probability of being filled: they are said to be degenerate. They possess, however, different spatial configurations (Figure 7.17). One group, the e_{g} orbitals, consisting of the d_{z^2} and the $d_{x^2-y^2}$ orbitals, has lobes directed along the Cartesian coordinates, while the t_{2_0} group, consisting of the $d_{xyy} d_{yyy}$ and d_{xz} possess lobes between the Cartesian axes. In a crystal field the 3d orbitals are no longer degenerate, and some have lower energy than others. Thus there can be a relative energy gain by preferentially filling low-energy *d* orbitals. This energy gain is traded off against the energy cost of placing two electrons in a single orbital. Depending on this trade off, preferentially filling low energy d orbitals can lower the overall energy of some transition metal ions, hence stabilizing them, in certain lattice configurations relative to other configurations. The effect is a lattice site preference of some transition metals that would not be predicted simply from consideration of ion charge and size.

In octahedral coordination, the transition metal

is surrounded by six identical ligands (Figure 7.18). Electrons in all orbitals are repelled by the negatively charged ligands, but electrons in the e_g orbitals, the orbitals directed toward the ligand, are repelled to a greater extent than those in the t_{2_0} orbitals. The energy separation between the t_{2g} and the e_g orbitals is termed octahedral crystal-field splitting and is denoted by Δ_0 . The t_{2g} orbitals are lowered by $2/5\Delta_0$ while the e_g orbitals are raised by $3/5\Delta_0$ relative to the mean energy of an unperturbed ion. Therefore each electron in a $t_{2_{\circ}}$ orbital stabilizes a transition metal ion by $^{2}/_{5}\Delta_{0}$

			High	Spin State		I	Low Spin	n State	
Number	of Ion	Elect	ronic	Unpaired	CFSE	Electron	ic Ūn	paired	CFSE
3d electro	ons	Config	uration	Electrons		Configura	tion Ele	ectrons	
		t _{2g}	eg			t _{2g}	eg		
0	Sc ³⁺ , Ti ⁴⁺	•		0	0			0	0
1	Ti ³⁺	Ϋ́ Υ		1	$2/5\Delta_{o}$	\uparrow		1	$2/5\Delta_{o}$
2	Ti ²⁺ , V ³⁺	Ϋ́		2	$4/5\Delta_{o}$	\uparrow \uparrow		2	$4/5\Delta_{\rm o}$
3	V^{2+} , Cr^{3+} , Mn^{4+}	T T T		3	$6/5\Delta_{o}$	$\uparrow \uparrow \uparrow$		3	$6/5\Delta_{o}$
4	Cr^{2+} , Mn^{3+}	$\uparrow \uparrow \uparrow$	\uparrow	4	$3/5\Delta_{o}$	$\uparrow \downarrow \uparrow \uparrow$		2	$8/5\Delta_{\rm o}$
5	$Mn^{2+} Fe^{3+}$	$\uparrow \uparrow \uparrow$	\uparrow \uparrow	5	0	$\uparrow \downarrow \uparrow \downarrow \uparrow$		1	$10/5\Delta$
6	Fe ²⁺ , Co ³⁺ ,Ni ³⁺	$\uparrow \downarrow \uparrow \uparrow$	\uparrow \uparrow	4	$2/5\Delta_{o}$	$\uparrow \downarrow \uparrow \downarrow \uparrow \downarrow \downarrow \downarrow$		0	$12/5\Delta_{o}^{\circ}$
7	Co^{2+}, Ni^{3+}	$\uparrow \downarrow \uparrow \downarrow \uparrow$	\uparrow \uparrow	3	$4/5\Delta_{\rm o}$	$\uparrow \downarrow \uparrow \downarrow \uparrow \downarrow \downarrow \downarrow$	\uparrow	1	$9/5\Delta_{o}$
8	Ni ²⁺	$\uparrow \downarrow \uparrow \downarrow \uparrow \downarrow \downarrow \downarrow$	$\uparrow \uparrow$	2	$6/5\Delta_{o}$	$\uparrow \downarrow \uparrow \downarrow \uparrow \downarrow \downarrow \downarrow$	\uparrow \uparrow	2	$6/5\Delta_{o}$
9	Cu^{2+}	$\uparrow \downarrow \uparrow \downarrow \uparrow \downarrow \downarrow \downarrow$	$\uparrow\downarrow\uparrow$	1	$3/5\Delta_{o}$	$\uparrow \downarrow \uparrow \downarrow \uparrow \downarrow \downarrow \downarrow$	$\uparrow\downarrow\uparrow$	1	$3/5\Delta_{o}$
10	Zn^{2+}	$\uparrow \downarrow \uparrow \downarrow \uparrow \downarrow$	$\uparrow \downarrow \uparrow \downarrow$	0	0	$\uparrow \downarrow \uparrow \downarrow \uparrow \downarrow$	$\uparrow \downarrow \uparrow \downarrow$	0	0

Table 7.6. Electronic Configurations and Crystal Field Stabilization Energies of First Transition Series Metal Ions in Octahedral Configura-TION

Geochemistry

CHAPTER 7: TRACE ELEMENTS



Figure 7.18. (a) Orientation of ligands and Cartesian coordinates for a metal ion in octahedral coordination. (b) Orientation of ligands (points) and d_{xy} (shaded) and $d_{x^2-y^2}$ (unshaded) orbitals in the x-y plane of a metal in octahedral coordination. (c) Energy level diagram for d orbitals of a free transition metal ion, and ion in a spherically symmetric crystal field and an octahedral crystal field. In an octahedral crystal field, the energy of the *d* orbitals projected between the coordinates and the ligands (the t_{2g} orbitals) are lowered relative to the energy of the orbitals projected toward the ligands (e_g orbitals). From Burns (1970).

and each electron in an e_g orbital diminishes stability by ${}^3/{}_5\Delta_o$. The resultant net stabilization energy, i.e., $\Sigma\Delta_o$, is called the *octahedral crystal field stabilization energy* or octahedral CFSE.

How electrons are distributed in an octahedrally coordinated transition metal is governed by two opposing tendencies. Coulomb forces between electrons cause them to be distributed over as many orbitals as possible, but crystal-field splitting favors the occupation of lowest energy orbitals. These two factors in turn depend on the number of *d* electrons possessed by the metal and the strength of the crystal field.

In ions having 1, 2 or 3 *d* electrons all electrons will be in only t_{2a} orbitals, regardless of the strength of the crystal field, since there is only 1 electron per orbit. In ions having 8, 9, or 10 electrons in *d* orbitals, each orbital must contain at least 1 electron and three orbitals must contain 2 electrons, so all the 3 t_{2a} orbitals will be filled even for weak ligands. But in ions having 4, 5, 6 and 7 d electrons there is a choice. If the crystal-field splitting is large, as it is in the case of strong field ligands, all electrons are in t_{2a} orbitals. This is the *low-spi*n case, because the spins of electrons are antialigned (recall the Pauli exclusion principle that electron can only occupy the same orbit if their spins are When the crystal-field opposite).

splitting is small, the energy cost of placing two electrons in the same orbit is greater than the energy gain from the octahedral CFSE, and electrons are distributed over both t_{2g} and e_g orbitals. This is known as the *high spin* case because the electrons will preferentially occupy different orbitals with their spins parallel. Apart from Co³⁺ and Ni³⁺, all first series transition metals exist in the high-spin state on the Earth's surface. Cr³⁺, Ni²⁺, and Co³⁺ have particularly high CFSE in octahedral coordination.

The distinction between high and low spin configurations is important in understanding magnetic properties of transition metal compounds because magnetism relates to spin alignment of electrons. Also, the crystal-field splitting energies are in the visible light band and hence relate to the coloration of transition-metal-bearing minerals. For example, consider titan-augite (a variety of clinopyroxene containing appreciable amounts of Ti). In Ti⁺³, the single *d* electron is normally in the t_{2g} orbital. Absorption of light of appropriate frequency $v = \Delta_0/h$ excites the electron into an e_g orbital. This energy corresponds to violet light, which is emitted when the electron returns to the t_{2g} orbital.

In tetrahedral coordination (Figure 7.19), the e_g orbitals become more stable than the t_{2g} orbitals, but the tetrahedral crystal-field splitting parameter, Δ_t , is smaller than Δ_o . Other things being equal,

Geochemistry

W. M. White

CHAPTER 7: TRACE ELEMENTS

 $\Delta_t = 4/9\Delta_o$.

The crystal field splitting parameter, Δ , depends on a number of things, but some generalizations can be made:

1. Values of Δ are higher for +3 ions than +2 ions.

2. The values are 30% higher for each succeeding transition element (left to right in the periodic table).

3. Δ depends on the ligands coordinating the transition metal. Δ may be arranged to increase as follows: halides $< OH^-$ (hydroxides) $< silicates < H_2O < SO_4$.

4. Δ depends on the symmetry of the coordinating ligands (as we have seen).

5. Δ varies with interatomic distance (inversely proportional to the 5th power).

In silicates, oxygen atoms frequently form distorted coordination polyhedra about cations and metal-oxygen interatomic distances are not constant. Distortion of coordinated polyhedra is expected from the *Jahn-Teller Theorem*, which states that if the ground state of a molecule is degenerate, it will distort to make one of the energy levels more stable (Figure 7.20). For example, if one of the *d* orbitals is filled or empty while another is half-filled, the environment about the transition metal ion is predicted to distort spontaneously to a different geometry in which a more stable



Figure 7.19. (a) Orientation of ligands and Cartesian coordinates for a metal ion in tetrahedral coordination. (b) Orientation of ligands (points) and d_{xy} (shaded) and $d_{x^2y^2}$ (unshaded) orbitals in the *x*-*y* plane of a metal in tetrahedral coordination. Black points are in front of plane of orbitals, gray points are behind the plane. Energy level diagram for *d* orbitals of a free transition metal ion, and ion in a spherically symmetric crystal field and a tetrahedral crystal field. From Burns (1970).

electronic configuration is attained by making the occupied orbital lower in energy. Further splitting of *d* orbital energy levels occurs when transition metal ions are in distorted coordination.

This can be illustrated for the case of Mn^{3+} in octahedral coordination with oxygen. The Mn^{3+} ion has the high spin configuration (Table 7.6) in which each t_{2g} orbital is occupied by 1 electron and the fourth electron may occupy either of the e_g orbitals. If the four oxygen atoms in the *x*-*y* plane move toward, and the two oxygens along the *z*-axis move away from, the central Mn^{3+} ion (Figure 7.20a), then the one e_g electron will favor the d_{z2} orbital in which the repulsion by the O ions is smaller than in the $d_{x^2-y^2}$ orbital. Thus the e_g orbital group separates into two energy levels with the d_{z2} becoming more stable. The t_{2g} orbital group is also split, with the d_{xz} and d_{yz} becoming more stable than the d_{xy} orbital. If the two O ions along the *z*-axis move closer to the Mn^{3+} ion (Figure 7.20c), the $d_{x^2-y^2}$ becomes more stable than the d_{z2} . In either of the distorted environments, the Mn^{3+} becomes more stable than in an undistorted octahedral site. Transition metals subject to Jahn-Teller distortions in octahedral coordination are those with d^4 , d^9 , and low-spin d^7 configurations, in which one or three electrons occupy e_g orbitals. Looking at Table 7.6, we can see that these are Cr^{2+} , Mn^{3+} , Cu^{2+} , Co^{2+} and Ni^{3+} ions.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

As noted, electronic transition energies are related to color. Because of the distortion, additional electronic transitions become possible. The differing probabilities of the various electronic transitions in polarized radiation in one of the causes of *pleochroism*^{*} in minerals.

Crystal-field effects lead to irregularities in the interatomic distances, or ionic radii of transition metals. As you might expect, they depend on the nature of the site, and for a given site, there need not be a smooth contraction of interatomic distances with increasing atomic number.

The application of crystal field theory is restricted to compounds where the transition metal forms a dominantly ionic bond with surrounding ligands. In sulfides, and related minerals, the effects of co-valent bonding, in which the orbitals become hybridized, must be considered, but such consideration is beyond the scope of our treatment.

7.5.1 Crystal Field Influences on Transition Metal Partitioning

We can now return to the transition metals and crystal-field theory to explain some of the peculiarities of their behavior. We noted that the energy of some *d* orbitals is reduced (stabilized) by the effects of the electrostatic field of coordinating ligands in both octahedral and tetrahedral sites and that the octahedral CFSE is always greater than the tetrahedral CFSE. We now introduce one more quantity: the octahedral site preference energy (OSPE), which is defined as the octahedral CFSE less the tetrahedral



Figure 7.19. Arrangement of ligands and energy levels for (a) an octahedral site distorted along the z axis. (b) an undistorted site, and (c) an octahedral site distorted along the y axis. After Burns (1970).

^{*} Pleochroism refers to property possessed by some crystals of exhibiting different colors when viewed along different axes in polarized light.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

CFSE. Table 7.7 lists these energies for various first transition series metals. Silicate melts contain both octahedral and tetrahedral sites, but transition metals, with a few exceptions, occupy only octahedral sites in silicate minerals. Thus the OSPE is a indicator of the preference of the transition ion for solid phases over liquid phases: the higher the OSPE, the more readily it is partitioned into a solid phase. Predicted order of uptake is: Ni>Co>Fe⁺²>Mn for 2+ ions and Cr>V>Fe⁺³ for 3+ ions, which agrees with observation. Since the number of octahedral sites in the liquid decrease with increasing SiO₂ concentration, crystal field theory explains why Ni partition coefficients are highly composition dependent, increasing with increasing SiO₂ concentration.

It should be emphasized that there are no crystal field effects for transition metals such as Sc^{2+} , Ti^{+4} , Y^{3+} , Zr^{4+} , Nb^{5+} , Hf^{4+} , and Ta^{5+} where the *d* electrons are not present in the ion, or where the *d* shell is completely filled (Zn^{2+}) in the usual valence state, at least when the electrons are in their ground state. However color, which arises from excitation of electrons into higher orbitals and subsequent decay to the ground state, may still relate to crystal field effects even when the *d* orbitals are not filled in the ground state. The second and third transition series metals for which crystal-field effects are expected are all highly siderophile or chalcophile and highly depleted in the Earth's crust and mantle. Little information is available on their behavior in silicate systems.

An understanding of crystal field theory solves an interesting dilemma. A phase (T-X) diagram for the binary system $Mg_2SiO_4 - Ni_2SiO_4$ is shown schematically in Figure 7.21. It is apparent from a quick glance that for any coexisting liquid and solid in the system, the solid will be poorer in Ni than the liquid, i.e., $(Ni/Mg)_{ol} < (Ni/Mg)_{liq}$. However, olivine crystallizing from basaltic liquids is always richer in Ni than the liquid. The reason for this is that in the pure olivine system, only octahedral sites exist in the melt and the solid, and thus Ni has no particular preference for the solid due to crystal field effects. But basaltic melts have both tetrahedral and octahedral sites, while olivine has only octahedral sites (available to Ni). The greater availability of octahedral sites in the solid provides an added incentive for Ni to partition into olivine relative to basaltic liquid.

Number of					CFSE	Octahedral	Tetrahedral	Octahedral
3 <i>d</i>		Electro	nic	Δ (cm ⁻¹)	hydrate	CFSE	CFSE	Site Preference
electrons	Ion	Configur	ation	$M(H_2O)_{6 (aq)}^{n+}$	(KJ/mole)	(kJ/mole)	(kJ/mole)	Energy (kJ)
1	Ti^{3+}	$(t_{2g})^1$		20,300	$2/5\Delta = 97.1$	87.4	58.6	28.9
2	V^{3+}	$(t_{2g})^2$		17,700	$4/5\Delta = 169.0$	160.2	106.7	53.6
3	Cr^{3+}	$(t_{2_g})^3$		17,400	$6/5\Delta = 249.4$	224.7	66.9	157.7
4	Cr^{2+}	$(t_{2g})^3$	$(e_{g})^{1}$	13,900	$3/5\Delta = 99.6$	100.7	29.3	71.1
4	Mn^{3+}	$(t_{2_g})^3$	$(e_{g})^{1}$	21,000	$3/5\Delta = 150.6$	135.6	40.2	95.4
5	Mn^{2+}	$(t_{2_g})^3$	$(e_{g})^{2}$	7,800	0	0	0	0
5	Fe^{3+}	$(t_{2g})^3$	$(e_{g})^{2}$	13,700	0	0	0	0
6	Fe^{2+}	$(t_{2_g})^4$	$(e_{g})^{2}$	10,400	$2/5\Delta = 49.8$	49.8	33.1	16.7
6	Co ³⁺	$(t_{2_g})^6$		18,600	$12/5\Delta = 533.5$	* 188.3	108.8	79.5
7	Co ²⁺	$(t_{2_g})^5$	$(e_{g})^{2}$	9,300	$4/5\Delta = 89.1$	92.9	61.9	31.0
7	Ni ³⁺	$(t_{2_g})^6$	$(e_{g})^{1}$	-	$9/5\Delta =$			
8	Ni ²⁺	$(t_{2_g})^6$	$(e_{g})^{2}$	8,500	$6/5\Delta = 29.6$	122.2	36.0	86.2
9	Cu^{2+}	$(t_{2_g})^6$	$(e_{g})^{3}$	12,600	$3/5\Delta = 21.6$	90.4	26.8	63.6

Table 7.7. Crystal-Field Splittings and Stabilization Energies in Transition-Metal Ions

* Low-spin complexes. The calculated CFSE must be reduced by the energy required to couple two electrons in a t_{2q} orbital. Data from Orgel (1966) and McClure (1957).

November 11, 2005

Geochemistry

CHAPTER 7: TRACE ELEMENTS



Figure 7.21. Schematic phase diagram for the system forsterite—Ni olivine showing Ni-poor olivine in equilibrium with Ni-rich liquid.

7.6 TRACE ELEMENT DISTRIBUTION DURING PARTIAL MELTING

In igneous geochemistry, trace elements are useful in understanding magmatic processes and in evaluating the composition of magma sources such as the mantle and lower crust. To make use of trace elements in such studies, we need to understand how magmatic processes such as partial melting and fractional crystallization will affect trace element abundances.

The task of the igneous geochemist is often to make inferences about the sources of magma, the mantle and lower crust, from the composition of the magmas themselves. This can be done through a mathematical model of the melting. In the following sections, we will consider two simple alternative models of melting: batch, or equilibrium, melt-

ing, and fractional melting. In fractional melting, the melt is extracted as soon as it is created, and only an infinitesimal increment of melt will be in equilibrium with the solid residue at any given time. In batch melting, a finite amount of melt, for example 5 or 10%, is produced and equilibrates completely with the solid residue.

Once a melt is created and begins to rise, it may further interact with the surrounding "wallrock". We will also consider one possible model of this interaction: "zone refining". Choosing between alternative models of partial melting requires a knowledge of how melting and melt extraction actually occurs. Unfortunately, melting and melt extraction in the Earth remain poorly understood because we are unable to observed them directly. Although melting experiments are useful in determining phase relationships, melting temperatures, and distribution coefficients, they do not provide much direct information on how melt is extracted. By and large, our knowledge of the melt extraction process comes from indirect inferences. Rarely, we can identify partial melting residues that have been tectonically emplaced at the surface of the Earth, and studies of these have provided some insights into the melting process. We will consider some of these insights in a subsequent section.

7.6.1 Equilibrium or Batch Melting

Equilibrium crystallization or melting implies complete equilibration between solid and melt. This means that the entire batch equilibrates with the residue before it is removed. From mass balance we may write:

$$C_i^o = C_i^s (1 - F) + C_i^{\ell} F$$
 7.41

where *i* is the element of interest, C° is the original concentration in the solid phase (and the concentration in the whole system), C^{ℓ} is the concentration in the liquid, C° is the concentration remaining in the solid and F is the melt fraction (i.e., mass of melt/mass of system). Since $D = C^{\circ}/C^{\ell}$,

and rearranging:

$$C_{i}^{o} = C_{i}^{\ell} D_{i}^{s/\ell} (1 - F) + C_{i}^{\ell} F$$

$$\boxed{\frac{C_{i}^{\ell}}{C_{i}^{o}} = \frac{1}{D^{s/\ell} (1 - F) + F}}$$
7.42

This equation is an extremely useful one and describes the relative enrichment or depletion of a trace element in the liquid as a function of degree of melting. Two approximations are often useful and give us a feel for this equation. First consider the case where D<<F. In this case $C^{\ell}/C^{\circ} \approx 1/F$, that is, the enrichment is inversely proportional to the degree of melting. This is the case for highly incompatible elements at all but the smallest degrees of melting. Now consider the case where F approaches 0. In this

Geochemistry

CHAPTER 7: TRACE ELEMENTS

case $C^{\ell}/C^{\circ} \approx 1/D$, the enrichment is inversely proportional to the partition coefficient. Thus the maximum enrichment possible in a partial melt is 1/D. For highly compatible elements, that is those with large D such as Ni, the depletion in the melt is 1/D when F is small and is relatively insensitive to F.

7.6.2 Fractional Melting

Now consider the case where only an infinitesimally small amount of melt equilibrates with the solid residue, in other words, imagine we remove the liquid as fast as we make it. If i^s is the mass of element *i* in the solid phase being melted, S the mass of the solid phase, L the mass of the liquid phase, i^i the mass of *i* in the liquid, S° the original mass of the solid (and mass of the system), and i° the original mass of *i* in the solid (and system), then:

$$C_i^s = \frac{i^s}{S} = \frac{i^\circ - i^t}{S^\circ - L} \qquad \text{and} \qquad C_i^\ell = \frac{1}{D_i} \frac{i^\circ - i^\ell}{S^\circ - L} = \frac{di^\ell}{dL}$$

This can be rearranged so that we can integrate it to obtain:

$$\frac{C_i^{\ell}}{C_i^o} = \frac{1}{D} (1 - F)^{1/D - 1}$$
7.43

If we subsequently mix the various melt fractions produced over a melt interval from 0 to F, the composition of this *aggregate liquid*, \overline{C} is:

$$\frac{\overline{C}_{i}^{\ell}}{C_{i}^{o}} = \frac{1}{F} \left(1 - (1 - F)^{1/D_{i}} \right)$$
7.44

Figure 7.22 illustrates the variation of the liquid enrichment (C^{ℓ}/C°) with degree of melting for both batch and fractional melting. The aggregate liquid of fractional melting, which may be the most realistic of the 3 equations we have considered so far, follows a trend close to that of batch melting.

7.6.3 ZONE REFINING

If melt percolates slowly through the source region, trace element fractionation may be best approximated by equations governing zone refining. The term zone refining comes from the industrial purifi-

cation process in which a melt zone is caused to migrate along a column of material. Several passes of this process efficiently extract low melting-temperature components. The relevant equation is:

$$\frac{C_i^{\ell}}{C_i^o} = \frac{1}{D_i} - \left(\frac{1}{D_i} - 1\right)e^{-D_i R} 7.45$$

where R is the ratio of host, or wallrock, to melt. Note that for large R, $C^l/C^{\circ} \sim 1/D$.

7.6.4 Multiphase Solids

The above equations are relevant when the solid undergoing melting is homogeneous. If it consists of several phases, we need to replace D with a *bulk* distribution coefficient, which is simply the weighted mean of the



Figure 7.22. Variation in C^{ℓ}/C^{0} with degree of melting, *F*, for various partition coefficients for batch and fractional melting.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

. M. White

individual solid/liquid partition coefficients:

$$\overline{D}_i = D_i^o = \sum_{\phi} m_{\phi} D_i^{\phi/\ell}$$
 7.46

where *m* is simply the *modal mass fraction* of phase ϕ , that is the fraction of phase ϕ as it exists in the rock.

In general, minerals do not enter the melt in the proportion in which they exist in a rock, i.e., their modal proportions. Thus a realistic melting model requires we modify our equations to account for this. We need to define a new parameter P, which is simply the average individual partition coefficients weighted according the proportions in which the minerals enter the melt:

$$P_i = \sum_{\phi} p_{\phi} D_i^{\phi/\ell}$$
 7.47

where *p* is the proportion in which phase ϕ enters the melt. It is often assumed that the p_{ϕ} values are constants. In reality, they will be functions of *F*. The equations for batch and fractional melting in this *non-modal melting* case become:

 $\frac{C_i^{\ell}}{C_i^{o}} = \frac{1}{D_i^{o}} \left(1 - \frac{P_i F}{D_i^{o}}\right)^{1/P_i - 1}$

Non-Modal Batch Melting

$$\frac{C_i^{\ell}}{C_i^{o}} = \frac{1}{F(1-P_i) + D_i^{o}}$$
7.48

The aggregate liquid for non-modal fractional melting is given by:

Aggregate:

$$\frac{\overline{C}_{i}^{\ell}}{C_{i}^{o}} = \frac{1}{F} \left(1 - (1 - \frac{P_{i}F}{D_{i}^{o}})^{1/P_{i}} \right)$$
7.50

These equations are from Shaw (1970).

Example 7.3. Modelling Partial Melting

Geochemists are often interested in the ratios of incompatible trace elements. For one thing, ratios of incompatible elements are less affected by fractional crystallization than are trace element abundances. To illustrate the effect of partial melting on a trace element ratio, we will calculate how the ratio of La to Sm varies in basalts produced by different extents of melting in the mantle.

We need to make some assumptions about the composition and mineralogy of the mantle. These assumptions are: (1) the mantle 'source' is composed of 58% olivine, 27% orthopyroxene, 12% clinopyroxene, and 3% spinel (2) these minerals enter the melt in the proportions 20% olivine, 25% orthopyroxene, 45% clinopyroxene, and 10% spinel, (3) that this source contains 1 ppm La and 1 ppm Sm. Using the partition coefficients for peridotite mantle listed below, we calculate the La and Sm concentrations at 2% and 10% melting using the batch melting model.

Our first step is to calculate the D_o and P values for each element using equations 7.46 and 7.47. Doing so, we find $D_{La} = 0.007$, $D_{Sm} = 0.047$, $P_{La} = 0.025$, $P_{Sm} = 0.166$. Using equation 7.48, we find that at 10% melting, [La] = 9.5 ppm and [Sm]= 7.66 ppm. So a 10% melt (F = 0.1) of a mantle having La/Sm = 1 will have La/Sm = 9.5/7.66 = 1.25. For the same calculation at 2% melting we obtain [La] = 37.3 ppm and [Sm] = 15.65 ppm and La/Sm = 2.38. Thus at a fairly large degree of melting (10%), the La/Sm ratio was only 25% greater than that of the source. At a small degree of

PARTITION COEFFICIENTS

melting, the La/ Sm ratio was m ore than a factor of 2 greater than that of the source.

	Ol	Opx	Срх	spinel
La	0.0003	0.002	0.053	0.01
Sm	0.0013	0.011	0.36	0.01

7.49

CHAPTER 7: TRACE ELEMENTS

7.6.5 Continuous Melting

In most circumstances, the way in which rock melts in the Earth is probably intermediate between our batch and fractional melting models: only part of the melt is extracted continuously, some fraction remains to fill the pore spaces between the mineral grains. This process has been called *continuous melting*. Let's look at how we can modify our fractional melting equation (equation 7.43) for this situation.

Consider a rock undergoing melting. We assume that it has a melt-filled porosity of ϕ , where ϕ is defined by mass. We can replace the partition coefficient in equation 7.43 with an effective partition coefficient, D', which takes account of a fraction of liquid, ϕ , in the rock with a partition coefficient of 1 (Albarède, 1995). Equation 7.43 thus becomes:

$$\frac{C_i^{\ell}}{C_i^{o}} = \frac{1}{D_i^{\prime}} (1 - F)^{1/D_i^{\prime} - 1}$$
 7.51



Figure 7.23. Comparison of continuous and fractional melting for D = 0.0001 and ϕ = 0.01. The aggregate melt is similar in both cases when F is greater than about 2%. A separate curve for continuous melting is shown for ϕ = 0.001.

F in this case is the fraction of melt extracted, which differs from the total amount of melt produced by an amount equal to ϕ . *D*['] is related to the usual partition coefficient, *D*, (equation 7.1) as:

$$D'_{i} = (1 - \phi)D_{i} + \phi$$
 7.52

The exponential term in 7.51, $1/D^{-1}$, is related to *D* by:

$$\frac{1}{D_i'} - 1 = \frac{1 - \phi}{(1 - \phi)D_i + \phi} (1 - D)$$
7.53

Substituting these back into equation 7.51, our expression for continuous melting written in terms of the usual partition coefficient is:

$$\frac{C_i^{\ell}}{C_i^o} = \frac{1}{(1-\phi)D_i + \phi} (1-F)^{\frac{(1-\phi)(1-D_i)}{(1-\phi)D_i + \phi}}$$
7.54

Porosity is normally defined in terms of volume, but the above equations use a porosity defined in terms of mass. The relationship between the mass porosity and the volume porosity is:

$$\phi = \frac{\rho_{\ell} \varphi}{\rho_{s} (1 - \varphi) + \rho_{\ell} \varphi}$$

$$7.55$$

where ϕ is the mass porosity, ϕ volume porosity, ρ_s is the density of the solid and ρ_ℓ is the density of the liquid.

We can also derive an equation for an aggregate continuous melt simply by replacing D with D' in equation 7.44. Figure 7.23 compares continuous and fractional melting for D = 0.0001 and ϕ = 0.001. Leaving residual melt in the pores has the effect of buffering the depletion of the solid, so that the concentration of an incompatible element does not decrease as fast in the case of continuous melting as for fractional melting. As Figure 7.23 shows, for high values of F, the aggregate melts produced by fractional and continuous melting have almost identical compositions. The compositions of the residual solids, however, will be far different, with the residue of fractional melting being far more depleted in incompatible elements than the residue of batch melting.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

7.6.6 CONSTRAINTS ON MELTING MODELS

To summarized the discussion above, we may say that the concentration of a trace element in a melt is a function of (1) the solid phases (i.e., minerals) present in the system before and during melting (2) the extent of melting (i.e., F), (3) the manner of melting (e.g., fractional vs. batch), (4) the concentration of the element in the original solid (i.e., C°), and (5) the partition coefficients. The partition coefficients, as we have seen, are functions of temperature, pressure, and composition of the phases involved. Two tasks of trace element geochemistry are to deduce something about the melting process and about the composition of the source of magmas. If we are to use trace elements for these purposes, it is essential we independently constrain some of these variables.

Most magmas are generated by partial melting of the upper mantle. Although temperature increases with depth in the mantle, the solidus temperature (i.e., the temperature where melting begins) increases more rapidly, so that the deep mantle is generally well below its solidus[†]. Though they have played a very important role in the evolution of the Earth, magmas produced by melting of the deep crust are very much rarer. So our discussion here will be limited to the melting process in the upper mantle. The phases present in the upper mantle, and their compositions, are discussed in more detail in Chapter 11, so we will omit that topic from the discussion here.

7.6.6.1 Relationship Between Melt Fraction and Temperature and Pressure

We can shorten our list of variables if we can somehow relate the degree of melting to temperature and ultimately to pressure. We can do this through a simplified thermodynamic analysis.

Most melting in the mantle, with the notable and important exception of subduction zones, appears to result from decompression: packets of mantle moving upward. Pressure in the Earth is related to depth, (h), by the simple relationship:

$$\frac{dP}{dh} = \rho g \tag{7.56}$$

where ρ is density and *g* is the acceleration of gravity. For a typical upper mantle density, pressures increases by about 1 GPa every 35 km depth.

Because of the scales generally involved (kilometers to hundreds of kilometers) and the low thermal conductivity of rock, it is reasonable to assume that a rising packet of mantle is adiabatic. As we learned Chapter 2, this means it can do work or have work done on it, but it does not exchange heat with its surroundings, i.e., dQ = 0. We also learned in Chapter 2 that an adiabatic system is an isoentropic one, i.e., dS = 0. The constraint that the system is isoentropic allows us to relate the amount of melting that will occur to the temperature and pressure of the rising mantle.

The variation of entropy with temperature and pressure can be expressed as:

$$dS = \left(\frac{\partial S}{\partial T}\right)_P dT + \left(\frac{\partial S}{\partial P}\right)_T dP$$
7.57

Substituting equations 2.105 and 2.106 into 7.57, we have:

$$dS = \frac{C_P}{T} dT - \alpha V dP$$
 7.58

Since the system is isoentropic, dS = 0, and we can solve 7.58 to obtain:

$$\left(\frac{\partial T}{\partial P}\right)_{S} = \frac{T\alpha V}{C_{P}}$$

$$7.59$$

This equation describes the *adiabat*, the P-T path that adiabatically rising mantle follows. (By the way, we can see that the adiabat will be curved, since its slope depends on *T*). The solidus temperature will

⁺ Recent seismic studies suggest the possible presence of melt pockets in the lowermost mantle, near the core-mantle boundary.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

also change with pressure. It's slope is given by the *Clapeyron Equation* (equation 3.3):

$$\left(\frac{dT}{dP}\right)_{sol} = \frac{\Delta V_m}{\Delta S_m}$$

The slope of the solidus is steeper than that of the adiabat, so that rising mantle will eventually intersect the solidus (Figure 7.24). For simplicity, let's assume the solid consists of a single phase. When the solidus is reached, the system will consist of two phases, solid and liquid, and we can write one version of equation 7.57 for the solid and one version for any melt that has formed. Now let's specify that the two phases coexist at equilibrium along a univariant reaction curve, whose slope in T-P space is $(dT/dP)_{2\phi}$. We can solve equation 7.57 to determine how entropy of each phase changes with pressure, for example, for the solid:

$$\frac{dS^s}{dP} = \frac{C_P^s}{T} \left(\frac{dT}{dP}\right)_{2\phi} - \alpha^s V^s$$
7.60

The total specific entropy (i.e., entropy per unit mass) of the system, S_{o} , can be expressed as the sum of the entropy of the solid and the melt.

$$FS^{\ell} + (1 - F)S^{s} = S_{0}$$
 7.61

where S^{ℓ} and S^{s} are specific entropies of the melt and solid respectively and *F* is the fraction of melt. If we solve 7.61 for *F*, we have:

$$F = \frac{S_0 - S^s}{S^\ell - S^s}$$
 7.62

The term S^{ℓ} - S^{s} is just entropy of melting, ΔS_{m} so 7.61 can be written as:

$$F = \frac{S_0 - S^s}{\Delta S_m}$$
7.63

So long as the melt is not extracted the system remains isoentropic and S_0 is a constant; however, neither S^ℓ nor S^s are necessarily constant. Let's assume for the moment, that ΔS_m is also constant (equivalent to assuming that S^ℓ and S^s change in an identical way). If we now differentiate 7.63 with respect to pressure, we have:

$$\left(\frac{\partial F}{\partial P}\right)_{S} = \frac{1}{\Delta S_{m}} \left(\frac{C_{P}^{s}}{T} \left(\frac{dT}{dP}\right)_{2\phi} - \alpha^{s} V^{s}\right)$$
 7.64

Equation 7.64 shows that even assuming that all the thermodynamic parameters in 7.64 are constant, the amount of melt produced by rising mantle will be a function of its temperature.

Once melting begins, rising mantle follows a P-T path that is steeper than adiabatic (Fig. 7.24), since some energy is consumed in melting. Let's call the temperature that the system would have attained had melting not occurred the potential temperature, T_{pot} . The difference between that temperature and the actual temperature *T* is related to the entropy change during melting ΔS_m . We can determine the entropy change due to the difference between *T* and T_{pot} by integrating equation 2.105:

$$\Delta S = \int_{T}^{T_{pot}} \frac{C_{P}}{T} dT$$
 7.65

Since we are interested in a simple, approximate analysis, let's



Figure 7.24. Representation of melting of an ascending packet of mantle in temperature and pressure space. Below the solidus, the mantle rises along the adiabat. Once packet intersects the solidus, the T-P path of the mantle packet is deflected by ΔT .

Geochemistry

CHAPTER 7: TRACE ELEMENTS

assume that C_P is constant. In that case, equation 7.65 becomes:

$$\Delta S == C_p \ln \frac{T_{pot}}{T}$$
7.66

To find a simple linear solution, let's approximate 7.66 with a Taylor Series expansion about T_{act} , which yields:

$$\Delta S \cong \frac{C_p}{T} (T_{pot} - T)$$
7.67

So long as melt has not been lost, the system remains isoentropic, so the entropy difference in 7.67 must simply be the entropy consumed in melting:

$$\Delta S = \Delta S_m F \tag{7.68}$$

equating the two, we have:

$$\Delta S_m F \cong \frac{C_p}{T} (T_{pot} - T)$$
7.69

Rearranging, and letting $\Delta T = (T_{pot} - T)$, we have:

$$T_{pot} - T \cong \frac{T}{C_p} \Delta S_m F$$

$$7.70$$

This difference, T_{pot} – T is the temperature deflection due to melting in the T-P path and is shown in Figure 7.24. If we differentiate 7.70 with respect to P (and still holding S constant), we have:

$$\left(\frac{\partial(T_{pot} - T)}{\partial P}\right)_{S} \cong \frac{\Delta S_{m}}{C_{p}} \left(\frac{\partial(TF)}{\partial P}\right)_{S} = \frac{\Delta S_{m}}{C_{p}} \left[F\left(\frac{\partial T}{\partial P}\right)_{S} + T\left(\frac{\partial F}{\partial P}\right)_{S}\right]$$
$$\left(\frac{\partial\Delta T}{\partial P}\right)_{S} \cong \frac{\Delta S_{m}}{C_{p}} F\left(\frac{\partial T}{\partial P}\right)_{S} + \frac{\Delta S_{m}}{C_{p}} T\left(\frac{\partial F}{\partial P}\right)_{S}$$
7.71

and finally:

Solving for $(\partial T / \partial P)_{S'}$ we have:

$$\left(\frac{\partial T}{\partial P}\right)_{S} \cong \left(\frac{\partial T_{pot}}{\partial P}\right)_{S} - \frac{T}{C_{p}}\Delta S_{m}\left(\frac{\partial F}{\partial P}\right)_{S}$$

$$7.72$$

The term $(\partial T_{pot}/\partial P)_s$ is just the adiabatic gradient, given by equation 7.58, and substituting that into 7.72 we have:

$$\left(\frac{\partial T}{\partial P}\right)_{S} \cong \frac{T\alpha V}{C_{P}} - \frac{T}{C_{p}}\Delta S_{m} \left(\frac{\partial F}{\partial P}\right)_{S}$$
7.73

Equation 7.73 describes the P-T path that a system undergoing isoentropic melting will follow as it rises.

The degree of melting will be a function of excess temperature, i.e., the difference between the solidus temperature and the actual temperature, which we shall call ΔT . In Figure 7.24, ΔT can be found by subtracting the solidus temperature from temperature path of the mantle packet, i.e.:

$$\left(\frac{\partial \Delta T}{\partial P}\right)_{S} \cong \left[\frac{T\alpha V}{C_{P}} - \frac{T}{C_{p}}\Delta S_{m}\left(\frac{\partial F}{\partial P}\right)_{S}\right] - \left(\frac{\partial T}{\partial P}\right)_{solidus}$$
7.74

There have been many attempts to determine the relationship between melting and temperature for mantle materials. Such melting curves are notoriously difficult to determine. Figure 7.25 shows an experimentally determined melting curve for a peridotite composition at 3.5 GPa. The curve has several

Geochemistry

ol+opx+cpx+liq

CHAPTER 7: TRACE ELEMENTS

breaks in slope that correspond to elimination of phases. Despite the kinks, one can extract from this kind of experiment a relationship between degree of melting and excess temperature, i.e., a value for $(\partial F/\partial T)_P$. For example, Langmuir et al. (1993) adopted a value of about 0.00285 for $(\partial F/\partial T)_P$ below 22% melting and 0.0015 for $(\partial F/\partial T)_P$ above 22% melting. We want to incorporate this information into our analysis. We do this as follows. First, we express the variation in temperature as a function of melt fraction and pressure:

$$dT = \left(\frac{\partial T}{\partial P}\right)_F dP + \left(\frac{\partial T}{\partial F}\right)_P dF \qquad 7.75$$

If we differentiate 7.75 with respect to pressure, specifying that entropy be held constant, we can derive the following relationship:

$$\left(\frac{\partial T}{\partial P}\right)_{S} = \left(\frac{\partial T}{\partial P}\right)_{F} + \left(\frac{\partial T}{\partial F}\right)_{P} \left(\frac{\partial F}{\partial P}\right)_{S}$$

This can be substituted into 7.72 to obtain:

$$\left(\frac{\partial T}{\partial P}\right)_{F} + \left(\frac{\partial T}{\partial F}\right)_{P} \left(\frac{\partial F}{\partial P}\right)_{S} \cong \frac{T\alpha V}{C_{P}} - \frac{T}{C_{p}}\Delta S_{m} \left(\frac{\partial F}{\partial P}\right)_{S}$$

$$7.76$$

1700

1650

1600

H

Rearranging:

$$\left(\frac{\partial T}{\partial P}\right)_{F} - \frac{T\alpha V}{C_{P}} \cong \left[-\frac{T}{C_{P}}\Delta S_{m} - \left(\frac{\partial T}{\partial F}\right)_{P}\right] \left(\frac{\partial F}{\partial P}\right)_{S}$$

and finally:

$$\left(\frac{\partial F}{\partial P}\right)_{S} \cong \frac{\frac{T\alpha V}{C_{P}} - \left(\frac{\partial T}{\partial P}\right)_{F}}{\frac{T}{C_{P}}\Delta S_{m} + \left(\frac{\partial T}{\partial F}\right)_{P}}$$
7.77

Equation 7.77 gives the *melt fraction as a function of pressure above the pressure where the mantle intersects the solidus.*

Let's now attempt to evaluate equation 7.77 by substituting some real values into it. The term $(\partial T/\partial P)_F$ is the slope in T-P space of lines of constant melt fraction. We can make two simplifying assumptions: (1) the lines of constant melt fraction are parallel to the solidus, and (2) the solidus can be adequately described by a Clapeyron slope, equation 3.3 (because the composition of both melt and solid can vary in the real mantle, the solidus will not be a simple univariant curved described by the Clapeyron equation), so 7.77 becomes:

$$\left(\frac{\partial F}{\partial P}\right)_{S} \approx \frac{\frac{T\alpha V}{C_{P}} - \frac{\Delta V_{m}}{\Delta S_{m}}}{\frac{T}{C_{P}} \Delta S_{m} + \left(\frac{\partial T}{\partial F}\right)_{P}}$$
7.78

The coefficient of thermal expansion, α , is about 3×10^{-5} K⁻¹, V is about 0.3175 cm³/g (=0.3175 J



ol+opx+log

ol+lia

Geochemistry

CHAPTER 7: TRACE ELEMENTS

MPa⁻¹g⁻¹), C_p is about 1.15 J K⁻¹g⁻¹. Thus the adiabatic gradient at 1673 K (1400°C) is about 14K/GPa. The term $(\partial T/\partial F)_p$ is, of course, just the inverse of $(\partial F/\partial T)_p$ and has a value of 1/0.00285 = 350.88 K. ΔV_m is about 0.0434 cm³/g (0.0434 J MPa⁻¹g⁻¹) and ΔS_m is about 0.362 J K⁻¹ g⁻¹, which corresponds to a slope of the solidus of about 120 K/GPa. From this we calculate a value for $(\partial F/\partial P)_S$ of about -0.12 GPa⁻¹, or about -1.2%/kbar (it is negative because the extent of melt *increases* as pressure *decreases*). Of course, we have greatly simplified matters here by neglecting the pressure and temperature dependencies of all



Figure 7.26. Relationship between dihedral angle, θ , and melt distribution at grain triple junctions. In case (a), the θ = 0, and the melt (shaded) is distributed along grain-grain boundaries as well as triple junctions. In case (b) the melt forms and interconnected network along grain triple junctions. In case (c) θ is greater than 60° and melt is present

thermodynamic functions. Thus this relationship is only approximate and considering the uncertainty in our assumptions and the thermodynamic parameters, this value could anywhere between -0.08/GPa and -0.2/GPa. So, for example, if a rising packet of mantle intersects the solidus at 100 km depth (\approx 3 GPa), upon reaching a depth of 30 km (\sim 1 GPa) that packet would have undergone 24% melting.

The solidus temperature of silicates can be substantially lowered by the addition of water and, at high pressures, of CO_2 . In the presence of either H₂O or CO₂, the melting curve will be different from that shown in Figure 7.23, and the relationship we deduced between melt, temperature, and pressure will also be different.

A final point to make is that once melt is extracted, the system is no longer isoentropic because the extracted melt carries away some of the entropy of the system. Thus our analysis would be strictly limited to batch melting, where the melt remains in equilibrium with the solid. A complete treatment of the thermodynamics of melting, including fractional melting can be found in Azimov et al. (1997). Morgan (2001) discusses the situation where the material undergoing melting is lithologically heterogeneous.

7.6.6.2 Mantle Permeability and Melt Distribution and Withdrawal

Whether the melting process approximates the batch (equilibrium) model or the fractional model depends on the permeability of the source region. If the source region is highly permeable, melt will flow out as it is created, approximating the fractional melting model; if it is impermeable, it will build up in place before ascending, approximating the equilibrium model. Permeability depends on the degree to which the melt is interconnected, and this in turn depends on the crystalliquid interfacial energy.

We explored the effects of interfacial energy on nucleation in Chapter 5 (section 5.5.3.3). We found that the difference in interfacial energy determined the geometry of nucleation. Here we wish to consider the case of how a liquid will distribute itself between grains of a solid undergoing melting. We assume that the solid consists of a single phase (e.g., olivine) and that the interfacial energy between these grains is σ_{ss} . Now consider the intersection between 3 such grains (Figure 7.26). When melt is present, there will also be an interfacial energy between the grains and the melt, σ_{sm} . If θ is the angle formed by a melt pocket at a grain triple junction, the balance

Geochemistry

CHAPTER 7: TRACE ELEMENTS

of forces may be described as:

$$\sigma_{ss} = 2\sigma_{sm}\cos\frac{\theta}{2}$$
 7.79

Rearranging, we have:

$$\cos\frac{\theta}{2} = \frac{\sigma_{ss}}{\sigma_{sm}}$$
 7.80

Depending on the value of θ , the melt will distribute itself in one of three possible ways, which are illustrated in Figure 7.26. The first case occurs when the solid-solid interfacial energy is twice that of the solid-melt interfacial energy; if so, then $\theta \approx 0$. In this case, solid-solid interfaces are energetically unfavorable and melt will form a thin film that coats all grain boundaries (Figure 7.26a). The second case is where $2\sigma_{sm} > \sigma_{ss} > 1.73\sigma_{sm}$, which corresponds to $0 < \theta < 60^{\circ}$ (Fig. 7.25b). In this case, the melt will form interconnected channels along grain triple junctions, as is illustrated in Figure 7.26, but is absent from grain-grain surfaces. The third case corresponds to $\sigma_{ss} < 1.73\sigma_{sm}$ and $\theta > 60$ (Fig. 7.25c). In this case, melt forms isolated pockets at junctions between 4 or more grains and is absent elsewhere. These pockets become connected only at relatively high melt fractions (several percent). Permeability will be high for either case a or b, allowing melt to readily migrate, and low for case c. The interfacial energy, and hence θ , depends on temperature, pressure, and the composition of the melt and solid phase, and hence will vary even within a single rock.

Scanning electron microscopy of experiments in which basaltic melt is allowed to come to textural equilibrium with olivine indicate show that θ is characteristically between 25° and 50°. The dihedral angle is larger, typically greater than 60°, for junctions between pyroxene grains and for H₂O and CO₂ fluids (though addition of water to a silicate rock reduces θ). Since the upper mantle consists of over

60% olivine, however, it is likely that melt forms an interconnected network such as that illustrated in Figure 7.27, resulting in high permeability. Experiments in which melt is induced to migrate, either a result of a gradient in melt fraction in the experimental charge or a result of stress, confirm that permeability of mantle material undergoing melting will be high. From our perspective, this means melt is likely to be extracted fairly quickly after it is created, and that very small melt fractions, perhaps as low as 0.1%, can segregate from the mantle. Thus the fractional melting model may more closely approximate melting in the mantle.

Laporte (1994) carried out similar experiments with quartz and hydrous silicate melts and found that the dihedral is in the range of 12 to 18° , indicting a high ratio of γ_{ss}/γ_{sm} . This in turn indicates that the permeability within regions of the crust undergoing melting will be relatively high. However, the rate at which melt segregates from its source depends on melt viscosity as well as permeability. Because the viscosity of even hydrous granitic magmas is 4 orders of magnitude greater than that of basalt, segregation of granitic melt requires a higher melt fraction than does segregation of basaltic melt. Nevertheless, Laporte argues that melt fractions as low as 10%



Figure 7.27. Three-dimensional network of formed by melt along triple junctions of olivine grains. From Riley and Kohlstedt (1990).

Geochemistry

CHAPTER 7: TRACE ELEMENTS

will segregate on time scales of 10⁵ yrs, whereas it had been previously believed that melt fractions as high as 30% would not segregate on reasonable geologic time scales.

7.6.6.3 Realistic Models of Mantle Melting

As we pointed out above, in most circumstances melting in the mantle occurs because of decompres-



Figure 7.28. (a). Melting regime under a mid-ocean ridge. Red lines show flow of mantle induced by passive spreading of overlying plates. Since melting results from decompression, no further melting occurs once motion becomes horizontal. Only those parts of the mantle directly under the ridge reach the maximum extent of melting. The melting regime along ridge can be assumed to uniform parallel to the ridge, hence the process is two-dimensional. The cartoon is, however, readily adapted to mantle-plume related volcanism by assuming radial symmetry. (b) Low degree melts generated between the "wet" and "dry" solidi could enrich higher degree melts from the normal mantle column (light shading) in highly incompatible elements. However, the volume of this region must be large (typically 10 times that of the normal mantle column) for this to be effective, requiring efficient transport and focussing of melt over scales of hundreds of kilometers. After Plank and Langmuir (1992).

sion. (A possible exception is in subduction zones; here the generation of melt is still poorly understood, but may ultimately be due to hydration of the mantle wedge. Addition of water lowers the melting temperature, so this is a form of flux melting.) Decompression melting is necessarily a dynamic process: a parcel of mantle will begin to melt at some depth and will continue to melt as it rises. The fraction of melt produced will increase with height above the initial melting depth. If, as we have argued above, melt segregates readily, melt will rise faster than the solid. As a result, once the parcel of mantle has risen above the depth where melting begins, melt from below will continually stream through it. The melt entering the parcel from below will initially not be in equilibrium with the solid within the parcel, having been produced as a smaller melt fraction at greater depth (and hence greater pressure and temperature). Thus melt passing through the parcel will react with the solid in an attempt to reach equilibrium with it. This is similar to the process we described above as zone refining.

The situation then is analogous to diagenesis in sediments, which we discussed in Chapter 5. There are some differences, however. In diagenesis, the fraction of solid relative to fluid does not change, except through expulsion of fluid. In the melting process, solid is converted to fluid by nature of the process. In the melting process, length scales are such that

Geochemistry

CHAPTER 7: TRACE ELEMENTS

diffusion does not significantly contribute to the flux and bioturbation does not exist, so advection is the only significant flux. Furthermore, our reduction of the problem to 1 dimension by assuming lateral uniformity will not be valid for the melting process. This is because the extent of melt will also decrease with distance from the some central point (a point under a volcano or under a spreading midocean ridge), and because melt will be focussed in from these peripheral regions toward the center. With these caveats, however, the diagenetic equation (equ. 5.171) may be directly applicable to the melting process.

Unfortunately, a truly thorough quantitative treatment of the melting process has not yet been undertaken. In one of the more thorough discussions to date, Langmuir et al. (1993) concluded that despite the complexity of the melting process, the batch melting equation gives a reasonably good approximation of incompatible element concentrations in the melt as a function of the *average* degree of melting. Beneath mid-ocean ridges, the average degree of melting will be less than the maximum degree of melting, because different parcels of mantle follow different paths. Only mantle directly beneath the ridge is able to rise the maximum amount, and hence melt the maximum amount. In the simple case illustrated in Figure 7.28a, the average extent of melting is one half the maximum extent. Other ratios are possible for other models of mantle flow.

There are two situations where batch melting may not be a good approximation of incompatible element concentrations. The first is where there is a large volume of mantle from which only very low degree melts are extracted. This situation might arise as a result of suppression of the solidus by H_2O or CO_2 fluid, a well-established phenomenon. If melting is such that a small fraction of melt, say 0.1% or so, is generated between the "wet solidus" (i.e., H_2O or CO_2 present) and the "dry solidus" and the temperature between the two is large, there could be a large region at great depth where very small degree melts are produced (Figure 7.28b). Mixing of these small degree melts with larger degree ones produced above the dry solidus then results in melt compositions different from those predicted by the batch melting region. For example, spinel is replaced by garnet as the aluminous phase in the mantle at about 60-90 km depth. These two minerals have very different partition coefficients for some elements (Table 7.5), hence melts produced in the garnet stability region will have different incompatible element concentrations (and in particular, different rare earth patterns) than those in the spinel stability region.

7.7 TRACE ELEMENT DISTRIBUTION DURING CRYSTALLIZATION

7.7.1 Equilibrium Crystallization

Equilibrium crystallization occurs when the total liquid and total solid remain in equilibrium throughout the crystallization. If we define X as the fraction of material crystallized, then

$$\frac{C_i^l}{C_i^0} = \frac{1}{DX + (1 - X)}$$
7.81

where C^{ℓ} is the concentration in the remaining liquid and C° is the concentration in the original liquid (we derive this equation in a manner exactly analogous to equation 7.42). The limit of trace element enrichment or depletion, occurs when X = 1 when $C_l/C_0 = 1/D$. Equilibrium crystallization requires the liquid keep contact with all crystals. Crystal interiors would have to maintain equilibrium through solid state diffusion, a slow process. Thus equilibrium crystallization is probably relevant only to a limited range of situations, such as the slow crystallization of an intrusion.

7.7.2 Fractional Crystallization

Fractional crystallization, which assumes only instantaneous equilibrium between solid and liquid, is a more generally applicable model of crystallization. In this case, trace element concentrations in the melt are governed by:

Geochemistry

CHAPTER 7: TRACE ELEMENTS

$$\frac{C_i^{\ell}}{C_i^o} = (1 - X)^{D - 1} \qquad 7.82$$

There is no limit to the enrichment or depletion of the liquid in this case. If *D* is very large, C'/C° approaches 0 as *X* approaches 1, and it approaches ∞ as *X* approaches 1 if *D* is very small. What happens when *D*=O?

For multiphase crystallization, we need to replace *D* in equations 7.81 and 7.82 with the bulk distribution coefficient as we defined it in equation 7.46, where m_{ϕ} in that equation would become the fraction of phase ϕ in the crystallizing mass.

Though fractional crystallization can, in principle, produce extreme trace element enrichment,



Figure 7.29. Variation of relative trace element concentration in a liquid undergoing crystallization.

this rarely occurs. A melt that has crystallized 90% or more (which would produce a ten-fold enrichment of a perfectly incompatible element in the melt) would have major element chemistry very different from its parent. From our knowledge of the compositional dependence of partition coefficients we could predict that incompatible elements would have partition coefficients close to 1 for such an acid melt[§]. This limits the enrichment of incompatible elements. However, highly *compatible* elements (elements with solid/liquid partition coefficients greater than 1 such as Ni) do have concentrations that approach 0 in fractionated melts (generally they disappear below detection limits). Variation of relative trace element concentration as a function of the fraction of liquid remaining is shown in Figure 7.29.

In summary, for moderate amounts of fractionation, crystallization has only a moderate effect on trace element concentrations, given that these concentrations vary by orders of magnitude.

7.7.3 IN SITU CRYSTALLIZATION

A magma chamber is likely to be substantially hotter than the rock surrounding it, which will result in a substantial thermal gradient at the margin of the chamber. Thus the margins, particularly the roof and walls, are likely to be cooler than the interior and it is here where crystallization will primarily occur. Crystals sloughed from the walls and roof would accumulate on the floor. When crystallization is restricted to marginal zones of the chamber, magma composition will evolve in a somewhat different manner than for simple fractional crystallization. Langmuir (1989) called this process *in situ crystallization*.

Imagine a magma chamber that is bounded by a zone of well-consolidated crystals at the margin. There may be some liquid within this zone, but we assume that the permeability is sufficiently low that it will never return to the magma chamber. Between this 'cumulate zone' and the free magma is a transition zone of higher permeability, which we call the 'solidification zone' (Figure 7.30). Magma is added to the solidification zone as crystallization advances into the chamber. As crystallization and compaction proceed within the solidification zone, liquid is expelled back into the central magma. The

[§] Silica-rich, or silicic, melts are sometimes referred to as "acidic" and Mg and Fe-rich ones as "basic". The reason is historical: it was once though that silica was present in melts as H_4SiO_4 . This is not the case, but the terminology persists.

Geochemistry

CHAPTER 7: TRACE ELEMENTS



flux of magma to the solidification zone we will designate dM_I and the return flux as dM_A . We let f be the fraction of interstitial liquid remaining after crystallization within the solidification zone, and (1 - f) be fraction of liquid crystallized within this zone. Some fraction f_T of the liquid remains to form the trapped liquid within the cumulate zone, and some fraction f_A returns to the magma, so that $f = f_T + f_A$. Hence:

$$d\mathbf{M}_{\mathrm{A}} = f_{\mathrm{A}} d\mathbf{M}_{\mathrm{I}}$$
 7.83

If the magma plus cumulates form a closed system, then the change in mass of liquid within the chamber is the difference between the flux into and out of the solidification zone:

$$dM_{\rm L} = dM_{\rm A} - dM_{\rm I} = dM_{\rm I}(f_{\rm A} - 1)$$
 7.84

If C_L is the concentration of some element in the magma and C_f is the concentration in the liquid returning from the solidification zone, then the change in mass of the element is:

$$d(M_LC_L) = C_L dM_L + M_L dC_L = C_f dM_A - C_L dM_I$$
 7.85
We define the parameter *E* as ratio of the concentration in the magma to that in the returning liquid:

$$E = C_f / C_L 7.86$$

E will depend on the partition coefficient for the element of interest and on the manner in which crystallization proceeds within the solidification zone. If, for example, we assume that there is complete equilibration between crystals and liquid within the zone, then from equ. 7.81:

$$E = \frac{1}{D(1-f)+f}$$
 7.87

Substituting equation 7.83, 7.85, and 7.86 into 7.84 and rearranging, we have:

$$\frac{dC_{L}}{C_{L}} = \frac{dM_{L}}{M_{L}} \left(\frac{f_{A}(E-1)}{f_{A}-1} \right)$$
 7.88

Assuming E and f_A are constants, we can integrate 7.88 to yield:

$$\frac{C_L}{C^0} = \left(\frac{M_L}{M^0}\right)^{f_A(E-1)/(f_A-1)}$$
 7.89

Figure 7.31 compares the change in concentration due to *in situ* crystallization and fractional crystallization for two values of D and several values of *f*. In general, *in situ* crystallization results in less enrichment of incompatible elements and less depletion of compatible elements for a given degree of crys-

tallization of a magma body than does fractional crystallization. The degree of enrichment depends on f, the fractional of liquid remaining when liquid is expelled from the solidification zone. For very small values of f, that is, complete crystallization within the solidification zone, the composition of the magma remains nearly constant. For f=1, that is, for no crystallization within the solidification zone, equation 7.89 reduces to equation 7.82, and the *in situ* curves in Figure 7.31 coincide with the fractional crystallization curves.

Figure 7.30. Cartoon of magma chamber undergoing *in situ* crystallization. The solidification zone is the transition region between consolidated cumulates (that nevertheless retain some trapped liquid) and the magma chamber. As crystallization proceeds, some liquid will be expelled from the solidification zone back into the magma. After Langmuir (1989).

November 11, 2005

Geochemistry

CHAPTER 7: TRACE ELEMENTS

7.7.4 Crystallization in Open System Magma Chambers

Thus far, we have treated crystallization only in closed systems, that is, where a certain volume of magma is intruded and subsequently cools and crystallizes without withdrawal or further addition of magma. This is certainly not a very realistic model of magmatism and volcanism. In the well-studied Hawaiian volcanoes for example, injections of new magma into crustal magma chambers are fairly frequent. Indeed, it appears that addition of magma to the Kilauean magma chamber is nearly continuous. Furthermore, many igneous rocks show petrographic evidence of mixing of differentiated magmas with more primitive ones. In this section, we consider the concentrations of trace elements in open magma chambers; that is, the case where new 'primary' magma mixes with a magma that has already undergone some fractional crystallization. Magma chambers where crystallization, eruption and addition of new magma are sometimes called RTF magma chambers, the RTF referring



refilling

Figure 7.32. Schematic illustration of a steady-state and periodically refilled, fractionally crystallized, and tapped magma chamber beneath a mid-ocean ridge.



Figure 7.31. Comparison of the effects of *in situ* and fractional crystallization on concentration for two values of the distribution coefficient, D, and several values of f. f_A is assumed to equal f, i.e., no trapped liquid remains in the cumulate zone. After Langmuir (1989).

to "refilled", "tapped", and "fractionated".

The extreme case of an RTF magma chamber is a *steady state* system, where magma resupply rate equals the rate of crystallization and eruption, thus maintaining a constant volume of liquid (Figure 7.32). In such a magma chamber, the concentrations of all elements eventually reach steady state after many cycles of refilling, eruption, and fractional crystallization. Steady state occurs when the rate of supply of the elements (due addition of new magma) becomes equal to the rate of loss (due to crystallization and withdrawal and eruption of magma).

To understand how steady state is achieved, consider a cyclic process where a volume *C* is lost by crystallization and a volume *T* is lost by eruption and a volume (T + C) is added to the magma chamber during each cycle. For incompatible elements, the concentration in the liquid initially increases because a greater mass of these elements is added by refilling than is lost by crystallization and eruption. As the concentration in the liquid increases, so

Geochemistry

W. M. White

CHAPTER 7: TRACE ELEMENTS

does the concentration in the solid since $C^s = D^{s/l} C^l$ ($D^{s/l}$ would be a bulk distribution coefficient if more than 1 phase is crystallizing). Eventually a point is reached where the concentration in the solid is so great, that loss of the element by crystallization and eruption equals the gain resulting from crystallization.

We can quickly derive an expression for the steadystate concentration of an element in the *equilibrium crystallization* case. In the steady-state, the losses of an element must equal gains of that element, so:

$$C^{\circ}(X+T) = TC^{ssl} + XC^{s} \qquad 7.90$$

where C° is the concentration in the primary magma being added to the chamber, C^{ssl} is the concentration in the steady-state liquid, and C^{s} is the concentration in the solid. Since $C^{s} = D^{s/l}C^{ssl}$:

$$C^{\circ}(X+T) = TC^{ssl} + XDC^{ssl}$$
 7.91

We can rearrange this to obtain the enrichment in the steady-state liquid relative to the primary magma:



Figure 7.34. La and Ni concentrations plotted against MgO concentration in a basalt undergoing closed system fractional crystallization (dashed line) and in a steady-state magma chamber where the mass of new magma equals the mass crystallized in each cycle.



Figure 7.33. Concentration of Ni and La in closed system fractional crystallization (dashed line) and open system crystallization (solid line) as a function of the fraction crystallized. In the open system case the mass injected into the chamber is equal to the mass crystallized (i.e., Y = 0).

$$\frac{C^{ssl}}{C^0} = \frac{X+T}{X+TD^{s/l}}$$
 7.92

For the fractional crystallization case, the enrichment factor is:

$$\frac{C^{ssl}}{C^0} = \frac{(X+T)(1-X)^{D-1}}{1-(1-X-T)(1-X)^{D-1}} \qquad 7.93$$

These equations are from O'Hara (1977).

Compatible element concentrations reach steady state after few cycles than do incompatible elements. This is illustrated in Figure 7.33, which shows how the concentrations of La, an incompatible element, and Ni, a compatible element, vary as function of the total amount crystallized in a steady-state system. The Ni concentration, which is 200 ppm in the primary magma, reaches a steady-state concentration of about 20 ppm after the equivalent of 5 magma chamber masses has crystallized. After the equivalent of 30 magma chamber masses crystallization, La has not quite reached its steady state concentration of around 25 ppm.

Unfortunately, it is never possible to measure

CHAPTER 7: TRACE ELEMENTS

the fraction of magma that has crystallized. In place of our parameter, X, petrologists often use the concentration of some 'index species', i.e., some species whose behavior is relatively well understood and whose concentration should vary smoothly as a function of the fraction crystallized. In basalts, MgO is commonly used as the index species, SiO_2 is a more common index in more acid magmas such as andesites and dacites. Figure 7.34 shows the La and Ni concentrations in the same steady-state system. Since MgO is a compatible element (though not a trace element, we could treat it using the same equations we have derived for trace elements) it quickly reaches a steady state concentration around 4.25%while the La concentration continues to increase. Lavas erupted from this magma chamber could thus have essentially constant MgO and Ni concentrations, but variable La concentrations. This apparent 'decoupling' of compatible and incompatible element concentrations is a feature of open magmatic systems.

7.8 SUMMARY OF TRACE ELEMENT VARIATIONS DURING MELTING AND CRYSTALLIZATION

For moderate amounts of crystallization, fractional crystallization does not have dramatic effects on *incompatible* elements concentrations. Concentrations of highly *compatible* elements are, however, dramatically affected by fractional crystallization. The RTF model does have significantly greater effects on incompatible element concentrations than simpler models, however.

Partial melting has much more dramatic effects on *incompatible* element concentrations. It is likely much of the incompatible element variations observed in magmas and magmatic rocks are related to variations in degree of melting. Depth of melting also has an effect, in that the phases with which melts equilibrate vary with depth. For example, the presence of garnet dramatically affects rare earth element (REE) abundances. The heavy rare earths are accepted into the garnet structure and have D's > 1. The lights are, however, highly rejected. The presence of garnet in the partial melting residua thus can lead to strong light rare earth enrichment of the melt.

Two simplifications are important for partial melting (both batch and fractional). When $D \ll F$, the enrichment is 1/F. Thus for small D, the enrichment is highly dependent on the degree of melting. If D is large, i.e., D > 1 and $D \gg F$, the depletion of the element in the melt is rather insensitive to F. In either case, when F approaches 0, the maximum enrichment or depletion is 1/D.

Highly *compatible* elements are, of course, depleted in a partial melt. But the degree of depletion is rather insensitive to the degree of melting for values of F likely to occur in the mantle (<.25). It must be emphasized that we have no good constraints on the absolute values of F.

This all works out nicely: compatible element are good qualitative indicators of the extent of fractional crystallization and incompatible elements are good indicators of the degree of melting.

Both geochemical and (experimental) petrological evidence indicates that alkali basalts and their kin are the result of lower degrees of melting than tholeiites. Highly undersaturated rocks such as nephelinites are probably produced by the smallest degrees of melting (1% or less). At the same time, alkali basalts are probably also the products of deeper melting.

Ratios of incompatible elements are generally less sensitive to fractional crystallization and partial melting than are absolute abundances, particularly if they are of similar incompatibility. For relatively large extents of melting, the ratio of two incompatible elements in a magma will be similar to that ratio in the magma source. For this reason, trace element geochemists are often more interested in the ratios of elements than in absolute abundances.

One approach commonly used is to plot the ratio of two incompatible elements against the abundance of the least compatible of the two. This kind of plot is sometimes referred to as a *process identification plot* because fractional crystallization and partial melting result in very different slopes on such a diagram. Figure 7.35 is a schematic version of a process identification plot. Crystallization, both fractional and equilibrium produce rather flat slopes on such a diagram, as does crystallization in an open system magma chamber. Partial melting produces a much steeper slope and the slope of produced by aggregates of fractional melts is similar to that of equilibrium partial melting. *In Situ* crystallization can produce a range of slopes depending on the value of *f*. Langmuir (1989) found a that a value for *f* of

M. White

CHAPTER 7: TRACE ELEMENTS

0.25 best matched the variation observed in the Kiglapait Intrusion in Labrador, a classic large layered intrusion. Assuming this value of f is typical, then the change in a trace element ratio due to in situ crystallization should be only moderately greater than for fractional crystallization.

References and Suggestions for Further Reading

- Albarède, F. 1995. Introduction to Geochemical Modeling. Cambridge: Cambridge Univ. Press.
- Azimov, P. D., M. M. Hirschmann and E. Stolper, 1997. An analysis of variations in isentropic melt productivity, Phil Trans R Soc Lond A, 355:255-281.
- Beattie, P. 1993. Olivine-melt and orthopyroxene-melt equilibria. Contrib. Mineral. Petrol. 115: 103-111.
- Beattie, P. 1994. Systematics and energetics of traceelement partitioning between olivine and silicate melts: implications for the nature of mineral/melt par-57-71.



Figure 7.35. Plot of the ratio of two incompatible elements (one with D=0.01, the other with D=0.1) vs. the concentration of the more incompatible element. Plot shows calculated effects of equilibrium partial melting and aggregate partial melting assuming concentrations of 1 in the source for both elements. Other lines show the effect of crystallization on the composition of a liquid produced by 10% equilibrium melting. Fractional crystallization, equilibrium crystallization, and open system crystallization (RTF magma chambers) produce less variation of the ratio than does partial melting. In Situ crystallization can mimic the effect of partial melting if the value of *f*, the fraction of titioning. Chem. Geol. 117: liquid returned to the magma, is sufficiently small.

- Beattie, P., M. Drake, J. Jones, W. Leeman, J. Longhi, et al. 1993. Terminology for trace-element partitioning. Geochim. Cosmochim. Acta. 57: 1605-1606.
- Ben Othman, D., W. M. White and J. Patchett. 1989. The geochemistry of marine sediments, island arc magma genesis, and crust-mantle recycling. Earth Planet. Sci. Lett. 94: 1-21.
- Blundy, J. D., J. C. Robinson and B. J. Wood, 1998. Heavy REE are compatible in clinopyroxene on the spinel lherzolite solidus, Earth Planet. Sci. Lett., 160:493-504.
- Blundy, J. and B. Wood, 1994. Prediction of crystal-melt partition coefficients from elastic moduli, Nature, 372:452-454.
- Blundy, J. and B. Wood, 2003. Partitioning of trace elements between crystals and melts, Earth Planet. Sci. Lett., 210: 383-397.
- Brice, J. C., 1975. some thermodynamic aspects of the growth of strained crystals, J. Cryst. Growth, 28:249-253.
- Brügmann, G. E., A. J. Naldrett, M. Asif, P. C. Lightfoot, N. S. Gorbachev, et al. 1993. Siderophile and chalcophile metals as tracers of the evolution of the Siberian Trap in the Noril'sk region, Russia. Geochim. Cosmochim. Acta. 57: 2001-2018.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

- Brügmann, G. E., N. T. Arndt, A. W. Hofmann and H. J. Tobschall. 1987. Noble metal abundances in komatiite suites from Alexo, Ontanrio, and Gorgona Island, Columbia. *Geochim. Cosmochim. Acta*. 51: 2159-2171.
- Burns, R. G. 1970. Mineralogical Applications of Crystal Field Theory. Cambridge: Cambridge Univ. Press.
- Burns, R. G. 1973. The partitioning of trace transition elements in crystal structures: a provocative review with applications to mantle geochemistry. *Geochimica et Cosmochimica Acta*. 37: 2395-2403.
- Carroll, M. R. and D. S. Draper. 1994. Noble gases as trace elements in magmatic processes. *Chem. Geol.* 117: 37-56.
- Chaussidon, M. and G. Libourel. 1993. Boron partitioning in the upper mantle: an experimental study. *Geochim. Cosmochim. Acta*. 57: 5053-5062.
- Chazot, G., M. Menzies and B. Harte. in press. Determination of partition coefficients between apatite, clinopyroxene, amphibole and melt in natural spinel lherzolites from Yemen: implications for wet melting of the lithospheric mantle. *Geochim. Cosmochim. Acta*.
- Drake, M. J. and D. F. Weill, 1975. Partition of Sr, Ba, Ca, Y, Eu²⁺, Eu³⁺, and other REE between plagioclase feldspar and magmatic liquid: an experimental study, *Geochim. Cosmochim. Acta*, 39, 689–712.
- Dunn, T. and C. Sen. 1994. Mineral/matrix partition coefficients for orthopyroxene, plagioclase, and olivine in basaltic to andesitic systems: a combined analytical and experimental study. *Geochim. Cosmochim. Acta*. 58: 717-734.
- Elderfield, H. and M. J. Greaves. 1982. The rare earth elements in seawater. Nature. 296: 214-219.
- Fryer, B. J. and J. D. Greenough. 1992. Evidence for mantle heterogeneity from platinum-group element abundances in Indian Ocean basalts. *Can. J. Earth Sci.* 29: 2329-2340.
- Gallahan, W. E. and R. L. Nielsen, 1992. The partitioning of Sc, Y and the rare earth elements between high-Ca pyroxene and natural mafic to intermediate lavas at 1 atmosphere, *Geochim. Cosmochim. Acta*, 56, 2387-2404.
- Green, T. H. 1994. Experimental studies of trace-element partitioning applicable to igneous petrogenesis: Sedona 16 years later. *Chem. Geol.* 117: 1-36.
- Hack, P. J., R. L. Nielsen and A. D. Johnston. 1994. Experimentally determined rare-earth element and Y partitioning behavior between clinopyroxene and basaltic liquids at pressures up to 20 kbar. *Chem. Geol.* 117: 89-105.
- Hanson, G. N., 1980. Rare earth elements in petrogenetic studies of igneous systems, Ann. Rev. Earth Planet. Sci., 8, 371-406.
- Hart, S. R. 1993. Equilibration during mantle melting: a fractal tree model. *Proc. Natl Acad. Sci. USA*. 90: 11914-11918.
- Hart, S. R. and C. Brooks. 1974. Clinopyroxene-matrix partitioning of K, Rb, Cs, Sr, and Ba. *Geochim. Cosmochim. Acta*. 38: 1799-1806.
- Hart, S. R. and C. J. Allègre, 1980. Trace element constraints on magma genesis, in *Physics of magmatic processes*, edited by R. B. Hargraves, p. 121-159, Princeton University Press, Princeton.
- Hart, S. R. and T. Dunn. 1993. Experimental cpx/melt partitioning of 24 trace elements. *Contrib. Mineral. Petrol.* 113: 1-8.
- Hauri, E. H., T. P. Wagner and T. L. Grove. 1994. Experimental and natural partitioning of Th, U, Pb and other trace elements between garnet, clinopyroxene and basaltic melts. *Chem. Geol.* 117: 149-166.
- Hie}stra, S. A. 1979. The role of collectgors in the formation of platinum deposits in the Bushveld Complex. *Can. Mineral.* 17: 469-482.
- Irving, A. J. 1978. A review of experimental studies of crystal/liquid trace element partitioning. *Geochim. Cosmochim. Acta.* 42: 743-770.
- Jochum, K. P., A. W. Hofmann and H. M. Seufert. 1993. Tin in mantle-derived rocks: Constraints on Earth evolution. *Geochim. Cosmochim. Acta*. 57: 3585-3595.
- Jochum, K.-P., A. W. Hofmann and H. M. Seifert. 1985. Sn and Sb in oceanic basalts and the depletion of the siderophile elements in the primitive mantle (abs). *EOS*. 66: 1113.
- Jones, J. H.1995. Experimental trace element partitioning. in *Rock Physics and Phase Relations: a Handbook* of *Physical Constants, AGU Reference Shelf 3*, ed. T. J. Ahrens. 73-104. Washington: AGU.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

- Kennedy, A. K., G. E. Lofgren and G. J. Wasserburg. 1993. An experimental study of trace element partitioning between olivine, orthopyroxene and melt in chondrules: equilibrium values and kinetic effects. *Earth Planet. Sci. Lett.* 115: 177-195.
- Kohlstedt, D. L., 1993. Structure, rheology and permeability of partially molten rocks at low melt fractions, in *Mantle flow and melt generation at mid-ocean ridges, Geophysical Monograph 71*, edited by J. P. Morgan, D. K. Blackman and J. M. Sinton, p. 103-121, AGU, Washington.
- Kohn, S. C. and P. F. Schofield. 1994. The implication of melt composition in controlling trace-element behavior: an experimental study of Mn and Zn partitioning between forsterite and silicate melts. *Chem. Geol.* 117: 73-87.
- Langmuir, C. H., 1989. Geochemical consequences of in situ crystallization, Nature, 340, 199-2059.
- Langmuir, C. H., E. M. Klein and T. Plank, 1993. Petrological systematics of mid-ocean ridge basalts: constraints on melt generation beneath oceanic ridges, in *Mantle flow and melt generation at mid-ocean ridges, Geophysical Monograph 71*, edited by J. P. Morgan, D. K. Blackman and J. M. Sinton, p. 183-280, AGU, Washington.
- Laporte, D. 1994. Wetting behavior of partial melts during crustal anatexis: the distribution of hydrous silicic melts in polycrystalline aggregates of quartz. *Contrib. Mineral. Petrol.* 116: 486-499.
- McClure, D. S. 1957. The distribution of transition metal cations in spinel. J. Phys. Chem. Solids. 3: 311-317.
- McClure, D. S., 1957. The distribution of transition metal ions in spinels, J. Chem. Solids. 3, 311-317.
- McLennan, S. M. 1989. Rare earth elements in sedimentary rocks: influence of provenance and sedimentary processes. in *Geochemistry and mineralogy of the rare earths, Reviews in Mineralogy 21*, ed. B. R. Lipin and G. A. McKay. 169-200. Washington: Mineral. Soc. Am.
- Mitchell, R. H. and R. R. Keays. 1981. Abundance and distribution of gold, palladium and iridium in some spinel and garnet lherzolites: implications for the nature and origin of precisou metal-rich intergranular components in the upper mantle. *Geochim. Cosmochim. Acta*. 45: 2425-2442.

Morgan, J. P., 2001. Thermodynamics of pressure release melting of a veined plum pudding mantle, *Geochem. Geophys. Geosyst.*, 2: paper number 2002GC000049.

Mysen, B. O., 1976. Partitioning of samarium and nickel between olivine, orthopyroxene, and liquid: preliminary data at 20 kbar and 1025°C, *Earth Planet Sci. Lett.*, 31: 1-7.

- Mysen, B. O., 1978. Experimental determination of nickel partition coefficients between liquid, paragasite, and garnet peridotite minerals and concentration limits to behavior according to Henry's Law at high pressure and temperature, *Am. J. Sci.*, 278: 217-243.
- Nagasawa, H., H. D. Schrieber and R. V. Morris. 1980. Experimental mineral/liquid partition coefficients of the rare earth elements (REE), Sc and Sr for perovskite, spinel, and melilite. *Earth Planet. Sci. Lett.* 46: 431-437.
- Nakamura, N., 1974. Determination of REE, Ba, Fe, Mg, Ba, and K in carbonaceous and ordinary chondrites, *Geochim. Cosmochim. Acta*, 38: 7577-775.
- Newsom, H. E., W. M. White, K. P. Jochum and A. W. Hofmann. 1986. Siderophile and chalcophile element abundances in oceanic basalts, Pb isotope evolution and growth of the Earth's core. *Earth Planet. Sci. Lett.* 80: 299-313.
- O'Hara, M. J., 1977. Geochemical evolution during fractional crystallization of a periodically refilled magma chamber, Nature, 266: 503-507.
- O'Hara, M. J., 1985. Importance of the 'shape' of the melting regime during partial melting of the mantle. *Nature*, 314: 58-62.
- Onuma, N., H. Higuchi, H. Wakita and H. Nagasawa, 1968. Trace element partition between two pyroxenes and the host lava, *Earth Planet. Sci. Lett.*, 5:47-51.
- Orgel, L. E. 1966. An Introduction to Transition Metal Chemistry: Ligand Field Theory. London: Methuen.
- Orgel, L. E., 1960. An Introduction to Transition-Metal Chemistry: Ligand-Field Theory, Methuen and Co., London.
- Piper, D. Z. 1994. Rare earth elements in the sedimentary cycle: a summary. Chem. Geol. 14: 285-304.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

- Plank, T. and C. H. Langmuir. 1992. Effects of the melting regime on the composition of the oceanic crust. *J. Geophys. Res.* 97: 19749-19770.
- Riley, G. N. and D. L. Kohlstedt, 1990. Kinetics of melt migration in upper mantle-type rocks, *Earth Planet. Sci. Lett.*, 105, 500-521.
- Rollinson, H. 1993. Using Geochemical Data: Evaluation, Presentation, Interpretation. Essex: Longman Scientific and Technical.
- Shaw, D.M., 1970. Trace element fractionation during anatexis, Geochim. Cosmochim. Acta, 34: 237-243.
- Stockman, H. W. 1982. Noble metals in the Ronda and Josephine peridotites, PhD dissertation. Cambridge. M.I.T., 1982.
- Taylor, S. R., and S. M. McLennan, 1985. *The Continental Crust: its Composition and Evolution*. 312 p., Blackwell Scientific Publications, Oxford.
- Van Westenen, W., J. D. Blundy and B. J. Wood, 2001. High field strength element/rare earth element fractionation during partial melting in the presence of garnet: Implications for identification of mantle heterogeneities, *Geochem. Geophys. Geosyst.*, 2: paper number 2000GC000133.
- Watson, E. B., 1976, Two-liquid partition coefficients: experimental data and geochemical implications, *Contrib. Mineral. Petrol.*, 56: 119-134.
- Wood, B. J., J. D. Blundy and J. C. Robinson, 1999. The role of clinopyroxene in generating U-series disequilibrium during mantle melting, *Geochim. Cosmochim. Acta*, 63:1613-1620.

Problems

1. For one element in the "other" category in Figure 7.3, write several paragraphs on its geochemistry. Include answers to the following: what valence state (or states) will it have in nature? What is its ionic radius in its most common valence state (preferably in octahedral coordination)? What is its electronegativity? What kinds of bonds will it most likely form? How will it behave, in particular what it its solubility, in aqueous solution? What element will it most easily substitute for in silicate rocks? Is it volatile or does it form volatile compounds? Is it siderophile or chalcophile? Will it behave as a compatible or incompatible element? What are its uses? What are the primary sources of the element for use by man?

2. Make rare earth plots for the following two samples. For the granite, plot it normalized to one of the chondritic values in Table 7.3. For the Mn nodule, make one plot normalizing it to chondrites and one plot normalizing it to average shale. Describe the features of the REE patterns.

	La	Ce	Pr	Nd	Sm	Eu	Gd	Tb	Dy	Но	Er	Tm	Yb	Lu
Granite	41	75		25	3.6	0.6	2.7		1.3				0.4	
Mn-nodule	110	858		116	24	4.9	24		24.1		14.4		13	1.92

3. Using the Blundy and Wood model, calculate the partition coefficients for the alkali metals in plagioclase at 1250° C. Assume that the site radius in 124 nm and that an ion of this radius would have a partition coefficient of 1. Assume that the value of Young's Modulus in plagioclase is 64 GPa and that ionic radii for the alkali are as follows: Li: 92 pm, Na 118 pm, K 151 pm, Rb 161 pm, and Cs 180 pm.

4. Relatively Ca-rich garnets (i.e., grossular-rich) appear to accept high field strength elements in both the X (normally occupied by Ca^{2+} , Mg^{2+} , Fe^{2+} , etc.) and Y (normally occupied by Al^{3+} , Fe^{3+} , etc.) crystallographic sites. The garnet/liquid partition coefficient should be the sum of the individual partition coefficients for each site. Assume that D_0 , r_0 , and *E* for the X site are 7, 91 pm and 1350 GPa respectively, and the corresponding values for the Y site are 1.9, 67 pm, and 920 GPa respectively. Ions in the X site will be in 8-fold coordination, while ions in the Y site will be in six-fold coordination. In 8-fold coordination, the ionic radii of Zr^{4+} , Hf^{4+} , Th^{4+} , and U^{4+} are 84 pm, 83 pm, 105 pm, and 100 pm respectively; in six fold coordination they will be 72 pm, 71 pm, 94 pm, and 89 pm respectively. Calculate the garnet/liquid partition coefficients for these 4 elements at 1250°C using the Blundy and Wood model.

Geochemistry

CHAPTER 7: TRACE ELEMENTS

5. Does Gallahan and Nielsen's (1992) equation for the cpx/liquid partition coefficient (equation 7.31) have the form you expect from thermodynamics? If so, what does the *a* parameter represent? What does the *b* parameter represent? Derive this equation from the thermodynamic relationships we developed in Chapter 3.

6. Calculate the cpx/melt partition coefficients for La and Sm using the method of Gallahan and Nielsen (1992) for the following basalt from Reunion Island:

SiO_2	46.99	
TiO ₂	2.66	
Al_2O_3	13.57	
Fe_2O_3	13.38	(total Fe as Fe ₂ O ₃)
MnO	0.18	
MgO	9.78	
CaO	9.51	
Na ₂ O	2.91	
K ₂ O	0.88	
P_2O_5	0.35	

7. Construct a table similar to Table 7.3 showing electronic configuration and CFSE (in terms of Δ_t) for both high spin and low spin states in *tetralhedral* coordination for Ti²⁺, V²⁺, Fe²⁺, Co²⁺, Ni²⁺, and Cu²⁺.

8. The Table below gives the spectroscopically measured octahedral crystal-field splitting parameter (Δ_0) of oxides for several transition metal ions.

(a) For each ion below, calculate the octahedral CFSE (crystal-field stabilization energy), in joules per mole, for the high spin state. The data are given in terms of wavenumber, which is the inverse of wavelength, λ . Recall from your physics that $\lambda = c/v$, and that e = hv. Useful constants are: $h = 6.626 \times 10^{-34}$ joule-sec, $c = 2.998 \times 10^{10}$ cm/sec, $N_A = 6.023 \times 10^{23}$ atoms/mole.

(b) Assuming $\Delta_t = 4/9\Delta_o$, use the table you constructed in problem 1 to calculate the tetrahedral CFSE for the high spin state.

(c) Calculate the OSPE (octahedral site preference energy) for each. (OSPE = Octa. CFSE – tetra. CFSE) Ion Δ_0 (cm⁻¹)

Ti ²⁺	16,100
V^{2+}	13,550
Fe ²⁺	11,200
Co^{2+}	8,080
Ni ²⁺	7,240
Cu ²⁺	12,600

9. Calculate the relative concentrations (i.e., C^{l}/C^{o}) in a partial melt at increments of F = 0.1, under the following assumptions:

(a.) Homogeneous solid phase, equilibrium (batch) melting for D = 0.01 and D = 10.

(b.) Calculate the relative concentration in the aggregate liquid for fractional melting for D = 0.01 and D = 10.

(c.) Plot C^{l}/C^{o} vs. F for a and b on the same graph, labeling each curve (use different colors or line types as well).

10. Calculate the enrichment in rare earth elements in an equilibrium partial melt of a mantle consisting of 10% cpx, 5% gar, 25% opx and 60% ol, assuming *modal* (phases enter the melt in the same proportion as they exist in the solid) melting for F = 0.02 and F = 0.10. Assume the concentrations in the

Geochemistry

CHAPTER 7: TRACE ELEMENTS

mantle are chondritic. Use the partition coefficients given in Table 7.5. Where data is missing in this table, *interpolate the values of partition coefficients*. Use only the following 8 rare earths: La, Ce, Nd, Sm, Eu, Gd, Dy, and Lu. Plot the results on a semi-log plot of chondrite-normalized abundance vs atomic number (i.e. typical REE plot). Draw a smooth curve through the REE, interpolating the other REE. (Hint: *work only with chondrite-normalized abundances, don't worry about absolute concentrations, so the C⁰ values will all be 1*).

11. Calculate the relative enrichments of La and Sm and the La/Sm ratio in an aggregate melt produced by continuous melting. Assume bulk distribution coefficients for these two elements of 0.01 and 0.05 respectively. Do the calculation for porosities (ϕ) of 0.001 and 0.01.

a.) Plot your results as a function of extent of melting, F, letting F vary from 0.001 to 0.1.

b.) Plot your results on a process identification diagram, i.e., plot La/Sm vs. La, assuming initial La and Sm concentrations of 1.

c.) Do the same calculation for and batch melting. Compare the two processes, aggregate continuous and batch on a plot of plot La/Sm vs. La.

12. Calculate the change in the La/Sm ratio of a melt undergoing *in situ* crystallization assuming bulk partition coefficients for La and Sm of 0.05 and 0.2 respectively. Assume the melt initially has a La/Sm ratio of 1. Do the calculation for values of *f* of 0.05 and 0.25 and assume that $f_A = f$.

a.) Plot your results as a function of M^1/M° .

b.) Plot your results on a process identification diagram, i.e., plot La/Sm vs. La, assuming initial La and Sm concentrations of 1.

W. M. White Chapter &: Radiogenic Isotope Geochemistry

Chapter 8: Radiogenic Isotope Geochemistry

8.1 INTRODUCTION

adiogenic isotope geochemistry had an enormous influence on geologic thinking in the twentieth century. The story begins, however, in the late nineteenth century. At that time Lord Kelvin (born William Thomson, and who profoundly influenced the development of physics and thermodynamics in the 19th century), estimated the age of the solar system to be about 100 million years, based on the assumption that the Sun's energy was derived from gravitational collapse. In 1897 he revised this estimate downward to the range of 20 to 40 million years. A year earlier, another Englishman, John Jolly, estimated the age of the Earth to be about 100 million years based on the assumption that salts in the ocean had built up through geologic time at a rate proportional their delivery by rivers. Geologists were particularly skeptical of Kelvin's revised estimate, feeling the Earth must be older than this, but had no quantitative means of supporting their arguments. They did not realize it, but the key to the ultimate solution of the dilemma, radioactivity, had been discovered about the same time (1896) by Frenchman Henri Becquerel. Only eleven years elapsed before Bertram Boltwood, an American chemist, published the first 'radiometric age'. He determined the lead concentrations in three samples of pitchblende, a uranium ore, and concluded they ranged in age from 410 to 535 million years. In the meantime, Jolly also had been busy exploring the uses of radioactivity in geology and published what we might call the first book on isotope geochemistry in 1908. When the dust settled, the evidence favoring an older Earth was deemed conclusive.

Satisfied though they might have been with this victory, geologists remained skeptical of radiometric age determinations. One exception was Arthur Holmes, who in 1913 estimated that the oldest rocks were at least 1600 million years old (Holmes was exceptional as well in his support for Alfred Wegener's hypothesis of continental drift). Many geologists were as unhappy with Holmes's age for the Earth as they had been with Kelvin's.

Until the end of World War II, the measurement of isotope ratios was the exclusive province of physicists. One name, that of Alfred Nier, stands out over this period. Nier determined the isotopic compositions of many elements and made the first measurements of geologic time based on isotope ratios rather than elemental abundances. Modern mass spectrometers, while vastly more sophisticated than those of a half century ago, have evolved from Nier's 1940 design. After World War II, mass spectrometers began to appear in geological laboratories. Many of these laboratories were established by former students and associates of Nier or Harold Urey of the University of Chicago. As this occurred, isotope geochemistry expanded well beyond geochronology, ultimately to have an impact in almost every branch of earth science.

Beyond providing precise ages of geologic events, radioactive decay is important because it provides natural tracers of geologic processes and because it provides information on the rates and pathways of geologic evolution. To understand the first point, consider a biologist who wishes to know how a nutrient, phosphorus for example, is utilized by an organism, a horse for example. The biologist can feed the horse grain doped with a small amount of radioactive phosphorus. Then by taking tissue and fluid samples and determining the amount of radioactive phosphorus present, he can trace phosphorus through various metabolic pathways. Similarly, an engineer might test a new automobile design by placing a model in a wind tunnel and using smoke as a tracer to follow the path of air around it. In principle at least, we could do a similar thing with the Earth. We might add dye to downwelling ocean water to trace deep ocean currents, or add a radioactive tracer to subducting lithosphere to trace mantle convection currents. In practice, however, even the contemplation of such experiments is a bit absurd. We would need far too much dye or radioactive tracer: the scales of distance and mass are simply too large for this kind of experiment. Even if we could overcome that obstacle, we would be long dead before any useful results came from our experiment: the rates of geologic processes are simply too slow.

Nature, however, has provided natural tracers in the form of the radiogenic isotopes, which are products of natural radioactivity, and these tracers have been moving through the Earth since its beginning. For example, subducting oceanic crust has a different ratio of ⁸⁷Sr to ⁸⁶Sr than does the mantle,

W. M. White Chapter &: Radiogenic Isotope Geochemistry

so we can use the 87 Sr/ 86 Sr ratio to trace the flow of subducting lithosphere through the mantle. Similarly, Atlantic water has a lower 143 Nd/ 144 Nd ratio than does Pacific water, so we can trace the flow of North Atlantic Deep Water into the Pacific using the 143 Nd/ 144 Nd ratio.

Geochemistry

To understand the second point, consider the continental crust, which has a much higher ratio of Rb to Sr than does the mantle. Through time, this has led to a higher ratio of ⁸⁷Sr, the product of radioactive decay of ⁸⁷Rb, to ⁸⁶Sr in the crust than the mantle. However, the ⁸⁷Sr/⁸⁶Sr ratio in crustal rocks is lower than it should be had these rocks had their present Rb/Sr ratio for 4500 million years. From this observation we can conclude that the crust has not existed, or at least has not had its present composition, for the full 4500 million year history of the Earth. The situation is just the opposite for the mantle: had it had its present Rb/Sr ratio for 4500 million years it should have a lower ⁸⁷Sr/⁸⁶Sr than it does. Apparently, the mantle has had a higher Rb/Sr ratio in the past. From these simple observations we can draw the inference that the crust has evolved from the mantle through time. With more quantitative observations we can use isotope ratios to estimate the *rate* of crustal evolution.

Two fundamental assumptions are involved in virtually all geologic uses of radiogenic isotope ratios. The first is that the rate of radioactive decay is independent of all external influences, such as temperature, pressure, etc. The second is that two isotopes of the same element are chemical identical and therefore that chemical processes cannot change, or fractionate, the ratio of two isotopes of the same elements. Neither of these assumptions holds in the absolute*. Nevertheless, all available evidence indicates violations of these assumptions are entirely negligible.

8.2 Physics of the Nucleus and the Structure of Nuclei

Nuclear physics is relevant to geochemistry for two reasons. First, the study of the distribution of isotopes forms an increasingly important part of geochemistry as well as earth science generally. Second, geochemistry concerns itself not only with the distribution of elements, but also with their origin, and the elements originated through nuclear processes.

Nuclei are made up of various numbers of neutrons and protons. We'll use N to represent the number of neutrons, the *neutron number*, and Z to represent the number of protons, or *proton number*. Z is also the *atomic number* of the element, because the chemical properties of elements depend almost exclusively on the number of protons (since in the neutral atom the number of electrons equals the number of protons). The sum of Z and N is the mass number A.

8.2.1 Nuclear Structure and Energetics

Not all possible combinations of protons and neutrons result in stable nuclei. Typically for stable nuclei, $N \approx Z$. Thus a significant portion of the nucleus consists of protons, which tend to repel each other. From the observation that nuclei exist at all, it is apparent that another force must exist which is stronger than coulomb repulsion at short distances. It must be negligible at larger distances, otherwise all matter would collapse into a single nucleus. This force, called the nuclear force, is a manifestation of one of the fundamental forces of nature (or a manifestation of the single force in nature if you prefer unifying theories), called the *strong* force. If this force is assigned a strength of 1, then the strengths of other forces are: electromagnetic 10^{-2} ; weak force 10^{-5} ; gravity 10^{-39} (we will discuss the weak nuclear force is mediated by the pion. The photon carries one quantum of electromagnetic force field; the pion carries one quantum of nuclear force field. The strong force also binds quarks together to form hadrons, a class of particles that includes neutrons and protons. The intensity of the strong force decreases rapidly with distance, so that at distances more than about 10^{-14} m it is weaker than the electromagnetic force.

^{*} There is a slight dependence of the rate of electron capture on pressure, and at extreme temperatures where nuclei become thermally excited there could be a dependence of decay rate on temperature (such temperatures, however, will only occur in the interiors of stars). There are subtle differences in the chemical behavior of the different isotopes of an element, which can be exploited for geological use, as we shall see in the next Chapter. For most radiogenic elements, 'isotopic fractionations' are small and corrections for them are easily and routinely made.

Geochemistry

W. M. White Chapter &: Radiogenic Isotope Geochemistry

One of the rules of thermodynamics was that the configuration with the lowest Gibbs Free Energy is the most stable. This is really just one example of the general physical principle that the lowest energy configuration is the most stable; this rule applies to electron orbital configurations, as we saw in crystal field theory, and to nuclei. We would thus expect that if ⁴He is stable relative to two free neutrons and two free protons, ⁴He must be a lower energy state compared to the free particles. If this is the case, then from Einstein's mass-energy equivalence:

$$E = mc^2$$
 8.1

we can predict that the ⁴He nucleus will have less mass that 2 free neutrons

and protons. It does in fact have less Figure 8.1 Binding energy per nucleon vs. mass number.

mass. From the principle that the lowest energy configurations are the most stable and the mass-energy equivaler We begin by calculating the nominal weight of an atom from the sum of the mass of the constituent particles:

proton: 1.007593 unified atomic mass unit (u) or Dalton (Da)^{*} = $1.6726231 \times 10^{-27}$ kg

neutron: 1.008982 u

electron: $0.000548756 \text{ u} = 9.10093897 \text{ x} 10^{-31} \text{ kg}$

Then we define the *mass decrement* of an atom as:

$$\delta = W - M \tag{8.2}$$

where W is the sum of the mass of the constituent particles and M is the actual mass of the atom. For example, W for ⁴He is $W = 2m_p + 2m_n + 2m_e = 4.034248$ u. The mass of ⁴He is 4.003873 u, so $\delta = 0.030375$ u. Converting this to energy using Equ. 1 yields 28.28 Mev. This energy is known as the *binding energy*. Dividing by A, the mass number, or number of nucleons, gives the *binding energy per nucleon*, E_b:

$$E_b = \frac{W - M}{A}c^2 \tag{8.3}$$

 E_b is a measure of nuclear stability: those nuclei with the largest binding energy per nucleon are the most stable. Figure 8.1 shows E_b as a function of mass. Note that the nucleons of intermediate mass tend to be the most stable.

Some indication of the relative strength of the nuclear binding force can be obtained by comparing the mass decrement associated with it to that associated with binding an electron to a proton in a hydrogen atom. The mass decrement above is of the order of 1%, 1 part in 10^2 . The mass decrement associated with binding an electron to a nucleus of the order of 1 part in 10^8 , so bonds between nucleons are about 10^6 times stronger than bonds between electrons and nuclei.

Why are some combinations of N and Z more stable than others? The answer has to do with the forces between nucleons and to how nucleons are organized within the nucleus. The structure and organization of the nucleus are questions still being actively researched in physics, and full treatment is certainly beyond the scope of this book, but we can gain some valuable insight to nuclear stability by considering two of the simplest models of nuclear structure. The simplest model of the nucleus is the *liquid-drop model*, proposed by Niels Bohr in 1936. This model assumes all nucleons in a nucleus have equivalent states. As its name suggests, the model treats the binding between nucleons as similar to the



^{*} The was formerly called an *atomic mass unit* or *amu*. The unified atomic mass unit or Dalton is defined as 1/12th the mass of a ${}^{12}C$ atom. $1 u = 1.6605402 \times 10^{-27}$ kg. Both u and Da are equivalent and acceptable SI units. The *unified atomic mass unit* seems preferred in physics and related fields while the *Dalton* seems preferred in biochemistry.

Geochemistry

CHAPTER O: RADIOGENIC ISOTOPE GEOCHEMISTRY

Example 8.1. Calculating Binding Energies

Calculate the binding energies of 50 V, 50 Cr, and 50 Ti. Which of these 3 nuclei is the least stable? Which is the most stable?

Answer: the nucleus of ⁵⁰V consists of 23 protons and 27 neutrons, that of ⁵⁰Cr consists of 24 protons and 26 neutrons, that of ⁵⁰Ti consists of 22 protons and 28 neutrons. Atoms of these elements also have, respectively 23, 24, and 22 electrons. First we calculate W for each:

 \overline{W} (⁵⁰V) = 23 × 1.007593 + 27 × 1.008982 + 23 × 0.000548756 = 50.429774 u

W $\binom{50}{2}$ Cr = 24 × 1.007593 + 26 × 1.008982 + 24 × 0.000548756 = 50.428934 u

W $({}^{\circ 0}\text{Ti}) = 22 \times 1.007593 + 28 \times 1.008982 + 22 \times 0.000548756 = 50.430615 \text{ u}$

The actual masses (M) of these nuclides are: ⁵⁰V: 49.947162 u; ⁵⁰Cr: 49.946047 u; ⁵⁰Ti: 49.944792 u. Using equation 8.2 we calculate the mass decrement, and then divide by 50 to calculate the mass decrement per nucleon. We convert the result to kg using the conversion factor 1 dalton = 1.66×10^{-27} kg. We then multiply by the square of the speed of light (2.998 × 10^8 m/sec) to obtain the binding energy in kg-m/sec or joules. We use $1 \text{ J} = 6.2415 \times 10^{12}$ MeV to convert our answer to MeV. The results are shown in the table below.

Nuclide	δ	δ	δ/ A	E _b	E _b
	u	k g	kg/nucleon	J/nucleon	MeV
50 V	0.482612	8.011449×10^{-28}	1.60229×10^{-29}	1.4401×10^{-12}	8.9885
⁵⁰ C r	0.482887	$8.016014 imes 10^{-28}$	1.60320×10^{-29}	1.4410×10^{-12}	8.9937
⁵⁰ T i	0.484568	8.043919 ×10 -28	1.60878×10^{-29}	1.4460×10^{-12}	9.0245

Our results indicate that of the three, ⁵⁰V is the least stable and ⁵⁰Ti the most stable, though the difference is not that great.



Figure 8.2. Binding energy per nucleon vs. mass number.

model, the total binding of nucleons is influenced by 3 effects: a volume energy, a surface energy, and a coulomb energy. The variation of these three forces with mass number and their total effect is shown in Figure 8.2. Looking again at Figure 8.1, we see

binding between molecules in a liquid

drop.

According to the liquid-drop

that, except for very light nuclei, the binding energy per nucleon is roughly constant, or that total binding energy is roughly proportional to the number of nucleons. Similarly, for a drop of liquid, the energy required to evaporate it, or unbind the molecules, would be proportional to the volume of the liquid. So the

volume effect contributes a constant amount of energy per nucleon.

The surface effect arises from saturation of the strong nuclear force: a nucleon in the interior of the nucleus is surrounded by other nucleons and exerts no force on more distance nucleons. But at the surface, the force is unsaturated, leading to a force similar to surface tension in liquids. This force tends to minimize the surface area of the nucleus. The surface force is strongest for light nuclei and becomes rapidly less important for heavier nuclei.

The third effect on nuclear stability considered by the liquid drop model is the repulsive force between protons. This force is a longer range one than the strong force and does not show saturation. It is proportional to the total number of proton pairs (Z(Z - 1)/2) and inversely proportional to radius. Figure 8.3 shows the stable combinations of N and Z on a plot of N against Z. Clearly, for light iso-

W. M. White Chapter &: Radiogenic Isotope Geochemistry

topes, N must roughly equal Z for a nucleus to be stable. For heavier isotopes the field of stability moves in the direction of N > Z. This effect may also be explained by the repulsive Coulomb force of the protons. The additional neutrons act to dilute the charge density (increase the radius) and thereby increase stability.

The liquid drop model can account for the general pattern of binding energy in Figure 8.1 and the general distribution of stable nuclei in Figure 8.3, but not the details. The model predicts a smooth variation in binding energy with mass number, but it is apparent from Figure 8.1 that this is not the case: certain maxima occur and some configurations are more stable than others. From this, we might guess that the nucleus has some internal structure.

Another interesting observation is the distribution of stable nuclei. Nuclei with even number of protons and neutrons are more stable than those with odd numbers of protons or neutrons. As Table 8.1 shows, stable even-even configurations are most common; stable odd-odd configurations are particularly rare. In addition, as can be seen in Figure 8.3, stable nuclei seem to be particularly common at *magic numbers*, i.e., when either N or Z equals 2, 8, 20, 28, 50, 82, and 126. These observations, the even number effect and the magic number effect, lead to the *shell model of the nucleus*. It is similar to the shell model of electron structure and is based on the same physical principles, namely the Pauli exclusion principle and quantum mechanics. The Pauli exclusion principle says that no state occupied by one nucleon can be occupied by another nucleon; i.e., a nucleon added to a nucleus must occupy a new state, or niche.

These states can be described by quantum numbers. One of these quantum numbers is spin. Two

protons can have the same spatial quantum numbers if their spins are anti-aligned (the situation is analogous to electrons sharing orbits). This is also true of neutrons. Apparently, nuclei are more stable when spins cancel (i.e., even number of protons or neutrons). The first proton and neutron shells are filled when occupied by 2 nucleons each. As in the atomic model, filling these shells produces a particularly stable configuration. The next shells are filled when 6 additional protons and neutrons are added neutron number for a total of 8 (each). This configuration is ¹⁶O. And so on, shells being filled with 2, 8, 20, 28, 50, 82, and 126 nucleons. These numbers of nucleons, which correspond to particularly stable nuclei, were called Ž 'magic numbers' and were an important clue leading to the shell model.

Another important aspect of the shell model is its prediction of nuclear angular momentum. Even-even nuclei have no angular momentum because the spins of the neutrons cancel by anti-alignment, as do the proton spins. And the angular orbital momentum is zero because the nucleons are in closed shells. In even-odd and oddeven nuclides one odd nucleon combines its half integral spin with the integral orbital angular momentum quantum number of the nucleus yielding half-integral angular momentum. In odd-odd nuclei, the odd proton and odd neutron each contribute a



Geochemistry

Figure 8.3. Neutron number vs. proton number for stable nuclides.

Geochemistry

W. M. White Chapter &: Radiogenic Isotope Geochemistry

		o. o.,		
Z	Ν	А	number of stable nuclei	number of very long-lived nuclei
		(Z + N)		
odd	odd	even	4	5
odd	even	odd	50	3
even	odd	odd	55	3
even	even	even	165	11

TABLE 8.1. NUMBERS OF S	TABLE ODD AND	EVEN NUCLEI
-------------------------	---------------	--------------------

spin of 1/2 yielding an integral angular momentum, which can combine with an integral orbital angular momentum quantum number to produce an integral angular momentum.

A slightly more complex model is called the *collective model*. It is intermediate between the liquiddrop and the shell models. It emphasizes the collective motion of nuclear matter, particularly the vibrations and rotations, both quantized in energy, in which large groups of nucleons can participate. Eveneven nuclides with Z or N close to magic numbers are particularly stable with nearly perfect spherical symmetry. Spherical nuclides cannot rotate because of a dictum of quantum mechanics that a rotation about an axis of symmetry is undetectable, and hence cannot exist, and in a sphere every axis is a symmetry axis. The excitation of such nuclei (that is, when their energy rises to some quantum level above the ground state) may be ascribed to the vibration of the nucleus as a whole. On the other hand, eveneven nuclides far from magic numbers depart substantially from spherical symmetry and the excitation energies of their excited states may be ascribed to rotation of the nucleus.

8.2.2 The Decay of Excited and Unstable Nuclei

Just as an atom can exist in any one of a number of excited states, so too can a nucleus have a set of discrete, quantized, excited nuclear states. The behavior of nuclei in transforming to more stable states is somewhat similar to atomic transformation from excited to more stable sites, but there are some important differences. First, energy level spacing is much greater; second, the time an unstable nucleus spends in an excited state can range from 10⁻¹⁴ sec to 10¹¹ years, whereas atomic life times are usually about 10⁻⁸ sec; third, excited atoms emit photons, but excited nuclei may emit other particles as well as photons. The photon emitted through the decay of unstable nuclei is called a gamma ray. Nuclear reactions must obey general physical laws, conservation of momentum, mass-energy, spin, etc. and conservation of nuclear particles. In addition to the decay of an excited nucleus to a more stable state, it is also possible for an unstable nucleus to decay to an entirely different nucleus, through the emission or absorption of a particle of non-zero rest mass.

Nuclear decay takes place at a rate that follows the law of radioactive decay. Interestingly, the decay rate is dependent only on the nature and energy state of the nuclide. It is independent of the past history of the nucleus, and essentially independent of external influences such as temperature, pressure, etc. Also, it is impossible to predict when a given nucleus will decay. We can, however, predict the probability of its decay in a given time interval. The probability of decay of a nucleus in some time interval, *dt*, is λ , where λ is called the decay constant. The probability of a decay among some number, *N*, of nuclides within *dt* is λN . Therefore, the rate of decay of N nuclides is:

$$\frac{dN}{dt} = -\lambda N$$
8.4

The minus sign simply indicates N decreases with time. Equation 8.4 is a first-order rate law that we will call the *Basic Equation of Radioactive Decay*. It is very much analogous to rate equations for chemical reactions (Chapter 5), and in this sense λ is exactly analogous to *k*, the rate constant, for chemical reactions, except that λ is independent of all other factors. 8.2.2.1 GAMMA decay

Gamma emission occurs when an excited nucleus decays to a more stable state. A gamma ray is simply a high-energy photon (i.e., electromagnetic radiation). Its frequency, v, is related to the energy difference by:

W. M. White Chapter &: Radiogenic Isotope Geochemistry

$$hv = E_u - E_l \tag{8.5}$$

Geochemistry

where E_u and E_l are the energies of the upper (excited) and lower (ground) states and *h* is Planck's constant. The nuclear reaction is written as:

 $NZ^* \rightarrow NZ + \gamma$ 8.6

where *Z* is the element symbol, *N* is the mass number, and γ denotes the gamma ray.

8.2.2.2 Alpha decay

An α -particle is simply a helium nucleus. Since the helium nucleus is particularly stable, it is not surprising that such a group of particles might exist within the parent nucleus before α -decay. Emission of an alpha particle decreases the mass of the nucleus by the mass of the alpha particle plus a mass equivalent to the energy lost during the decay, which includes the kinetic energy of the alpha particle (constant for any given decay) and the remaining nucleus (because of the conservation of momentum, the remaining nucleus recoils from the decay reaction), and any gamma ray emitted.

The escape of the α particle is a bit of a problem, because it must overcome a very substantial energy barrier, a combination of the strong force and the coulomb repulsion, to get out. For example, α particles with energies below 8 Mev are scattered from the ²³⁸U nucleus. However, the α particle emerges from the decaying ²³⁸U with an energy of only about 4 Mev. This is an example of a quantum mechanical effect called *tunneling* and can be understood as follows. Quantum mechanics holds that we can never know exactly where the α particle is (or any other particle, or you or I for that matter), we only the know the probability of its being in a particular place. This probability is determined by the square of the particle's wave function, ψ . Although the wave is strongly attenuated through the potential energy barrier, it nevertheless has a small but finite amplitude outside the nucleus. Anything that can occur ultimately will, so sooner or later the alpha particle escapes the nucleus.

The daughter may originally be in an excited state, from which it decays by γ decay. Figure 8.4 shows an energy-level diagram for such a decay.

Alpha-decay occurs in nuclei with masses above the maximum in the binding energy curve of Figure 8.1, located at ⁵⁶Fe. Quite possibly, all such nuclei are unstable relative to alpha-decay, but the half-lives of most of them are immeasurably long.



Figure 8.4. Nuclear energy-level diagram showing decay of bismuth 212 by alpha emission to the ground and excited states of thallium 208.

8.2.2.3 BETA DECAY

Beta decay is a process in which the charge of a nucleus changes, but not the number of nucleons. If we plotted Figure 8.3 with a third dimension, namely energy of the nucleus, we would see that stable nuclei are located in an energy valley. Alpha-decay moves a nucleus down the valley axis; beta decay moves a nucleus down the walls toward the valley axis. Beta-decay results in the emission of an electron or positron (a positively charged electron), depending on which side of the valley the parent lies. Consider the 3 nuclei in Figure 8.5. These are known as *isobars*, since they have the same number of nucleons(12; isotopes have the
Geochemistry

W. M. White Chapter &: Radiogenic Isotope Geochemistry

same number of protons, *isotones* have the same number of neutrons). From what we know of nuclear structure, we can predict that the ¹²C nucleus is the most stable of these three, because the spins of the neutrons and protons cancel each other. This is the case: ¹²B decays to ¹²C by the creation and emission of a β^- particle and the conversion of a neutron to a proton. ¹²N decays by emission of a β^+ and conversion of a proton to a neutron.



Figure 8.5. Proton and neutron occupation levels of boron 12, carbon 12 and nitrogen 12.

The discovery of beta decay presented physicists with a problem. Angular mo-

mentum must be conserved in the decay of nuclei. The ¹²C nucleus has integral spin, as do ¹²B and ¹²N. But the beta particle (an electron or positron) has 1/2 quantum spin units, hence β decay apparently resulted in the loss of 1/2 spin units. The solution, proposed by Enrico Fermi[†], was another, essentially massless, particle called the *neutrino*, with 1/2 spin to conserve angular momentum. (Whether the neutrino has mass remains unresolved. The most recent observations and experiments have shown that the mass of the neutrino must be extremely small, and it is most likely massless.) It is also needed to balance energy. The kinetic energies of alpha particles are discrete. Not so for betas: they show a spectrum with a characteristic maximum energy for a given decay. Since energy must be conserved, and the total energy given off in any decay must be the same, it is apparent that the neutrino must also carry away part of the energy. The exact distribution of energy between the beta and the neutrino is random: it cannot be predicted in an isolated case, though there tends to be a fixed statistically distribution of energies, with the average observed beta energies being about 1/3 the maximum value (the maximum value is the case where the beta carries all the energy).

Beta decay involves the weak nuclear force. The weak force transforms a neutral particle into a charged one and visa versa. Both the weak and the electromagnetic force are now thought to be simply a manifestation of one force that accounts for all interactions involving charge (in the same sense that electric and magnetic forces are manifestations of electromagnetism). This force is called electroweak. In β^+ decay, for example, a proton is converted to a neutron, giving up its +1 charge to a neutrino, which is converted to a positron. This process occurs through the intermediacy of the W+ particle in the same way that electromagnetic processes are mediated by photons. The photon and W particles are members of a class of particles called *bosons* that mediate forces between the basic constituents of matter. However, W particles differ from photons in having a substantial mass.

8.2.2.4 Electron Capture

Another type of reaction is electron capture. This is sort of the reverse of beta decay and has the same effect, more or less, as β^+ decay. Interestingly, this is a process in which an electron is added to a nucleus to produce a nucleus with less mass than the parent! The missing mass is carried off as energy by an escaping neutrino, and in some cases by a γ . In some cases, a nucleus can decay by either electron capture, β^- , or β^+ emission. An example is the decay of ⁴⁰K, which decays to ⁴⁰Ar by β^+ or electron capture and to ⁴⁰Ca by β^- . In Example 8.1, we found that ⁵⁰V was less stable than it's 2 isobars: ⁵⁰Cr and ⁵⁰Ti. In fact, a ⁵⁰V atom will eventually decay to either a ⁵⁰Cr atom by β^- decay or to ⁵⁰Ti by electron capture. The half-life for this decay is 1.4×10^{17} years, so that the decay of any single atom of ⁵⁰V is extremely improbable.

 β decay and electron capture often leave the daughter nucleus in an excited state. In this case, it will decay to its ground state (usually very quickly) by the emission of a γ -ray. Thus γ rays often ac-

[†]Enrico Fermi (1901-1954) had the unusual distinction of being both an outstanding theorist and an outstanding experimentalist. He made many contributions to quantum and nuclear physics and won the Nobel Prize in 1938. Interestingly, the journal *Nature* rejected the paper in which Fermi made this proposal!

company β decay. A change in charge of the nucleus necessitates a rearrangement of the electrons in their orbits. This rearrangement results in x-rays being emitted from electrons in the inner orbits.

Geochemistry

8.2.2.5 Spontaneous Fission

Fission is a process in which a nucleus splits into two or more fairly heavy daughter nuclei. In nature, this is a very rare process, occurring only in the heaviest nuclei, ²³⁸U, ²³⁵U, and ²³²Th (it is, however, most likely in ²³⁸U). This particular phenomenon is perhaps better explained by the liquid drop model than the shell model. In the liquid drop model, the surface tension tends to minimize the surface area while the repulsive coulomb energy tends to increase it. We can visualize these heavy nuclei as oscillating between various shapes. The nucleus may very rarely become so distorted by the repulsive force of 90 or so protons, that the surface tension cannot restore the shape. Surface tension is instead minimized by the splitting the nucleus entirely. Since there is a tendency for the N/Z ratio to increase with A for stable nuclei, the parent is neutron-rich. When fission occurs, some free neutrons are produced and nuclear fragments (the daughters, which may range from A=30, zinc, to A=64, terbium) are too rich in neutrons to be stable. The immediate daughters will decay by β^- decay until enough neutrons have been converted to protons that it has reached the valley of energy stability. It is this tendency to produce unstable nuclear by-products, rather than fission itself, which makes fission in bombs and nuclear reactors such radiation hazards.

Some unstable heavy nuclei and excited heavy nuclei are particularly subject to fission. An important example is ²³⁶U. Imagine a material rich in U. When one ²³⁸U undergoes fission, some of the released neutrons are captured by ²³⁵U nuclei, producing ²³⁶U in an excited state. This ²³⁶U then fissions producing more neutrons, etc. – a sustained chain reaction. This is the basis of nuclear reactors and bombs (actually, the latter more often use some other nuclei, like Pu). The concentration of U, and ²³⁵U in particular, is not high enough for this sort of thing to happen naturally – fisson chain reactions require U enriched in ²³⁵U. However, the concentration of ²³⁵U was higher in the ancient Earth and at least one sustained natural chain reaction is known to have occurred about 2 billion years ago in the Oklo uranium deposit in Gabon, Africa. This deposit was found to have an anomalously low ²³⁵U/²³⁸U ratio, indicating some of the ²³⁵U had been 'burned' in a nuclear chain reaction. Anomalously high concentrations of fission-produced nuclides confirmed that this had indeed occurred.

Individual fission reactions are less rare. When fission occurs, there is a fair amount of kinetic energy produced, the fragments literally flying apart. These fragments damage the crystal structure through which they pass, producing 'tracks', whose visibility can be enhanced by etching. This is the basis of *fission track dating*.

Natural fission also can produce variations in the isotopic abundance of the daughter elements. In general, however, the amount of the daughter produced is so small relative to that already present in the Earth, that these isotopic variations are immeasurably small. An important exception is xenon, whose isotopic composition can vary slightly due to contributions from fission of U and the extinct radionuclide ²⁴⁴Pu.

8.3 Basics of Radiogenic Isotope Geochemistry

The basic equation of radioactive decay is:

$$\frac{dN}{dt} = -\lambda N \tag{8.4}$$

 λ is the decay constant, which we defined as the probability that a given atom would decay in some time dt. It has units of time⁻¹. Let's rearrange equation 8.4 and integrate:

$$\int_{N_0}^N \frac{dN}{N} = \int_0^t -\lambda t \tag{8.7}$$

where N_0 is the number of atoms of the radioactive, or parent, isotope present at time *t*=0. Integrating, we obtain:

$$\ln\frac{N}{N_0} = -\lambda t \tag{8.8}$$

November 14, 2005

W. M. WhiteGeochemistryCHAPTER &: RADIOGENIC ISOTOPE GEOCHEMISTRY

This can be expressed as:
$$\frac{N}{N_0} = e^{-\lambda t}$$
 or $N = N_0 e^{-\lambda t}$

Suppose we want to know the amount of time for the number of parent atoms to decrease to half the original number, i.e., *t* when $N/N_0 = 1/2$. Setting N/N_0 to 1/2, we can rearrange 8.8 to get:

$$\ln \frac{1}{2} = -\lambda t_{1/2} \quad \text{or} \quad \ln 2 = \lambda t_{1/2}$$
$$t_{1/2} = \frac{\ln 2}{\lambda} \quad 8.10$$

and finally:

This is the definition of the *half-life*, t_{12} .

Now the decay of the parent produces some daughter, or *radiogenic*, nuclides. The number of daughters produced is simply the difference between the initial number of parents and the number remaining after time *t*:

$$D = N_0 - N \tag{8.11}$$

Rearranging 8.9 to isolate N_0 and substituting that into 8.11, we obtain:

$$D = Ne^{\lambda t} - N = N(e^{\lambda t} - 1)$$
8.12

This tells us that the number of daughters produced is a function of the number of parents present and time. Since in general there will be some atoms of the daughter nuclide around to begin with, i.e., when t = 0, a more general expression is:

$$D = D_0 + N(e^{\lambda t} - 1)$$
8.13

where D_0 is the number of daughters originally present.

An exponential function can be expressed as a Taylor Series expansion:

$$e^{\lambda t} = 1 + \lambda t + \frac{(\lambda t)^2}{2!} + \frac{(\lambda t)^3}{3!} + \dots$$
 8.14

Provided $\lambda t \ll 1$, the higher order terms become very small and can be ignored; hence for times that are short compared to the decay constant (i.e., for $t \ll 1/\lambda$), equation 8.13 can be written as:

$$D \cong D_0 + N \,\lambda t \tag{8.15}$$

Let's now write equation 8.13 using a concrete example, such as the decay of ⁸⁷Rb to ⁸⁷Sr:

77
Sr = 87 Sr₀ + 87 Rb($e^{\lambda t} - 1$) 8.16

As it turns out, it is generally much easier, and usually more meaningful, to measure to ratio of two isotopes than the absolute abundance of one. We therefore measure the ratio of ⁸⁷Sr to a non-radiogenic isotope, which by convention is ⁸⁶Sr. Thus the useful form of 8.16 is:

$$\frac{{}^{87}Sr}{{}^{86}Sr} = \left(\frac{{}^{87}Sr}{{}^{86}Sr}\right)_0 + \frac{{}^{87}Rb}{{}^{86}Sr}\left(e^{\lambda t} - 1\right)$$
8.17

Similar expressions can be written for other decay systems.

Equation 8.17 is a concise statement of Sr isotope geochemistry: the 87 Sr/ 86 Sr ratio in a system depends on: (1) the 87 Sr/ 86 Sr at time t = 0, (2) the 87 Rb/ 86 Sr ratio of the system (in most cases, we can assume the 87 Rb/ 86 Sr ratio is directly proportional to the Rb/Sr ratio), and (3) the time elapsed since *t* = 0. Actually, the best summary statement of isotope geochemistry was given by Paul Gast (Gast, 1960):

In a given chemical system the isotopic abundance of ⁸⁷Sr is determined by four parameters: the isotopic abundance at a given initial time, the Rb/Sr ratio of the system, the decay constant of ⁸⁷Rb, and the time elapsed since the initial time. The isotopic composition of a particular sample of strontium, whose history may or may not be known, may be the result of time spent in a number of such systems or environments. In any case the isotopic composition is the time-integrated result of the Rb/Sr ratios in all the past environments. Local differences in the Rb/Sr will, in time, result in local differences in the abundance of ⁸⁷Sr. Mixing of material during processes will tend to homogenize these local variations. Once homogenization occurs, the isotopic composition is not further affected by these processes.

This statement can, of course, be equally well applied to the other decay systems.

8.9

Table 8.2 lists the radioactive decay schemes of principal geologic interest. The usefulness and significance of each of the decay schemes are different and depend on the geochemical behavior of the parent and daughter, the half-life and the abundance of the parent. We will shortly consider the geological significance of each one.

8.3.1 Geochronology

Geochronology is not our main interest here, but it is one of the most important applications of isotope geochemistry and the two are often closely intertwined. Let's rewrite equation 8.17 in more general terms:

$$R = R_0 + R_{P/D}(e^{\lambda t} - 1)$$
 8.18

Geochemistry

where R_0 is the initial ratio and $R_{P/D}$ is the parent/daughter ratio. Measurement of geologic time is based on this equation or various derivatives of it. First let's consider the general case. Given a measurement of an isotope ratio, R, and a parent daughter ratio, $R_{P/D}$, two unknowns remain in equation 8.18: t and the initial ratio. We can calculate neither from a single pair of measurements. But if we can measure R and $R_{P/D}$ on a second system for which we believe t and R_0 are the same, we have two equations and two unknowns and subtracting the two equations yields:

$$\Delta R = \Delta R_{P/D} (e^{\lambda t} - 1)$$
8.19

which we can solve for *t*. Rearranging:

$$\frac{\Delta R}{\Delta R_{P/D}} = e^{\lambda t} - 1 \tag{8.20}$$

$$t = \frac{\ln\left(\frac{\Delta R}{\Delta R_{P/D}} + 1\right)}{\lambda}$$
8.21

t may be then be solved for as:

We can obtain the ratio $\Delta R/\Delta R_{P/D}$ from any two data points, regardless of whether they are related or not. To overcome this problem in practice, many pairs of measurements of *R* and $R_{P/D}$ are made. How do we calculate the age when many pairs of measurements are made? Note that equation 8.18 has the form y = a + bx, where *y* is *R*, *a* is the intercept, R_0 , *b* is the slope, $e^{\lambda t}$ - 1, and *x* is $R_{P/D}$. An equation of this form is, of course, a straight line on a plot of *R* vs. $R_{P/D}$, such as Figure 8.6. The slope of the line passing through the data is thus related to the age of the system and is called an *isochron*. When multiple measurements of the daughter isotope ratio and parent-daughter ratio are available, the slope, $\Delta R/\Delta R_{P/D}$, can be calculated the statistical technique of *linear regression*, which we mentioned several times previously, and which is described in Appendix III. The age is then obtained by substituting the value of the slope into equation 8.21. Regression also yields an intercept, which is simply the initial ratio R_0 since, as may

Table 8.2: Long-Lived Radioactive Decay Systems of Geochemical Interest

Parent	Decay Mode	λ	Half-life	Daughter	Ratio
40 K	β^+ , e.c, β^-	5.543 x 10 ⁻¹⁰ y ⁻¹	1.28 x 10 ⁹ yr	⁴⁰ Ar, ⁴⁰ Ca	$^{40}Ar/^{36}Ar$
⁸⁷ Rb	β-	$1.42 \times 10^{-11} y^{-1}$	4.8 x 10 ¹⁰ yr	⁸⁷ Sr	⁸⁷ Sr/ ⁸⁶ Sr
¹³⁸ La	β-	2.67 x 10 ⁻¹² y ⁻¹	2.59 x 10 ¹¹ yr	¹³⁸ Ce	¹³⁸ Ce/ ¹⁴² Ce, ¹³⁸ Ce/ ¹³⁶ Ce
¹⁴⁷ Sm	ά	6.54 x 10 ⁻¹² y ⁻¹	$1.06 \ge 10^{11} \text{yr}$	¹⁴³ Nd	¹⁴³ Nd/ ¹⁴⁴ Nd
$^{176}Lu^{\ddagger}$	β-	$1.94 \ge 10^{-11} y^{-1}$	$3.6 \times 10^{10} \text{y}$	^{176}Hf	¹⁷⁶ Hf/ ¹⁷⁷ Hf
¹⁸⁷ Re	β-	$1.64 \ge 10^{-11} y^{-1}$	4.23 x 10 ¹⁰ y	¹⁸⁷ Os	¹⁸⁷ Os/ ¹⁸⁶ Os
²³² Th	α	4.948 x 10 ⁻¹¹ y ⁻¹	$1.4 \ge 10^{10} y$	²⁰⁸ Pb, ⁴ He	²⁰⁸ Pb/ ²⁰⁴ Pb, ³ He/ ⁴ He
²³⁵ U	α	9.849 x 10 ⁻¹⁰ y ⁻¹	$7.07 \times 10^{8} y$	²⁰⁷ Pb, ⁴ He	²⁰⁷ Pb/ ²⁰⁴ Pb, ³ He/ ⁴ He
²³⁸ U	α	1.551 x 10 ⁻¹⁰ y ⁻¹	$4.47 \times 10^9 y$	²⁰⁶ Pb, ⁴ He	²⁰⁶ Pb/ ²⁰⁴ Pb, ³ He/ ⁴ He

Note: the branching ratio, i.e. ratios of decays to ⁴⁰Ar to total decays of ⁴⁰K is 0.117. The production of ⁴He from ¹⁴⁷Sm decay is insignificant compared to that produced by decay of U and Th.

 \pm Exact value of λ is in dispute. Recent determinations range from 1.87 to 1.98 x 10⁻¹¹ yr⁻¹.



Figure 8.6. A Rb-Sr isochron. Five analyses from a clast in the Bholghati meteorite fall on an isochron, whose slope is related to the age of the system. Data from Nyquist et al. (1990).

be seen from 8.18, $R = R_0$ when $R_{P/D} = 0$.

In special cases, we can calculate *t* from a single pair of measurements. One such case is where there is no significant amount of the daughter element present at *t* = 0, and we can drop the initial ratio term from equation 8.18. An example is K-Ar dating of volcanic rocks. Ar, the daughter, is lost upon eruption, and the only Ar present is radiogenic Ar. Another example is minerals that concentrate the parent and exclude the daughter. Zircon, for example, accepts U but not Pb; micas accept Rb but not Sr. Even in these cases, however, precise age determination generally requires we make some correction for the very small amounts of the daughter initially present. But since $R >> R_0$, simple assumptions about R_0 gen-

Geochemistry

erally suffice.

There are two important assumptions built into the use of equation 8.18. (1) The system of interest was at isotopic equilibrium at time t = 0. *Isotopic equilibrium in this case means the system had a homogeneous, uniform value of* R_0 . (2) The system as a whole and each analyzed part of it was closed between t = 0 and time t (usually the present time). Violation of these conditions is the principal source of error in geochronology.

Appendix III sets out the equations for conventional least-squares linear regression. In geochronology, it is important to also make estimates of the error in the age and the error on the initial ratio. These can be estimated from the error on the slope and intercept respectively. However, in practical geochronology, an approach called two-error regression (York, 1969), which takes account of measurement errors in both R and $R_{P/D}$, is generally used. The details of this method are, however, beyond the scope of this book.

8.4 DECAY SYSTEMS AND THEIR Applications

8.4.1 Rb-Sr

This decay system was one of the first to be widely used in geochronology and remains one of the most useful geochemical tracers. An important advantage of the system is relatively large variations of the Rb/Sr ratio in rocks. Because of the difference in geochemical properties of the two elements, Rb/Sr can vary by several orders of magnitude. Since the accuracy of an age determination depends heavily on the spread of measured ratios, this makes the Rb-Sr system a useful geochronological tool. As we noted in Chapter 7, Rb is a highly soluble, highly incompatible element. Sr is also relatively soluble and is fairly incompatible in mafic and, particularly, ultramafic systems. However, it is relatively compatible in silica-rich igneous systems, partitioning preferentially into plagioclase. The result is that the mantle has a relatively uniform and low ⁸⁷Sr/⁸⁶Sr ratio, and the continental crust has a much more variable, and, on average, higher ratio.

The Sr isotopic evolution of the Earth and its major silicate reservoirs (the continental crust and mantle) is illustrated in Figure 8.7, which is a plot of ⁸⁷Sr/⁸⁶Sr vs. time (in giga-annum, or billions of years). Such a plot is called an *isotope evolution diagram*. A characteristic of such diagrams is that a closed reservoir will evolve along a line whose slope is proportional to the parent-daughter ratio, in this case ⁸⁷Rb/⁸⁶Sr. That this is so is easy to show from equation 8.17. Where $t <<1/\lambda$, then equation 8.17 becomes:

$$\frac{{}^{87}Sr}{{}^{86}Sr} = \left(\frac{{}^{87}Sr}{{}^{86}Sr}\right)_0 + \frac{{}^{87}Rb}{{}^{86}Sr}\lambda t$$
8.22

W. M. White

Geochemistry

CHAPTER O: RADIOGENIC ISOTOPE GEOCHEMISTRY

Example 8.2. Calculating Isochrons and Ages

The following ⁸⁷ Rb/ ⁸⁶ Sr- ⁸⁷ Sr/ ⁸⁶ Sr data were measured by Nyquist et al.	87 Rb/ 86 Sr	⁸⁷ Sr/ ⁸⁶ Sr
(1990) on a clast within the achondritic meteorite Bholghati. These data	0.01015	0.69966
are plotted in Figure 8.6. What is the age and error on the age of this clast?	0.17510	0.71052
<i>Answer:</i> We can calculate the age using equation 8.21 above. The value of	0.02166	0.70035
$\Delta R / \Delta R_{P/D}$ is the 'slope' of the isochron, i.e., the slope of a line through the	0.02082	0.70029
data on a plot of ⁸⁷ Sr/ ⁸⁶ Sr vs. ⁸⁷ Rb/ ⁸⁶ Sr. We will use simple least squares	0.01503	0.69988
linear regression to obtain the slope and intercept; the latter will be the		-
initial ⁸⁷ Sr/ ⁸⁶ Sr ratio.		

Least squares regression is a calculation intensive procedure. Many computer programs are available for calculating regression, and regression functions are built into some hand calculators. Some spreadsheets, such as Microsoft ExcelTM, also have built in regression functions. However, to

t	4.57237	±	0.0228	Ga
		*((1/n)+n	neanx^2/SSC	(Dev)*1/(n-2))
		-(CP-mea	nx*meany*i	n)^2/SSQDev)
sigma a	0.000026	SQRT(((S	SSQy-meany	^2*n)
		1/(SSQI	Dev(n-2)))	· · · ·
		-(CP-mea	nx*meany*i	n)^2/SSODev)
sigma b	0.000320	SORT(((S	SSOy-meanv	^2*n)
a (initia	0.69892	meany-b*	meanx	/ \
b	0.066204	(CP-mear	nx*meany*n)/SSODev
SSODev	0.020105	(SSOX-x-	-bar^2*n)	
СР	0.171782			
SSOx	0.031892	SSOv	2.465085	
n	5			
meanx	0.048552	meanv	0.702139	
	0.01503		0.699876	
	0.02082		0.70029	
	0.02166		0.70035	
	0 1751		0.71052	
	0.01015		0.69966	
	87Rb/86Sr		87Sr/86Sr	

illustrate the procedure, we will use a spreadsheet to calculate regression using basic spreadsheet functions and the equations in Appendix III.

Our spreadsheet is shown at left. To make it easier to read, we have defined names for those cells containing calculated parameters; the names are show to the left of the cell. We first calculate the parameters that we will need to use several times in our equations. These are the mean of x and y, the number of values, n, the sum of squares of x and y (Σx_i^2 , Σy_i^2) and the sum of the cross products, $\Sigma_{x_iy_i}$. From these parameters, we also calculate the sum of squares of deviations for x ($\sum x_i^2 - \overline{x}^2 n$). From these parameters, it is fairly straightforward to calculate the slope, b, the intercept, *a*, and the errors on the slope and intercept. Our result is that the age of the clast is 4.57±0.02 Ga (Ga is the abbreviation for gigayears, 10⁹ years). At that time, the ⁸⁷Sr/⁸⁶Sr ratio of the clast was 0.69892±0.00003.

This equation has the form:

 87 Sr/ 86 Sr = *a* + *bt*

i.e., the equation of a straight line on a plot of ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ vs. *t* with slope $\lambda {}^{87}\text{Rb}/{}^{86}\text{Sr}$ and intercept $({}^{87}\text{Sr}/{}^{86}\text{Sr})_0$ (to the degree that the Taylor Series approximation is not exact, the line will actually be curved, but the approximation is fairly good in this case). In Figure 8.7 we have plotted geologic age rather than the time since *t* = 0; in other words, we have transformed *t* to (4.55-*t*), where 4.55 Ga is the assumed age of the Earth. Thus the intercept occurs at *t* = 4.55 rather than 0.

The initial ⁸⁷Sr/⁸⁶Sr ratio of the Earth can be estimated from the initial ratio in meteorites under the assumption that the entire solar system had a uniform ⁸⁷Sr/⁸⁶Sr at the time it formed. Once it forms, the Earth is a closed system, so it evolves along a straight line with slope proportional to the bulk Earth ⁸⁷Rb/⁸⁶Sr. This ratio has been estimated in various ways to be about 0.085 (as we noted, the ⁸⁷Rb/⁸⁶Sr is proportional to the Rb/Sr ratio; the bulk Earth Rb/Sr is about 0.029).

Now suppose that a portion of the mantle partially melts at 3.8 Ga to form a segment of continental crust. The new crust has a higher Rb/Sr ratio than the mantle because Rb is more incompatible than Sr. Thus, this crust evolves along a steeper line than the bulk Earth. The residual mantle is left with a

Geochemistry

CHAPTER O: RADIOGENIC ISOTOPE GEOCHEMISTRY

W. M. White



Figure 8.7. Sr isotopic evolution of the bulk Earth, evolution of high Rb/Sr crust created at 3.8 Ga, evolution of the resulting residual mantle and the evolution of a mantle being continuously depleted. 'BABI' stands for basaltic achondrite best initial, and is the assumed initial 87 Sr/ 86 Sr of the solar system.

tial ⁸⁷Sr/⁸⁶Sr ratios.

Because Rb is volatile, we can only guess at the Rb/Sr ratio of the Earth. The present-day chondritic 87 Sr/ 86 Sr ratio is about 0.725 whereas the 87 Sr/ 86 Sr ratio of the Earth is thought to be around 0.705.

8.4.2 Sм-Nd

¹⁴³Nd is produced by α-decay of ¹⁴⁷Sm. The Sm-Nd system is in many ways opposite of the Rb-Sr system. First, the Sm/Nd ratio of the mantle is higher than that of the crust and hence the ¹⁴³Nd/¹⁴⁴Nd ratio is higher in the mantle than in the crust. Second, the Sm/Nd ratio of the crust and siliceous igneous rocks are relatively uniform, making this system unsuitable for dating such rocks, but variable in mafic and ultramafic rocks (the opposite is true of Rb/Sr). Third, neither Sm nor Nd is particularly mobile, so ages and initial ratios are relatively insensitive to weathering and metamorphism. Fourth, Nd has a short seawater residence time and ¹⁴³Nd/¹⁴⁴Nd is not uniform in seawater.

One interesting and useful feature of this system arises from the long residence time of Sr in seawater and its easy substitution in calcium carbonate. Because of its long residence time, the ⁸⁷Sr/⁸⁶Sr r ratio of seawater is homogeneous at any given time. It does vary with time, however, as is shown in Figure 8.8. The ⁸⁷Sr/⁸⁶Sr ratio of seawater is controlled by the relative input of Sr from the continents and ridge-crest hydrothermal activity. The ratio of these will vary with mean spreading rate, erosion rates, and plate geometry. The variation of ⁸⁷Sr/⁸⁶Sr in seawater through the Phanerozoic has been determined from the analysis of carbonate and phosphate fossils, so that ages can be determined simply by determining the ⁸⁷Sr/⁸⁶Sr ratio of carbonate precipitated from seawater. This is particularly true of the Tertiary, where evolution of ⁸⁷Sr/⁸⁶Sr in seawater has been particularly precisely determined and there has been a rapid and steady rise in ⁸⁷Sr/⁸⁶Sr.

lower Rb/Sr ratio than the bulk Earth, and evolves along a shallow slope. If the melting in the mantle and creation of crust were a continuous process, the mantle would evolve along a continuously decreasing slope, i.e., a concave downward path.

The relative mobility of these elements, particularly Rb, can be a disadvantage for geochronology because the closed system assumption is violated. Even very young rocks can be contaminated during weathering, hydrothermal activity, etc. Mobility of Rb or Sr can result in imprecise ages or incorrect initial ratios. On the other hand, the large range in Rb/Sr ratios in siliceous igneous and metamorphic rocks means that Rb-Sr ages are somewhat insensitive to variations in ini-



Figure 8.8. ⁸⁷Sr/⁸⁶Sr in seawater through Phanerozoic time determined from the analysis of phosphate and carbonate fossils. Dashed line shows the composition of modern seawater, grayed areas represent uncertainties.

The Sm/Nd ratio of the bulk Earth can be assumed to be the same as in chondritic meteorites (see Chapters 10 and 11). From the bulk Earth Sm/Nd ratio, the ¹⁴³Nd/¹⁴⁴Nd ratio of the bulk Earth can be calculated. That being the case, it is often more meaningful to consider variations in ¹⁴³Nd/¹⁴⁴Nd relative to the bulk Earth, or chondritic, value. This is the purpose of ε notation. ε_{Nd} is simply the relative deviation of the ¹⁴³Nd/¹⁴⁴Nd ratio from the chondritic ratio in parts in 10⁴:

$$\varepsilon_{Nd} = \frac{\binom{^{143}Nd}{^{144}Nd} - \binom{^{143}Nd}{^{144}Nd}_{chon}}{(^{143}Nd)^{^{144}Nd}_{chon}} \times 10^4$$
8.23

the present-day chondritic value of ¹⁴³Nd/¹⁴⁴Nd is 0.512638 (when corrected for mass fractionation to a ¹⁴⁶Nd/¹⁴⁴Nd ratio of 0.7219). An additional value of the ε_{Nd} notation is that most rocks will have ε_{Nd} values in the range of –20 to +10, and it is certainly easier to work with a one or two digit number than a five or six digit one. Furthermore, we learn something about the history of a rock just by knowing whether its ε_{Nd} is positive or negative. A negative value of ε_{Nd} implies that on average over the history of the Earth the Sm/Nd ratio of that rock or its precursors has been lower than chondritic. This in turn implies the rare earth pattern of the rock or its precursors was light rare earth-enriched. We can draw the opposite inference from a positive ε_{Nd} value. It is often useful to compare the ¹⁴³Nd/¹⁴⁴Nd of some ancient rock with the bulk Earth value at the time of its formation, i.e., it's initial ¹⁴³Nd/¹⁴⁴Nd or ε_t .

can calculate ε_t simply by substituting the chondritic value appropriate to that time into equation 8.23.

Figure 8.9a illustrates the gross features of the Nd isotope evolution of the Earth. The Earth as a whole has a 147 Sm/ 144 Nd ratio identical to that of chondrites (0.1967), and the initial ¹⁴³Nd/¹⁴⁴Nd of the Earth is also the same as that of chondrites. Thus the bulk Earth evolves along the same line as chondrites. Because Nd is more incompatible than Sm and will therefore be more enriched in a partial melt than Sm, the crust, formed by partial melting of the mantle, has a low Sm/Nd ratio and evolves along a line of lower slope than the bulk Earth. Partial melting depletes the mantle more in Nd than in Sm, so the mantle is left with a high Sm/Nd ratio and evolves along a steeper slope.

By replacing ¹⁴³Nd/¹⁴⁴Nd with ε_{Nd} we do not change the fundamental features of the isotope evolution diagram, but the slopes change (Figure 8.9b). Because at any time, ε_{Nd} is the deviation from chondritic ¹⁴³Nd/¹⁴⁴Nd, the bulk Earth, as well as chondritic meteorites always have $\varepsilon_{Nd} = 0$. Thus they evolve along a line with slope 0. Crustal rocks, with low Sm/Nd, evolve along negative slopes and the mantle, with high Sm/Nd, evolves along a positive slope.



Geochemistry

Figure 8.9. (a) Nd isotope evolution in mantle and crust. Bold line shows the evolution of the bulk earth or CHUR (chondritic uniform reservoir); also shown is the evolution of crust created at 3.5 Ga, the corresponding residual mantle, and the evolution of a continuously depleted mantle. (b) Evolution of bulk earth, crust, and mantle when ¹⁴³Nd/¹⁴⁴Nd is transformed to ϵ_{Nd} .

Geochemistry

A relatively large fractionation of Sm/Nd is involved in crust formation. But after a crustal rock is formed, its Sm/Nd ratio tends not to change (indeed the ¹⁴⁷Sm/¹⁴⁴Nd ratio of most crustal rocks is uniform at around 0.13). This leads to another useful concept, the *model age* or *crustal residence time*. Given

Example 8.3. Calculating Nd Model Ages

DePaolo (1981) reported that the Pike's Pike granitic batholith has a 147 Sm/ 144 Nd ratio of 0.1061 and a 143 Nd/ 144 Nd of .51197. What are it's CHUR model and depleted mantle model ages? *A n s w e r* : The closed system isotopic evolution of any sample can be expressed as:

$$({}^{143}Nd / {}^{144}Nd)_{sam} = ({}^{143}Nd / {}^{144}Nd)_0 + ({}^{147}Sm / {}^{144}Nd)_{sam} (e^{\lambda t} - 1)$$
 8.24

The chondritic evolution line is:

$$({}^{143}Nd / {}^{144}Nd)_{chon} = ({}^{143}Nd / {}^{144}Nd)_0 + ({}^{147}Sm / {}^{144}Nd)_{chon}(e^{\lambda t} - 1) \qquad 8.25$$

The CHUR model age of a system is the time elapsed, $t=\tau$, since it had a chondritic ¹⁴³Nd/¹⁴⁴Nd ratio, assuming the system has remained closed. We can find τ by subtracting equation 8.25 from 8.24:

$${}^{143}Nd / {}^{144}Nd_{sam} - {}^{143}Nd / {}^{144}Nd_{chon} = \{{}^{147}Sm / {}^{144}Nd_{sam} - {}^{147}Sm / {}^{144}Nd_{chon}\}(e^{\lambda t} - 1) \qquad 8.26$$

Another way of thinking about this problem is to imagine a $^{143}Nd/^{144}Nd$ vs. time plot: on that plot, we want the intersection of the sample's evolution curve with the chondritic one. In terms of the above equations, this intersection occurs at $(^{143}Nd/^{144}Nd)_0$.

Solving equ. 8.26 for
$$\tau$$
. $\tau_{CHUR} = \frac{1}{\lambda} \ln \left(\frac{{}^{143}Nd / {}^{144}Nd_{sam} - {}^{143}Nd / {}^{144}Nd_{chon}}{{}^{147}Sm / {}^{144}Nd_{sam} {}^{147}Sm / {}^{144}Nd_{chon}} + 1 \right)$ 8.27

The chondritic ¹⁴³Nd/¹⁴⁴Nd and ¹⁴⁷Sm/¹⁴⁴Nd ratios are 0.512638 and 0.1967 respectively. Substituting these, the data for the Pike's Pike batholith given above, and λ_{147} from Table 8.2, we have:

$$\tau_{CHUR} = \frac{1}{6.54 \times 10^{-12}} \ln \left(\frac{0.51197 - 0.512638}{0.1061 - 0.1967} + 1 \right) = 1.12Ga$$

To calculate the depleted mantle model age, τ_{DM} , we use the same approach, but this time we want the intersection of the sample evolution line and the depleted mantle evolution line. So equation 8.27 becomes:

$$\tau_{DM} = \frac{1}{\lambda} \ln \left(\frac{{}^{143}Nd / {}^{144}Nd_{sam} - {}^{143}Nd / {}^{144}Nd_{DM}}{{}^{147}Sm / {}^{144}Nd_{sam} {}^{147}Sm / {}^{144}Nd_{DM}} + 1 \right)$$
 8.28

The depleted mantle (as sampled by MORB) has an average ε_{Nd} of about 10, or $^{143}Nd/^{144}Nd = 0.51315$. The simplest possible evolution path, and the one we shall use, would be a closed system evolution since the formation of the Earth, 4.55 Ga ago (i.e., a straight line on a $^{143}Nd/^{144}Nd$ vs. time plot). This evolution implies a $^{147}Sm/^{144}Nd$ of 0.2137. Substituting these values into 8.34:

$$\tau_{DM} = \frac{1}{6.54 \times 10^{-12}} \ln \left(\frac{0.51197 - 0.51315}{0.1061 - 0.2137} + 1 \right) = 1.67 Ga$$

The age of the Pik e's Peak batholith has been dated through conventional geochronology as 1.02 Ga (the 'crystallization age'), only slightly younger than the τ_{CHUR} . If we assume the mantle is chondritic in its ¹⁴³Nd/¹⁴⁴Nd and ¹⁴⁷Sm/¹⁴⁴Nd ratios, then we would conclude that Pike's Peak batholith could have formed directly from mantle-derived material, perhaps by f ractional crystallization of basalt. In that case, it represents a new addition of material from the mantle to continental crust. However, DePaolo (1981) pointed out that the mantle has not maintained chondritic ¹⁴³Nd/¹⁴⁴Nd and ¹⁴⁷Sm/¹⁴⁴Nd ratios, its evolution is more closely approximated by the depleted mantle evolution model described above. Since the τ_{DM} is much older than the crystallization age, he concluded that Pike's Peak batholith originated by melting of pre-existing crustal material; material which had already resided in the crust for some 0.67 Ga before the batholith formed.

a measured ¹⁴³Nd/¹⁴⁴Nd and ¹⁴⁷Sm/¹⁴⁴Nd, and assuming the ¹⁴⁷Sm/¹⁴⁴Nd ratio was constant from the time of formation, we can estimate the 'age' or crustal residence time, i.e., the time the rock has spent in the crust. This assumes we know how the mantle ¹⁴³Nd/¹⁴⁴Nd has been evolving with time. A first order assumption is that it evolved like chondrites (i.e., constant ¹⁴⁷Sm/¹⁴⁴Nd = 0.1967). Ratios calculated under this assumption are termed chondritic or *CHUR model age* (τ_{CHUR}). More often, a slightly more complex evolution is assumed, in which case the term *depleted mantle model age* is used (τ_{DM}). In either case, the model age is calculated by extrapolating the ¹⁴³Nd/¹⁴⁴Nd back to the intersection with the mantle growth curve, as illustrated in Figure 8.10. Example 8.3 illustrates just how such model ages are calculated.

An interesting feature of the behavior of the rare earths is that there is little change in the relative abundances of these elements even in the production of fine-grained sediments from crystalline rocks. This means that meaningful crustal residence ages can be calculated even from sedimentary rocks. The crustal residence age tells how long the sample of Nd has spent in the crust. In other words, a crustal residence time, or age, of 2 Ga measured on particular sediment means the crystalline precursor was derived from the mantle 2.0 by ago. This has led to some interesting inferences about rates of crust formation, which we will discuss later.

Comparatively large variations in Sm/Nd and relatively high Sm/Nd ratios in mafic rocks, and in minerals composing them, make the Sm-Nd decay system a good tool for dating such rocks. The insensitivity of these elements to weathering and metamorphism is an important bonus. On the other hand, the low Sm/Nd ratios of granitic rocks make them unsuitable for Sm/Nd dating. Nevertheless, initial ¹⁴³Nd/¹⁴⁴Nd ratios of older rocks have been very useful in tracing the evolution of the crust and mantle. It is now clear that even before 3.5 Ga, the mantle had ¹⁴³Nd/¹⁴⁴Nd ratios greater than chondritic (i.e., positive ε_{Nd}). This implies a Sm/Nd greater than chondritic and depletion of the mantle in incompatible elements very early in Earth's history. On the other hand, ¹⁴³Nd/¹⁴⁴Nd ratios in the mantle have not increased as rapidly as expected, implying the Sm/Nd ratio of the upper mantle, and perhaps incompatible element concentrations generally, is somehow buffered. We will discuss these issues more fully in a subsequent chapter.

Compared to other parent-daughter ratios, the range in Sm/Nd ratios observed in the Earth is small. That, combined with the long half-life of ¹⁴⁷Sm, means that the ¹⁴³Nd/¹⁴⁴Nd ratio must be measured to very high precision, with errors no more than 5 parts in 10⁵. This level of precision is, however, readily achieved with modern mass spectrometers.

Because of the small range of Sm/Nd ratios, even in mafic rocks, problems can arise in using the Sm-Nd system for geochronology. This small Sm/Nd range results in a small range in ¹⁴³Nd/¹⁴⁴Nd ratios



Figure 8.10. Sm-Nd model ages. The ¹⁴³Nd/¹⁴⁴Nd is extrapolated backward (slope depending on Sm/Nd) until it intersects a mantle or chondritic growth curve. In this example, the *CHUR model age* is 3.05 Ga while the *depleted mantle model age* is 3.3 Ga.

and a calculated age can be influenced by small nonuniformities in the initial 143 Nd/ 144 Nd ratio. A classic example is the debate that arose over the age of volcanic rocks in Kambalda, Australia. An initial date of 2.7 Ga was obtained by the Sm-Nd technique, but a subsequent study obtained a 3.2 Ga Sm-Nd age. A third study concluded the age was actually 2.7 Ga, and that the 3.2 Ga age resulted from a correlated variation in Sm/Nd and $^{143}\mathrm{Nd}/^{144}\mathrm{Nd}$ in the rocks at the time of formation (due to assimilation of crust). Zircon and Pb-Pb ages published at about the same time supported this latter interpretation.

Geochemistry

Combined use of Sr and Nd isotope ratios can be a particularly powerful geochemical tool. Figure 8.11 shows the Sr and Nd isotopic characteristics of the ma-

jor terrestrial reservoirs. Sr and Nd isotope ratios tend to be anti-correlated in the mantle and in mantle-derived magmas. This part of the diagram is sometimes called the "mantle array". The inverse correlation reflects the relative compatibilities of the parent-daughter pairs. Rb is less compatible than Sr, while Sm is more compatible than Nd. Thus magmatic processes affect the Rb/Sr and Sm/Nd ratios in opposite ways. In time, this leads to an inverse relationship between ⁸⁷Sr/⁸⁶Sr and ε_{Nd} . The anticorrelation between Sr and Nd isotope ratios in the mantle reflects the dominance of magmatic processes in the chemical evolution of the mantle. Most of the mantle has ε_{Nd} higher and ⁸⁷Sr/⁸⁶Sr lower than the bulk Earth. This in turn implies that Sm/Nd ratios have been high and Rb/Sr ratios have been low in the mantle. Mid-ocean ridge basalts (MORB) tend to have lowest ⁸⁷Sr/⁸⁶Sr and highest ε_{Nd} . This implies the part of the mantle from which MORB is derived must be highly depleted in incompatible elements, presumably through past episodes of partial melting.

Geochemistry

The continental crust has, on average, higher ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ and lower ϵ_{Nd} than the bulk Earth. The anticorrelation between ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ and ϵ_{Nd} is much weaker in crustal rocks. In part, this reflects the important role of non-magmatic processes have played in evolution of the crust. It also reflects a weaker inverse relationship in Rb/Sr and Sm/Nd in siliceous igneous rocks. The lower continental crust appears to have somewhat lower ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ than does the upper continental crust, but it should be emphasized that there is no clear division in isotopic composition between the two, and that lower crustal rocks may plot anywhere within the crustal field. Continental basalts range to much higher ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ and



Figure 8.11. Sr and Nd isotope ratios in major geochemical reservoirs. The isotopic composition of the bulk Earth lies at the intersection of the horizontal line at $\epsilon_{Nd} = 0$ and the dashed vertical line at ${}^{87}\text{Sr}/{}^{86}\text{Sr} = 0.705$ (the latter is dashed because the bulk Earth ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ is somewhat uncertain). Arrows labeled 'enriched' and 'depleted' show where incompatible element enriched and depleted reservoirs would plot.

W. M. White

CHAPTER O: RADIOGENIC ISOTOPE GEOCHEMISTRY

Example 8.4. Modeling the SR and Nd Isotope Evolution of the Crust and Mantle

A simple first order model for the evolution of the crust and mantle might be that the crust was created early in Earth's history by partial melting of the mantle (the mass of the crust is 1.3% that of the mantle). Let's briefly explore the isotopic implications of this model. Assume that the crust was created as a 1.3% equilibrium partial melt of the mantle 3.5 Ga ago. Using partition coefficients of 0.0014, 0.023, 0.194, and 0.0813 for Rb, Sr, Sm, and Nd respectively, calculate the present ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ and ${}^{143}\text{Nd}/{}^{144}\text{Nd}$ of the crust and mantle. Assume that the bulk silicate Earth (i.e., crust plus mantle) ${}^{87}\text{Rb}/{}^{86}\text{Sr}$ and ${}^{147}\text{Sm}/{}^{144}\text{Nd}$ ratios are 0.085 and 0.1967 respectively, and that the bulk silicate Earth ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ and ${}^{143}\text{Nd}/{}^{144}\text{Nd}$ ratios are at present 0.705 and 0.512638 respectively. Also assume that the silicate Earth evolved as a single, uniform, closed system up to 3.5 Ga.

Answer: The first step is to calculate the Rb/Sr and Sm/Nd fractionations during partial melting. Using equation 7.42, we find that the Rb, Sr, Sm, and Nd concentrations in the melt (the crust) are 67.2, 30.2, 4.98, and 10.72 times those in the original source respectively. Since $D = C^{s}/C^{\ell}$, we calculate the (relative) concentrations in the residual solid (the mantle) by multiplying the melt concentrations by the partition coefficients. From these, we calculate the Rb/Sr and Sm/Nd ratios in the melt to be respectively 2.22 and 0.0465 times those in the original source; these ratios in the residual solid are respectively 0.132 and 1.11 times those in the original source. Multiplying these values times the bulk silicate Earth Rb/Sr and Sm/Nd, we obtain ⁸⁷Rb/⁸⁶Sr and ¹⁴⁷Sm/¹⁴⁴Nd ratios in the crust of 0.1892 and 0.0914 respectively; those in the mantle are 0.0113 and 0.2184 respectively.

The next step is to calculate the initial ratios at 3.5 Ga. We can calculate the initial ratios at 3.5 Ga by noting that the bulk silicate Earth is still a closed system and hence by solving equation 8.18 for R_0 at t = 3.5 Ga and substituting the values given above for bulk silicate earth parent/daughter ratios and the decay constants in Table 8.2. Doing this, we find the initial ratios (at 3.5 Ga) to be 0.7007 for ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ and 0.5081 for ${}^{143}\text{Nd}/{}^{144}\text{Nd}$. Finally, we calculate present ratios for crust and mantle using equation 8.18 using these initial ratios and the $R_{P/D}$ ratios we calculated above. In this way, we find present ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ ratios of 0.7103 and 0.7012 for the crust and mantle respectively, and present ${}^{143}\text{Nd}/{}^{144}\text{Nd}$ ratios of 0.5102 and 0.5131 for crust and mantle respectively.

lower ε_{Nd} than do oceanic basalts. For the most part, this is due to assimilation of continental crust by these magmas. In some cases, however, it may reflect derivation of these magmas from incompatible element enriched regions of the subcontinental mantle.

Because ¹⁴³Nd/¹⁴⁴Nd ratios are not uniform in seawater, ¹⁴³Nd/¹⁴⁴Nd ratios in ancient sediments cannot be used for dating the way ⁸⁷Sr/⁸⁶Sr ratios are. However, ¹⁴³Nd/¹⁴⁴Nd ratios have been used to identify different water masses and to attempt to constrain ocean circulation.

8.4.3 LU-Hf

The Lu-Hf system is in many respects similar to the Sm-Nd system: (1) in both cases both elements are relatively immobile; (2) in both cases both elements are refractory so we can assume the bulk silicate earth has a Lu/Hf ratio close to chondritic; (3) in both cases the daughter is preferentially enriched in the crust so both $^{143}Nd/^{144}Nd$ and $^{176}Hf/^{177}Hf$ ratios are lower in the crust than in the mantle.

Because of (2), we can usefully define an ε_{Hf} , which is the relative deviation in parts in 10⁴ from the chondritic value. ε_{Hf} is calculated as

$$\varepsilon_{Hf} = \frac{{}^{176}Hf / {}^{177}Hf - {}^{176}Hf / {}^{177}Hf_{chon}}{{}^{176}Hf / {}^{177}Hf_{chon}} \times 10^4$$
8.29

The present-day chondritic value is about 0.28280±0.00003: the exact value is still disputed.

The correlation between Nd and Hf isotope ratios in mantle-derived rocks is somewhat better than between Sr and Nd isotope ratios, as one might expect from the chemical behavior of the elements in-

volved: Lu is after all a rare earth like Sm and Nd, and the behavior of Hf is somewhat similar to that of the rare earths. The fractionation of Lu from Hf tends to be somewhat greater than between Sm and Nd and as a result there is greater variation in ¹⁷⁶Hf/¹⁷⁷Hf than in ¹⁴³Nd/¹⁴⁴Nd. Figure 8.11 shows how Hf isotope ratios relate to Sr and Nd isotope ratios in oceanic basalts. Hf and Nd isotope ratios scatter about a trend of $\epsilon_{Hf} = 1.4 \times \epsilon_{Nd}$, as is illustrated in Figure 8.12.

Geochemistry

Because they vary more, Hf isotope ratios should be a better geochemical tool than Nd isotope ratios, at least in principle. However, Hf is a difficult element to analyze in conventional thermal ionization mass spectrometers because it does not ionize readily. Until the development of multi-collector inductively coupled plasma mass spectrometers (MC-ICP-MS), which utilize an Ar plasma for ionization, the difficulty of analysis of Hf isotope ratios meant that few such measurements were made. In the last decade, however, increasing use of the Lu-Hf system has been made in both isotope geochemistry and geochronology. However, some disadvantages relative to the Sm-Nd system remain. Among these are difficulties in establishing the Lu/Hf ratio of the Earth, uncertainty about the present day chondritic value as mentioned above, and even uncertainty as to the exact value of the decay constant. These are all topics of active current research.

In the sedimentary reservoir, the similarity of the Lu-Hf and Sm-Nd systems breaks down. This is because Hf is geochemically very similar to Zr and as a result is concentrated in zircon in crustal rocks. Zircon is very stable mechanically and chemically and concentrates in coarse-grained sediments. Rare earths, including Lu tend to be more concentrated in fine-grained clays. As a result, Hf and Nd isotope ratios are not correlated in sedimentary rocks (e.g., Patchett et al., 1984; White et al., 1986).



Figure 8.12. ϵ_{Hf} vs. 87 Sr / 86 Sr and ϵ_{Nd} in MORB and oceanic island basalts.

Geochemistry

CHAPTER O: RADIOGENIC ISOTOPE GEOCHEMISTRY

W. M. White

8.4.4 Re-Os

¹⁸⁷Re decays to ¹⁸⁷Os by β^- decay with a half-life of 42 billion years. Unlike the other decay systems of geological interest, Re and Os are both siderophile elements. Thus, they are depleted in the silicate Earth by about 2 orders of magnitude compared to chondritic meteorites. The missing Re and Os are presumably in the core. The resulting very low concentration levels (sub-ppb) make analysis extremely difficult. Although the first Re-Os isotope studies were published in the early 1960's, isotopic analysis of Os proved so difficult that there were no follow-up studies until 1980, when an analysis technique using the ion probe was developed (Luck et al., 1980). However, it was not until the development of a new analytical technique (Creaser et al., 1991), which employed a conventional thermal ionization mass spectrometer to analyze OsO, that interest in this system really blossomed.

Several aspects of this decay system make it particularly useful in addressing a variety of geological processes. The first is the siderophile/chalcophile nature of these elements, making this a useful system to address questions of core formation and ore genesis. Second, whereas all the other radioactive and radiogenic elements are incompatible ones, and hence enriched in melts, Os is a highly compatible element (bulk D ~ 10) and is enriched in the residual solid. This makes Os isotope ratios particularly useful in studies of the mantle. Third, while Os is highly compatible, Re is moderately incompatible and is slightly enriched in the melt. For example, mantle peridotites have average Re/Os close to the chondritic value of 0.08 whereas the average Re/Os in basalts is 8.9. Thus partial melting appears to produce an increase in the Re/Os ratio by a factor of $\sim 10^2$. As a consequence, the range of Os isotope ratios in the Earth is enormous compared to other radiogenic elements (thus analytical precision need not be as high as for elements such as Sr, Nd, and Hf). The mantle



Figure 8.13. a.) Schematic evolution of Os isotope ratios in the mantle and crust. b.) ¹⁸⁷Os/¹⁸⁸Os evolution in the mantle as deduced from analyses of osmiridum and laurite grains in ancient mantle peridotites, initial ratios in peridotites, and modern oceanic peridotites and basalts. The pattern is consistent with other isotopic systems: The asthenospheric mantle is more incompatible element-depleted than mantle plumes, which produced oceanic island basalts. The mantle root of continents (lithospheric mantle), appears to have been severely depleted of its basaltic component and Re by melt extraction.

has a ¹⁸⁷Os/¹⁸⁸Os ratio close to the chondritic value of 0.1275 (Allègre and Luck, 1980), whereas the crust appears to have a Re/Os isotope ratio of 1.2 to 1.3 (Esser and Turekian, 1993). By contrast, the difference in ¹⁴³Nd/¹⁴⁴Nd ratios between crust and mantle is only about 0.5%.

Geochemistry

In the early work on this system, Hirt et al. (1963) reported the isotope ratio as ¹⁸⁷Os/¹⁸⁶Os (normal-



Figure 8.14. Histogram comparing γ_{Os} in peridotites of the subcontinental lithosphere, the upper oceanic mantle and oceanic island basalts (OIB). Samples from subcontinental lithosphere typically have γ_{Os} of -10, those of the oceanic upper mantle are typically - 2. OIB average γ_{Os} of +10. Those OIB with the most extreme γ_{Os} may have assimilated oceanic crust during eruption.

ized for fractionation to ¹⁹²Os/¹⁸⁸Os of 3.08271). However, ¹⁸⁶Os is itself radiogenic, being the product of α-decay of ¹⁹⁰Pt. ¹⁹⁰Pt is sufficiently rare, and its half-life sufficiently long, that in most cases the amount of radiogenic ¹⁸⁶Os is in-Significant radiogenic ¹⁸⁶Os was, significant. however, observed by Walker et al. (1991) in copper ores from Sudbury, Ontario, which prompted them to adopt the convention of using ¹⁸⁸Os as the normalizing isotope. Eventually, this convention was adopted by other laboratories as well and today Os isotope ratios are reported as ¹⁸⁷Os/¹⁸⁸Os, but there is a considerable amount of data in the literature using the old convention. ¹⁸⁷Os/¹⁸⁶Os ratios may be converted to 187 Os/ 188 Os ratios by multiplying by 0.12035.

Perhaps the first significant result from Re-Os studieswas the realization the Re/Os ratio of the mantle is nearly chondritic (Allègre and Luck, 1980). This is a bit surprising if most of Re and Os have been extracted to the core. If the core and mantle are in equilibrium, then mantle concentrations will be determined by metal-silicate partition coefficients, which are large for both elements. Since the partition coefficients of these two elements are different, the ratio of the two in the mantle should be quite different from chondritic. That the Re/Os ratio of the mantle is chondritic thus has important implications for how and when the core formed, which we will discuss in Chapter 11. Figure 8.13 illustrates the evolution of Os isotope ratios in the crust and mantle.

Since the silicate Earth appears to have a nearchondritic ¹⁸⁷Os/¹⁸⁸Os ratio, it is useful to define a parameter analogous to ε_{Nd} and ε_{Hf} that measures the deviation from chondritic. Walker et al. (1989) defined γ_{Os} as:

$$\gamma_{Os} = \frac{\frac{187Os}{188Os} - \frac{187Os}{188Os}}{\frac{187Os}{188Os} - \frac{187Os}{188Os}} \times 100 \qquad 8.30$$

where the present day chondritic value is 0.12753. Thus γ_{Os} is the *percent* deviation from the chondritic value. Just as with ε_{Nd} , we can calculate values of γ_{Os} for times other than the present by using initial ¹⁸⁷Os/¹⁸⁸Os and an chondritic value appropriate to that time.

It is important to remember that, unlike the Sm-Nd and Lu-Hf systems, there is no reason to believe that the Re/Os and ¹⁸⁷Os/¹⁸⁸Os ratios of the silicate Earth are *exactly* chondritic. Observa-

tional evidence suggests the 187 Os/ 188 Os is within a few percent of the chondritic value, but more than that cannot be said.

Geochemistry

Because of the extremely low concentrations of Os in most rocks, most early studies of Os isotope ratios focused on analysis of Os-rich grains, such as osmiridium alloy and laurite (RuS₂) in peridotites. Improvement in analytical techniques eventually allowed the analysis of peridotites and other ultramafic rocks such as komatiites, and Os-rich sediments such as black shales and Mn nodules, where Os concentrations exceed a part per billion. Although it remains analytically challenging, the negative ion method of Creaser et al. (1991) allows the analysis of basalts and clays with concentrations in the parts per trillion range. Consistent with Sr and Nd isotope systematics, these studies have shown that MORB and related peridotites have lower ¹⁸⁷Os/¹⁸⁸Os ratios than basalts from oceanic islands (e.g., Martin, 1991; Hauri and Hart, 1993; Reisberg et al., 1993). These results are consistent with the Nd and Sr isotope evidence we have already discussed that the source of MORB is more incompatible-element depleted than mantle plumes, which are thought to produce oceanic island volcanoes such as Hawaii. However, MORB typically have higher ¹⁸⁷Os/¹⁸⁸Os ratios than peridotites dredged from mid-ocean ridges (so-called abyssal peridotites). This is somewhat surprising since the peridotites are thought to be residual from the melting that produces MORB. If isotopic equilibrium is achieved, both should have the same isotope ratio. The difference may mean that isotopic equilibrium is not achieved during partial melting, at least for Os. Alternatively, the difference between MORB and abyssal peridotites could just reflect pervasive contamination of MORB by seawater, which has much higher ¹⁸⁷Os/¹⁸⁸Os. The concentration of Os in MORB is extremely low, far lower than in peridotites, meaning relatively small amounts of contamination could shift the isotope ratios.

Another interesting result of Re-Os studies has been the evidence for systematically low ¹⁸⁷Os/¹⁸⁸Os in the subcontinental lithosphere. Studies of pieces of subcontinental mantle carried to the surface as xenoliths in magmas have revealed that much of this mantle is poor in clinopyroxene and garnet and hence depleted in its basaltic component (the term "infertile" is often used to refer to such compositions), presumably as a result of previous episodes of melting. Surprisingly, these xenoliths often show evidence of incompatible element enrichment, including high ⁸⁷Sr/⁸⁶Sr and low ε_{Nd} . This latter feature is often attributed to reaction of the mantle lithosphere with very small degree melts percolating upward through it (a process termed "mantle metasomatism"). This process, however, apparently leaves the Re-Os system unaffected, so that ¹⁸⁷Re/¹⁸⁸Os and ¹⁸⁷Os/¹⁸⁸Os remain low. Figure 8.14 compares the γ_{Os} values of subcontinental lithosphere, oceanic periodotites (presumably representing the isotopic com-

position of the upper asthenospheric mantle, though to be the source of MORB), and oceanic basalts. Thus Re-Os studies are proving to be important in understanding the evolution of the subcontinental lithosphere, and its role in volcanism. We shall return to these topics in Chapter 11.

There has also been considerable interest in the Os isotope composition of seawater. The ¹⁸⁷Os/¹⁸⁸Os ratio of modern seawater is about seawater is about 8. Like that of ⁸⁷Sr/⁸⁶Sr, ¹⁸⁷Os/¹⁸⁸Os depends on



Figure 8.15. Os isotope composition of seawater over the last 80 Ma, from Peuker-Ehrenbrink et al. (1995). Gray field represents the authors best estimates of seawater Os isotopic composition.

W. M. WhiteGeochemistryCHAPTER &: RADIOGENIC ISOTOPE GEOCHEMISTRY

the balance of continental fluxes (e.g., rivers, with ¹⁸⁷Os/¹⁸⁸Os ~ 1.2) and oceanic crustal fluxes (e.g., hydrothermal activity, with ¹⁸⁷Os/¹⁸⁸Os ~ 0.13). In addition, however, cosmic fluxes (¹⁸⁷Os/¹⁸⁸Os ~ 0.13), which include both cosmic dust, which continually settles through the atmosphere into the oceans, and large meteorite impacts, may be significant for Os. Because of the low concentrations of Os in seawater, the Os isotopic composition of seawater cannot be measured directly. It can, however, be measured indirectly by analyzing the authigenic component in seawater, such as Mn nodules and the leachable fraction of clays.

Like Sr, the Os isotopic composition of seawater has changed over time (Figure 8.15). There are obvious similarities between the Os isotopic and Sr isotopic evolution of seawater, most notably the increase of both through the Tertiary period (the last 65 million years). This may in part reflect a decreasing hydrothermal flux resulting from decreasing sea floor spreading rates. There are also differences, which reflect the differing geochemical behavior of Sr and Os. The geochemical behavior of both at the surface of the Earth is related to carbon, but while Sr is concentrated in carbonates, Os is concentrated in organic-rich sediments. The very low Os isotope ratios at the Cretaceous-Tertiary boundary (65 Ma), may reflect a sudden input of meteoritic Os as a result of the impact of a large meteorite that apparently occurred then.

8.4.5 LA-CE

This is another system that is analytically challenging. ¹³⁸La is a rather rare isotope (an odd-odd), and its long half-life does not help matters. Worse, ¹³⁸Ce is also a trace isotope of Ce, which would be O.K. except for the problem of analyzing for a very rare isotope in the presence of a very common one (¹⁴⁰Ce). Lastly, ¹³⁸Ba is among the most abundant heavy isotopes (because it has 82 neutrons, a magic number), and can interfere with the analysis.

As Figure 8.16 shows, ¹³⁸Ce/¹⁴²Ce (or ¹³⁸Ce/¹³⁶Ce, a convention has not evolved) and ¹⁴³Nd/¹⁴⁴Nd are negatively correlated. This correlation is exactly what is expected because the parent, La, is more in-



Figure 8.16. Ce and Nd isotope ratios in oceanic basalts (Dickin, 1987)

compatible than the daughter, Ce, but the daughter of the Nd–Sm system is more incompatible than the parent. This leads to an expected anticorrelation between La/Ce and Sm/Nd ratios and, ultimately, between 138 Ce/ 142 Ce and 143 Nd/ 144 Nd. How well they are correlated is not yet clear.

8.4.6 U-TH-Pb

The U-Th-Pb system is somewhat of a special case since there are 3 decay schemes producing isotopes of Pb. In particular two U isotopes decay to two Pb isotopes, and since the two parents and two daughters are chemically identical, we get two decay systems for the price and one and together they provide a particularly powerful tool.

Let's explore the mathematics of this. Following convention, we will designate the 238 U/ 204 Pb ratio as μ , and the 232 Th/ 238U ratio as μ . The ratio 238 U/ 235 U is constant in the Earth at 137.88 (except in nuclear reactors and the Oklo deposit mentioned earlier). Now, we can write two versions of equation 8.18:

$${}^{207}Pb / {}^{204}Pb = {}^{207}Pb / {}^{204}Pb_0 + \frac{\mu}{137.88}(e^{\lambda_{235}t} - 1)$$
8.31

W. M. WhiteGeochemistryCHAPTER &: RADIOGENIC ISOTOPE GEOCHEMISTRY

and

$${}^{206}Pb / {}^{204}Pb = {}^{206}Pb / {}^{204}Pb_0 + \mu(e^{\lambda_{238}t} - 1)$$
8.32

Equations 8.31 and 8.32 can be rearranged by subtracting the initial ratio from both sides. For example, using Δ to designate the difference between the initial and the present ratio, equation 8.32 becomes:

$$\Delta^{206} Pb / {}^{204} Pb = \mu(e^{\lambda_{238}t} - 1)$$
8.33

Dividing the equivalent equation for ${}^{235}\text{U}$ - ${}^{207}\text{Pb}$ by equation 8.33 yields:

$$\frac{\Delta^{207}Pb/^{204}Pb}{\Delta^{206}Pb/^{204}Pb} = \frac{1}{137.88} \frac{(e^{\lambda_{235}t} - 1)}{(e^{\lambda_{238}t} - 1)}$$
8.34

The left-hand side of this equation is the slope of line on a plot of $^{207}\text{Pb}/^{204}\text{Pb}$ vs. $^{206}\text{Pb}/^{204}\text{Pb}$ such as Figure 8.17. The slope depends only on time and three constants (λ_{238} , λ_{235} , and $^{238}\text{U}/^{235}\text{U}$). Because its slope depends on time, a line on a plot of $^{207}\text{Pb}/^{204}\text{Pb}$ vs. $^{206}\text{Pb}/204\text{Pb}$ is an isochron, analogous to the isochrons in a plot such as Figure 8.6. We derived equation 8.34 by subtracting the initial ratio from the present one, but we could derive the identical equation by subtracting equations for two separate samples that share common initial ratios and time. Equation 8.34 differs from the conventional isochron equation (equation 8.18) in that terms for the initial values and the parent-daughter ratios do not appear. The point is, if we measure the Pb isotope ratios in a series of samples that all started with the same initial isotopic composition of Pb at some time t_0 , and which all remained closed to gain or loss of Pb or U since then, we can determine the age of this system without knowing the parent-daughter ratio.

The bad news is that equation 8.34 cannot be solved directly for t. However, we can guess a value of t, plug it into the equation, calculate the slope, compare the calculated slope with the observed one, revise our guess of t, calculate again, etc. Pretty laborious, but using a computer to make 'educated guesses' of t, it is pretty easy. If fact, there are algorithms available that converge to a high degree of accuracy after a few iterations^{*}. This kind of isochron is called a Pb-Pb isochron and the age derived from it is called a Pb-Pb age. A disadvantage of Pb-Pb isochrons is that we cannot determine initial ratios from them.

Because the half-life of ²³⁵U is much shorter than that of ²³⁸U, ²³⁵U decays more rapidly. As a result, on a plot ²⁰⁷Pb/²⁰⁴Pb vs. ²⁰⁶Pb/²⁰⁴Pb, Pb isotopic evolution follows curved paths. The exact path that is followed depends upon μ . Three such evolution curves are shown in Figure 8.15. All systems that begin with a common initial isotopic composition at time t_0 along a straight line at some later time t. This line is the Pb-Pb isochron.

As an example, let's consider the various bodies of the solar system. We assume that (and there has been no evidence to the contrary) when the solar system formed 4.55 billion years ago, it had a single, uniform Pb isotope composition, which we will refer to as primordial Pb. Planetary bodies formed shortly after the solar system formed, and we can reasonably assume that they have remained closed since their formation (i.e., they have neither gained nor lost Pb or U). There is no reason to assume that all planetary bodies started with the same ratio of U to Pb, i.e., the same value of μ . Indeed, there is good evidence that they did not. So Pb in each planetary body would evolve along a separate path that would depend on the value of μ in that body. However, at any later time *t*, the ²⁰⁷Pb/²⁰⁴Pb and ²⁰⁶Pb/²⁰⁴Pb ratios of all planetary bodies plot on a unique line. This line, called the *Geochron*, has a slope whose age corresponds to the age of the solar system, and it passes through the composition of primordial Pb.

Equation 8.34, in which the parent/daughter ratio term does not appear, turns out to be very useful for several reasons. First, for geochronological applications, U is a rather mobile element, so the closed system assumption is often violated. Often this mobility occurs late (perhaps occurring as the rock is unburied and reaches the surficial weathering zone). So a normal U-Pb isochron often gives erroneous results. But if the loss (or gain) has been sufficiently recent, Pb isotope ratios will not be significantly affected. Because of this, Pb-Pb ages calculated with equation 8.34 are robust with respect to recent mobility of U, and good ages can be obtained this way even when they cannot be obtained from a conventional U-Pb isochron (i.e., from the slope of 206 Pb/ 204 Pb vs. 238 U/ 204 Pb).

^{*} This kind of algorithm is available as a standard Add-In call the "Solver" in Microsoft Excel.

Geochemistry

Chapter O: Radiogenic Isotope Geochemistry

W. M. White



Figure 8.17. Evolution of Pb isotope ratios. The curve lines represent the evolutionary paths for systems having μ values of 8, 9, and 10. The hash marks on the evolution curves mark Pb isotope compositions 1.0, 2.0, and 3.0 Ga ago.

TABLE 8.3. INITIAL ISOTODIC COMPOSITIONS AND PAR-ENT-DAUGHTER RATIOS FOR THE EARTH

System	R_0	$R_{P/D}^{*}$	Source
⁸⁷ Rb- ⁸⁷ Sr	0.69898	~0.085	basaltic achondrites
¹⁴⁷ Sm- ¹⁴³ Nd	0.50670	0.1967	chondrites
¹⁷⁶ Lu- ¹⁷⁶ Hf	0.2797 [£]	0.0334	eucrites, chondrites
¹⁸⁷ Re- ¹⁸⁷ Os [†]	0.09677	0.397	chondrites
¹⁸⁷ Re- ¹⁸⁷ Os [‡]	0.8041	3.299	chondrites
²³⁵ U- ²⁰⁷ Pb	10.294	~0.058	Canyon Diablo troilite
²³⁸ U- ²⁰⁶ Pb	9.314	~8	Canyon Diablo troilite
²³² Th- ²⁰⁸ Pb	29.476	~32	Canyon Diablo troilite

*The parent daughter ratio is given as the present day value.

[†]Ratios for ¹⁸⁷Os / ¹⁸⁸Os and ¹⁸⁷Re / ¹⁸⁸Os

*Ratios for ¹⁸⁷Os/¹⁸⁶Os and ¹⁸⁷Re/¹⁸⁶Os

[£]Exact value disputed between 0.279628 and 0.279742

Reservoirs within the Earth, such as the crust and mantle and individual rock units, have clearly not been closed systems for 4.55 Ga, so their isotope ratios need not fall on the Geochron. The mean of all such reservoirs should, however, fall on the Geochron.

The initial isotopic composition of the Earth can be estimated from that in meteorites, as is the case for other isotope systems. Most terrestrial reservoirs do plot fairly close to the Geochron and close to an evolution curve corresponding to a value of μ of about 8. Table 8.3 lists estimated values for the initial isotopic composition of the Earth and solar system, as well as estimated parent/daughter ratios for the bulk Earth.

A second important application is 'mantle isochrons'. It is reasonable to assume that the isotope ratios in a volcanic rock will be the same as the isotope ratios of its source because the different isotope of an element, for example ²⁰⁶Pb and ²⁰⁴Pb, are chemically identical and cannot be fractionated by magmatic processes (which is to say that their partition coefficients are identical). But isotopes of different elements, for example ²³⁸U and ²⁰⁶Pb, are not chemically identical and can be fractionated. Thus the U/Pb or ²³⁸U/²⁰⁴Pb ratio of a volcanic rock is not necessarily the same as that of its source. As a result, we cannot calculate an 'age' of the source of a series of volcanic rocks using the normal isochron equation, 8.31 or 8.32. Since the parent-daughter ratio does not occur in equation 8.34, we can calculate an age using equation 8.34. Such 'mantle isochron' ages for

oceanic basalt sources are typically 1 to 2 Ga.

We cannot assume that the U-Pb ratio of the Earth is chondritic, indeed the evidence is that it is not. Thus unlike the Sm-Nd and Lu-Hf systems, we don't have any *a priori* constraints on the U/Pb ratio of the Earth. However, assuming the Earth formed at the same time and from the same materials as meteorites, and that it has remained closed to gain or loss of U and Pb, its ²⁰⁷Pb/²⁰⁴Pb and ²⁰⁶Pb/²⁰⁴Pb must fall on the Geochron just like any other planetary body. This applies only to the Earth as a whole. W. M. White

Geochemistry

CHAPTER O: RADIOGENIC ISOTOPE GEOCHEMISTRY



Figure 8.18. Pb isotope ratios in major terrestrial reservoirs. Typical lower continental crust and upper continental crust are represented by lower crustal xenoliths and modern marine sediments respectively (these somewhat underestimate the total variance in these reservoirs). MORB and oceanic islands represent the isotopic composition of upper mantle and deep mantle respectively.

If a system has experienced a decrease in U/Pb at some point in the past, its Pb isotopic composition will lie to the left of the Geochron; if its U/Pb ratio increased, its present Pb isotopic composition will lie to the right of the Geochron. U is more incompatible than Pb. This being the case, incompatible element depleted reservoirs should plot to the left of the Geochron, incompatible element enriched reservoirs should plot to the right of the Geochron (Figure 8.18). From the isotopic composition of other radiogenic elements, we would therefore predict that continental crust should lie to the right of the Geochron, as expected, but surprisingly, Pb isotope ratios of mantle-derived rocks also plot mostly to the right of the Geochron (Figure 8.18). This indicates the U/Pb ratio in the mantle has increased, not decreased as expected. This phenomenon is known as the *Pb paradox* and it implies that a simple model of crust-mantle evolution that involves only transfer of incompatible elements from crust to mantle through magmatism is inadequate. We will discuss this in greater detail in Chapter 11.

The only large geochemical reservoir known to plot to the left of the Geochron is the lower continental crust. It is as yet unclear whether the four reservoirs represented on Figure 8.18, the upper crust, the lower crust, the upper mantle (as represented by MORB), and the deep mantle (as represented by oceanic islands) average to a composition that lies on the Geochron. If not, there must be some other reservoir within the Earth that has unradiogenic Pb (i.e., Pb falling to the left of the Geochron).

We can combine the growth equations for ${}^{208}\text{Pb}/{}^{204}\text{Pb}$ and ${}^{206}\text{Pb}/{}^{204}\text{Pb}$ in a way similar to equation 8.34. We end up with:

$$\frac{\Delta^{208}Pb/^{204}Pb}{\Delta^{206}Pb/^{204}Pb} = \frac{\kappa(e^{\lambda_{23}t} - 1)}{(e^{\lambda_{23}t} - 1)}$$
8.35

Geochemistry

where, again, κ is the ²³²Th/²³⁸U ratio. Provided κ is constant in a series of cogenetic rocks (by cogenetic we mean that at *t*=0 they all had identical Pb isotope ratios), the slope of an array on a plot of ²⁰⁸Pb/²⁰⁴Pb and ²⁰⁶Pb/²⁰⁴Pb will depend only on *t* and κ , and not on the parent-daughter ratios. We can calculate κ from the slope of the data on a plot of ²⁰⁸Pb/²⁰⁴Pb—²⁰⁶Pb/²⁰⁴Pb. U and Th are both highly incompatible elements. They have similar, but not identical, partition coefficients in most minerals. It seems unlikely that they will behave identically, and hence unlikely that κ will very often be constant. It can be shown that if κ varies and is positively correlated with variations in μ then equation 8.35, when solved for κ , will actually overestimate it.

8.4.7 U and TH Decay Series Isotopes

U and Th do not decay directly to Pb, rather the transition from U and Th to Pb passes through many intermediate radioactive daughters (Figure 8.19). Many of these daughters have very short half-lives, ranging from milliseconds to minutes, and are of little use to geochemists. However, a number of these intermediate daughters have half-lives ranging from days to hundreds of thousands of years and do provide useful information about geological processes. Table 8.4 lists half-lives and decay constants of some of the most useful of these nuclides.

The half-lives of all of these daughter isotopes are short enough so that any atoms present when the Earth formed have long since decayed

(to Pb). They exist in the Earth (and in all other bodies of the solar system) only because they are continually produced by the decay of U and Th. The abundance of such an isotope depends on the balance between its own radioactive decay and its production by the decay of its parent:

$$\frac{dN_D}{dt} = \lambda_p N_p - \lambda_D N_D \quad 8.36$$

Table 8.4. Half-lives and decay constants of long-lived U and TH daughters

Nuclide	Half-life, yrs	Decay constant, yr ⁻¹	Parent
²³⁴ U	246,000	2.794 x 10 ⁻⁶	²³⁸ U
²³¹ Pa	32,480	2.134 x 10 ⁻⁵	²³⁵ U
²³⁰ Th	75,200	9.217 x 10 ⁻⁶	²³⁸ U
²²⁶ Ra	1,622	4.272 x 10 ⁻⁴	²³⁸ U
²²⁸ Ra	6.7	1.06 x 10 ⁻¹	²³² Th
²¹⁰ Pb	22.26	3.11 x 10 ⁻²	²³⁸ U

where subscripts P and D refer to par-

ent and daughter respectively. This equation says simply that the rate of change of the abundance of the daughter isotope is equal to the rate of production less the rate of decay. This can be integrated to give:

$$N_D = \frac{\lambda_P}{\lambda_D - \lambda_P} N_P^0 \left(e^{-\lambda_P t} - e^{\lambda_D t} \right) + N_D^0 e^{-\lambda_D t}$$
8.37

Scientists dealing with the intermediate daughters of U and Th (and it is the daughters of ²³⁸U that are of the most interest), generally work with activities, measured in the number of decays per unit time, rather than atomic abundances. One reason for this is that the abundances of these isotopes are generally determined by detecting their decay. Indeed, the shorter-lived ones are so rare they cannot be detected any other way. The other reason will become apparent shortly. Activities are related to atomic (molar) abundances by the basic equation of radioactive decay:

W. M. White

CHAPTER O: RADIOGENIC ISOTOPE GEOCHEMISTRY



Figure 8.19. Part of the chart of the nuclides showing the series of decays that occur as ²³⁸U, ²³⁵U, and ²³²Th are transformed to ²⁰⁶Pb, ²⁰⁷Pb, and ²⁰⁸Pb respectively. Nuclides of the ²³⁸U series are shown in red, the ²³²Th series in gray, and the ²³⁵U series in black.

$$\frac{dN_i}{dt} = -\lambda_i N_i \tag{8.38}$$

where dN_i/dt is the *activity* of *i*. Thus so long as we know the decay constant, we can calculate activity from concentration and visa versa. Activity is most commonly expressed in units of *disintegrations per minute* (dpm) per gram or per mole and is denoted by writing parentheses around the isotope (or the isotope ratio). Thus (²³⁴U) is the activity of ²³⁴U.

The state of *radioactive equilibrium* is the condition where their *activities* of the daughter and the parent are equal, i.e.:

$$\frac{dN_D}{dt} = \frac{dN_P}{dt}$$
8.39

It shares the same fundamental characteristic as the chemical equilibrium state: it is the state that will eventually be achieved if the system is not perturbed (remains closed). We can demonstrate this is so in two ways. The first is a simple mathematical demonstration. The equilibrium state is the steady state where the abundance of the daughter does not change, i.e., where the left hand side of 8.36 is zero:

$$0 = \lambda_P N_P - \lambda_D N_D \tag{8.40}$$

We substitute the dN/dt terms for the λN terms in 8.40, rearrange, and we obtain 8.39. QED.

The second demonstration is a thought experiment. Imagine a hopper, a grain hopper with an open top and a door in the bottom. The door is spring loaded such that the more weight placed on the door, the wider it opens. Suppose we start dropping marbles into the hopper at a constant rate. The weight of marbles accumulating in the hopper will force the door open slightly and marbles will start falling out at a slow rate. Because the marbles are falling out more slowly than they are falling in, the number and weight of marbles in the hopper will continue to increase. As a result, the door will continue to open. At some point, the door will be open so wide that marbles are falling out as fast as they are falling in. This is the steady state. Marbles can no longer accumulate in the hopper and hence the door is not forced to open any wider. The marbles falling into the door are like the decay of the parent isotope.

The marbles in the hopper represent the population of daughter isotopes. Their decay is represented by their passing through the bottom door. Just as the number of marbles passing through the door depends on the number of marbles in the hopper, the activity (number of decays per unit time) of an isotope depends on the number of atoms present.

Geochemistry

If the rate of marbles dropping into the hopper decreases for some reason, marbles will fall out of the hopper faster than they fall in. The number of marbles in the hopper will decrease; as a result the weight on the door decreases and it starts to close. It continues to close (as the number of marbles decreases) until the rate at which marbles fall out equals the rate at which marbles fall in. At that point, there is no longer a change in the number of marbles in the hopper and the position of the door stabilizes. Again equilibrium has been achieved, this time with fewer marbles in the hopper, but nevertheless at the point where the rate of marbles going in equals the rate of marbles going out. The analogy to radioactive decay is exact.

Thus when a system is disturbed it will ultimately return to equilibrium. The rate at which it returns to equilibrium is determined by the decay constants of the parent and daughter. If we know how far out of equilibrium the system was when it was disturbed, we can determine the amount of time that has passed since it was disturbed by measuring the present rate of decay of the parent and daughter. Equilibrium is approached asymptotically, and these dating schemes are generally limited to time scales less than 5-10 times the half-life of the daughter. At longer times, the difference between actual activities and equilibrium ones becomes too small to measure reliably.

There are several useful dating methods based on the degree of disequilibria between U decay series nuclides. Our first example is ²³⁴U-²³⁸U dating of sediments. As may be seen from Figure 8.19, ²³⁴U, which has a half-life of 246,000 years, is the 'great-granddaughter' of ²³⁸U. For most purposes, the half-lives of the two intermediate daughters ²³⁴Th and ²³⁴Pa are so short that they can be ignored (because they quickly come into equilibrium with ²³⁸U). As it turns out, ²³⁴U and ²³⁸U in seawater are not in equilibrium, i.e., the ²³⁴U/²³⁸U activity ratio is not 1. It is fairly constant, however, at about 1.15. The reason for this disequilibrium is that ²³⁴U is preferentially leached from rocks because ²³⁴U is present in rocks in damaged sites. It occupies the site of a ²³⁸U atom that has undergone α -decay. The α particle and the recoil of the nucleus damage this site. Since it occupies a damaged site, it is more easily leached or dissolved from the crystal during weathering than ²³⁸U. The oceans collect this 'leachate', hence they are enriched in ²³⁴U. When U enters a sediment it is isolated from seawater (not necessarily immediately) and ²³⁴U decays faster than it is created by decay of ²³⁸U, so it slowly returns to the equilibrium condition where (²³⁴U/²³⁸U) = 1.

Let's now consider the problem from a mathematical perspective and derive an equation describing this return to equilibrium. We can divide the 234 U activity in a sample into that which is *supported* by 238 U, i.e., that amount in radioactive equilibrium with 238 U and that amount that is excess, i.e., *unsupported* by 238 U:

$$(^{234}U) = (^{234}U)_{s} + (^{234}U)_{u}$$
8.41

Subscripts *s* and *u* denote supported and unsupported. The activity of the excess 234 U decreases with time according to equation 8.4, which we can rewrite as:

$${}^{234}U)_U = ({}^{234}U)_U^0 e^{-\lambda_{234}t}$$
8.42

where the superscript naught denotes the initial unsupported activity (at t = 0). We can also write:

$${}^{234}U)^0_U = ({}^{234}U)^0 - ({}^{234}U)_s$$
8.43

which just says that the initial unsupported activity of ²³⁴U is equal to the total initial activity of ²³⁴U less the (initial) supported activity of ²³⁴U. *Since to a very good approximation the activity of the parent,* ²³⁸U, *does not change over times on the order of the half-life of* ²³⁴U *or even ten half-lives of* ²³⁴U, *the present* ²³⁸U *activity is equal to the activity at* t = 0 (we make the usual assumption that the system is closed). By definition the supported activity of ²³⁴U is equal to the activity of ²³⁴U is equal to the activity of ²³⁸U, both now and at t = 0, hence, 8.41 can be expressed as:

$${}^{(234}U) = {}^{(238}U) + {}^{(234}U)_{\mu} \qquad 8.44$$

and 8.43 becomes
$$({}^{234}U)^0_u = ({}^{238}U) + ({}^{234}U)_u$$
 8.45

November 14, 2005

Substituting 8.45 into 8.42 yields:

$${}^{(234}U)_{u} = \left[{}^{(234}U)^{0} - {}^{(238}U) \right] e^{-\lambda_{234}t}$$
8.46

Substituting 8.46 into 8.44, we have:

$${}^{(234}U) = {}^{(238}U) + \left[{}^{(234}U)^0 - {}^{(238}U) \right] e^{-\lambda_{234}t}$$
8.47

Just as for other isotope systems, it is generally most convenient to deal with ratios rather than absolute activities (among other things, this allows us to ignore detector efficiency provided the detector is equally efficient at all energies of interest), hence we divide by the activity of ²³⁸U:

$$\left(\frac{^{234}U}{^{238}U}\right) = 1 + \left[\frac{(^{234}U)^0 - (^{238}U)}{(^{238}U)}\right]e^{-\lambda_{234}t}$$
8.48

or since ²³⁸U = ²³⁸U°:
$$\left(\frac{^{234}U}{^{238}U}\right) = 1 + \left[\left(\frac{^{234}U}{^{238}U}\right)^0 - 1\right]e^{-\lambda_{234}t}$$
 8.49

Corals, for example, concentrate U. If we measure the $(^{234}U/^{238}U)$ ratio of an ancient coral, and can assume that the seawater in which that coral grew had a $(^{234}U/^{238}U)$ the same as modern seawater (1.15), then the age of the coral can be obtained by solving equation 8.49 for *t*.

Because the disequilibrium between ²³⁰Th and ²³⁸U can be much larger than between ²³⁴U and ²³⁸U, ²³⁰Th-²³⁸U disequilibria is a more commonly used dating scheme than is ²³⁴U-²³⁸U. ²³⁰Th is the daughter of ²³⁴U (the decay chain is ²³⁸U \rightarrow ²³⁴Th + α , ²³⁴Th \rightarrow ²³⁴Pa + β^{-} , ²³⁴Pa \rightarrow ²³⁴U + β^{-} , ²³⁴U \rightarrow ²³⁰Th + α). To simplify the math involved, let's assume ²³⁴U and ²³⁸U are in radioactive equilibrium. In high temperature systems, this is a very good assumption because α -damaged sites, which cause the disequilibrium noted above, are quickly annealed. With this assumption, we can treat the production of ²³⁰Th as if it were the direct decay product of ²³⁸U. We write an equation analogous to 8.41 and from it derive an equation analogous to 8.47:

$${}^{(230}Th) = {}^{(238}U) + \left[{}^{(230}Th)^0 - {}^{(238}U) \right] e^{-\lambda_{230}t}$$
8.50

We divide by the activity of ²³²Th:

$$\left(\frac{{}^{230}Th}{{}^{232}Th}\right) = \left(\frac{{}^{238}U}{{}^{232}Th}\right) + \left[\left(\frac{{}^{230}Th}{{}^{232}Th}\right)^0 - \left(\frac{{}^{238}U}{{}^{232}Th}\right)\right]e^{-\lambda_{230}t}$$
8.51

and rearranging:

$$\left(\frac{^{230}Th}{^{232}Th}\right) = \left(\frac{^{230}Th}{^{232}Th}\right)^0 e^{-\lambda_{230}t} + \left(\frac{^{238}U}{^{232}Th}\right)(1 - e^{-\lambda_{230}t})$$
8.52

 230 Th/ 238 U is commonly used to date sediments and to determine sedimentation rates. However, unlike the case of (234 U/ 238 U)°, (230 Th/ 232 Th)° is not *a priori* known, so there are 2 unknowns in equation 8.52. As was the case for isochrons, however, we can solve for these 2 unknowns if we have a series of two or more measurements on sediments with the same initial (230 Th/ 232 Th). Example 8.5 demonstrates how this is done for the case of a Mn nodule. In other cases, corals, for example, we can assume that (230 Th/ 232 Th)° is the same as in modern seawater.

Th–U disequilibria may also be used for dating lavas, and we now turn our attention briefly to this application. Equation 8.52 has the form of a straight line on a $(^{230}\text{Th}/^{232}\text{Th}) - (^{238}\text{U}/^{232}\text{Th})$ plot; the first term is the intercept and $(1 - e^{-\lambda_{230}t})$ is the slope. Since the slope is proportional to time, a line on a $(^{230}\text{Th}/^{232}\text{Th}) - (^{238}\text{U}/^{232}\text{Th})$ plot is an isochron, though unlike a conventional isochron the intercept also changes with time (Figure 8.21).

To understand how this works, imagine a crystallizing magma with homogeneous (230 Th/ 232 Th) and (238 U/ 232 Th) ratios. Th and U partition into different minerals differently, so the minerals will have variable (238 U/ 232 Th) ratios but constant (230 Th/ 232 Th) ratios (assuming crystallization occurs quickly

Geochemistry

W. M. White

CHAPTER O: RADIOGENIC ISOTOPE GEOCHEMISTRY

Example 8.5. Determining the growth rate of a Mn nodule

The tops of manganese nodules grow by precipitation of Mn-Fe oxides and hydroxides from seawater. They are known to grow very slowly, but how slowly? If we assume the rate of growth is constant then depth in the nodule should be proportional to time. If z is the depth in the nodule, and s is the growth (sedimentation) rate, then:

$$t = z/s \tag{8.53}$$

and 8.52 becomes:

$$\left(\frac{^{230}Th}{^{232}Th}\right) = \left(\frac{^{230}Th}{^{232}Th}\right)^0 e^{-\lambda_{230}z/s} + \left(\frac{^{238}U}{^{232}Th}\right)(1 - e^{-\lambda_{230}z/s}) \qquad 8.54$$

At the surface of the nodule, z = 0, so the exponential terms both go to 1 and the measured activity ratio is initial activity ratio. Having a value for (²³⁰Th/²³²Th)°, equ. 8.54 can then be solved for s, the growth rate if measurements are made at some other depth.

In practice, however, it is difficult to obtain a sample exactly at the surface: a finite amount of



material is required for analysis, and this often translates into a la yer of sev eral mm thickness. Equation 8.54 is solved in that instance by less direct means. For example, consider the d ata shown in Figure 8.20 on a Pacific manganese nodule reported by Huh and Ku (1984). In this plot of $(^{230}Th/^{232}Th)$ vs. depth, the initial ratio is the intercept of the best-fit line through the data. A growth rate was obtained by obtaining an in itial guess of the initial $(^{230}\text{Th}/^{232}\text{Th})$, then iteratively refining the solution to 8.54 by minimizing the difference between computed and observed activity ratios. A growth rate of 49.5 mm/Ma and a $(^{230}\text{Th}/^{232}\text{Th})$ of 77.7 was found to best fit the observations.

compared to the half-life of Th) since these two isotopes are chemically identical. Thus the minerals will plot on a horizontal line in Figure 8.21 at t = 0. After the system closes, ²³⁸U and ²³⁰Th will begin to come to radioactive equilibrium (either ²³⁰Th will decay faster than it is produced or visa versa, depending on whether (²³⁰Th/²³⁸U) is greater than or less than 1, the equilibrium value). Thus the original horizontal line will rotate, as in a conventional isochron diagram, but unlike the conventional isochron diagram, the intercept also changes. The rotation occurs about the point where (²³⁰Th/²³²Th) = (²³⁸U/²³²Th), which is known as the *equipoint*. As *t* approaches infinity, the exponential term approaches 1 and:

$$\lim_{t \to \infty} \left(\frac{{}^{230}Th}{{}^{232}Th} \right) = \left(\frac{{}^{238}U}{{}^{232}Th} \right)$$
8.55

Thus the equilibrium situation, the situation at $t=\infty$, is $(^{230}\text{Th}/^{232}\text{Th}) = (^{238}\text{U}/^{232}\text{Th})$. In this case, all the minerals will fall on a line, having a slope of 1. This line is known as the *equiline*.

²²⁶Ra is another relatively long-lived nuclide ($t_{1/2} = 1600$ yr) that has proved useful in dating igneous rocks. The fundamentals are precisely analogous to those we have discussed for ²³⁴U and ²³⁰Th, with one exception. Unfortunately, Ra has no stable isotope to which one can ratio ²²⁶Ra. Therefore, the assumption is made that Ra behaves as Ba, and the abundance of Ba is used to form a ratio.

Figure 8.22 shows ²³⁰Th–²³⁸U and ²²⁶Ra–²³⁰Th isochrons obtained on mineral separates, groundmass, and a whole rock sample from the Castle Creek andesite of Mt. St. Helens. Conventional dating tech-

Geochemistry

Geochemistry

CHAPTER O: RADIOGENIC ISOTOPE GEOCHEMISTRY

W. M. White



Figure 8.21. (a) 230 Th— 238 U isochron diagram. The (238 U/ 232 Th) of the source is given by the intersection of the isochron with the equiline. (b) shows how the slope changes as a function of time. After Faure (1986).

niques indicate this lava is only about 2000 years old, but both the ²³⁰Th–²³⁸U and ²²⁶Ra-²³⁰Th give much older ages. The 230 Th/ 232 Th- 238 U/ 232 Th data define a 27±12 ka isochron. This suggests the crystals in the magma may have crystallized as much as 20,000 years or more before eruption. The age²²⁶Ra-²³⁰Th is less precise because the slope is only slightly less than 1, and all samples seem to contain excess ²²⁶Ra. Nevertheless, the slope of the data is clearly greater than that of a 2 ka isochron, so these data also suggest the crystallization significantly predated According to Volpe and eruption. Hammond (1991) the difference in the ²³⁰Th-²³⁸U and ²²⁶Ra-²³⁰Th ages may reflect changes in melt composition, perhaps due to assimilation of country rock or addition of a Ra-rich gas or fluid phase to the magma. 230 Th $-^{238}$ U and 226 Ra $-^{230}$ Th ages of the more recent eruptive products of Mt. St. Helens, including the 1980-1986 eruptive phase, were closer to the eruption ages. For example, Volpe and Hammond (1991) obtained a 230 Th- 238 U age of 6±4 ka on the 1982 dacite. Volpe and Hammond (1991) suggested that there was a significant volume of older magma still present in the magma chamber when the most recent eruptive stage of Mt. St. Helens began about 4000 years ago, and that this older magma has been progressively replaced by fresh magma over the past 4000 years.

There are also important geochemical, as opposed to geochronological, inferences that can be made from ratios of (²³⁰Th/²³²Th) and (²³⁸U/²³²Th) in zero-age igneous rocks. The degree of disequilibrium between ²³⁰Th and ²³⁸U in such rocks depends on the extent of fractionation between Th and U during melting and the rate at which melting occurs. This in turn depends on the mineralogy in the mantle undergoing melting, the extent of melting, rates of melt percolation, and other factors. Thus some interesting inferences about melting in the mantle can be made from these isotope ratios. We will discuss these in subsequent chapters.

In the past, the activities of U-decay series isotopes were measured by alpha-counting and fairly large quantities of material are necessary. Improvements in mass spectrometry made it possible to measure 234 U, 230 Th, and other key nuclides such as 231 Pa and 226 Ra on smaller quantities of material and with much better precision than α counting. This has led to a considerable expansion of applications of U-decay series isotopes.

8.4.8 Isotopes of He and Other Rare Gases

Helium

Alpha particles are, of course, simply ⁴He nuclei, and therefore ⁴He is produced by alpha decay of U and Th. Thus the ratio of ⁴He/³He varies in the Earth. Unlike virtually all other elements, He is not conserved on the Earth: much of the He present when the Earth formed has been subsequently lost. Be-

Geochemistry

W. M. White CHAPTER &: RADIOGENIC ISOTOPE GEOCHEMISTRY

ing a rare gas, He does not combine chemically with anything, and it also diffuses rapidly (Ever had a He balloon? How long did it last?) Helium brought to the Earth's surface by magmatism eventually finds its way into the atmosphere. Once there, it can escape from the top of atmosphere because of its low mass*. (H² also escapes from the atmosphere, but most H is bound up as water or hydrous minerals, so relatively little is lost.) Since ⁴He is continually produced and, for all practical purposes, ³He is not, it should not be surprising that the ⁴He/³He ratio in the Earth is a very large number. For some reason, perhaps because geochemists like to deal with very small numbers rather than very large ones, or perhaps because it actually ³He that is most interesting, the He isotope ratio is generally expressed as ³He/⁴He, in contradiction to the normal convention of placing the radiogenic isotope in the numerator. We will adhere to this particular convention of not adhering to the convention and use the ${}^{3}\text{He}/{}^{4}\text{He}$ ratio.

The ${}^{3}\text{He}/{}^{4}\text{He}$ ratio of the atmosphere is 1.384 x 10⁻⁶. Since this ratio is uniform and atmospheric He available to all laboratories, it is a convenient standard and moreover provides the basis for a convenient normalization. He ratios are often reported and discussed relative to the atmospheric Figure 8.22. ²³⁰Th-²³⁸U and ²²⁶Ra-²³⁰Th mineral iso- R_A indicates the atmospheric ratio.

duced in the Earth. It is produced in very small Hammond (1991). through the nuclear reaction: quantities



value. The units are called simply R/R_A , where chrons for the Castle Creek and esite of Mt. St. Helens. ¹⁴C and tree ring chronology indicate this lava Actually, it is not quite true that ³He is not pro- was erupted 1700-2200 years ago. After Volpe and

 ${}^{6}\text{Li}(n,\alpha) \rightarrow {}^{3}\text{H}(\beta) \rightarrow {}^{3}\text{He}$, which is to say ${}^{6}\text{Li}$ is excited by the capture of a neutron that has been produced by U fission, and decays through the emission of an alpha particle to tritium, which beta decays to ³He. As a result He in crustal rocks has an $R/R_A \approx 0.1 - 0.01$ (the exact ratio varies with the Li/U ratio). ³He can also be produced in the atmosphere and surface of the Earth by cosmic ray spallation reactions. Up to about 25 years ago, all He was thought to be a product of these processes (radiogenic and cosmogenic). Then helium with R/R_A around 1.22 was discovered in Pacific deep water in 1969 (Clark et al., 1969). Subsequently, Lupton and Craig (1975) discovered that mid-ocean ridge basalt glasses had high ${}^{3}\text{He}/{}^{4}\text{He}$ ratios. The gases are trapped in the glassy rims of basalts by the combined effects of hydrostatic pressure and rapid quenching. Measurements in 1974 and 1975 found such basalts had R/R_A \approx 9. It was clear that a 'primordial' He component was trapped in the Earth's interior and is constantly escaping. Subsequent measurements have shown that the ³He/⁴He of MORB is quite uniform around $R/R_A \approx 8.5\pm 1$. This is only primordial in a relative sense: gas-rich meteorites have an $R/R_A \approx 200$. Thus even in MORB, He is 95% radiogenic.

^{*}At equilibrium the average kinetic energy of all variety of molecules in a gas will be equal. Since kinetic energy is related to velocity by $E=1/2mv^2$, and since the mass of He is lower than for other species, the velocities of He atoms will be higher and more likely to exceed the escape velocity. In actuality, however, escape of He is more complex than simple thermal acceleration and involves other processes such as acceleration of He ions by the Earth's magnetic field.



Figure 8.23. He isotope ratios in various terrestrial and solar system materials. "Solar" and "Planetary" refer to the solar and planetary components in meteorites. "Crustal rocks" shows the values expected for *in situ* production from α decay and neutron-induced nuclear reactions.

Figure 8.23 illustrates the He isotopic composition of various terrestrial reservoirs. Oceanic islands and other hotspots often have even higher ratios, up to about $R/R_A \approx 40$, though some islands, Tristan da Cunha for example, have lower ratios: $R/R_A \approx 5$ (e.g., Kurz et al., 1982). This suggests that most oceanic island basalts are derived from a less degassed, and in that sense more primordial, reservoir than MORB. This is consistent with the mantle plume hypothesis.

Geochemistry

Island arc volcanics (IAV) seem also to be fairly uniform with $R/R_A \approx 6$ (Lupton, 1983). Ratios lower than MORB suggest the presence of a slabderived component, but also indicate most of the He in IAV comes from the mantle wedge (a conclusion similar to the one reached from other radiogenic isotopes).

Two interesting developments are the discovery of cosmogenic He in Hawaiian basalts (Kurz, 1987), and very high R/R_A in some very slowly accumulating marine sediments. The former is a result of the ⁶Li(n,α) \rightarrow ³H(β) \rightarrow ³He reaction instigated by cosmic-ray produced neutrons and spallation*. The cosmogenic component decreases rapidly with depth of the rock (it is largely restricted to the top meter) and increases with elevation above sealevel (the effect was noticed at the summit of Haleakala at 2000m). Thus one should be suspicious of high ³He/⁴He ratios in old (100,000 yrs or more) subareal

rocks. R/R_A as high as 226 (and as low as 0.03) have also been observed in diamonds. The meaning of this was not immediately clear, but the origin of these diamonds is unknown, and the most probable explanation is that the diamonds were mined from a placer and that the He is cosmogenic. Very high

 R/R_A in some very slowly accumulating marine sediments is probably an effect of accumulation of cosmic dust in ⁴⁰Ar/³⁶Ar ratios sediment. lower than atmospheric have also been observed in deep-sea sediment. A final development was the realization that He ratios of magma could change, through the combined effect of diffusion out of the magma and radiogenic growth of He, on the time scale of residence time of magma in a magma chamber (Zindler and Hart, 1986). This means we must be cautious of low ³He/⁴He ratios as well.



Figure 8.24. Contours of relative ³He concentrations in seawater over the East Pacific Rise (Lupton and Craig, 1981).

^{*} Spallation is the process in which a necleus breaks into smaller nucleii as a result of a collision with a very high energy particle such as a cosmic ray.

Geochemistry

Since He is injected at mid-ocean ridges, particularly in the Pacific, at depths of 2500-3000 meters, He isotopes can be used to trace water mass movements. Figure 8.24 shows the He plume of hydrothermal activity on the East Pacific Rise. Interestingly, the plume indicates water is flowing in the direction opposite to that which the physical oceanographers had thought. Another application is prospecting for ridge-crest hydrothermal activity. Several hydrothermal areas (including the first one) have been discovered from the He isotope anomalies in the water column. ⁴He is also used as a tracer in U prospecting.

8.4.8.1 NEON

Neon has perhaps the most puzzling isotopic systematics of the noble gases. Ne has three isotopes, ²⁰Ne, ²¹Ne, and ²²Ne. Though none of them are radiogenic, the isotopic composition of neon varies in the Earth as well as extraterrestrial materials (Figure 8.25). Two processes have apparently produced these isotopic variations: mass-dependent fractionation (which we cover in the next chapter) and nuclear reactions such as ¹⁸O(α ,n) \rightarrow ²¹Ne, ²⁴Mg(n, α) \rightarrow ²¹Ne, and ²⁵Mg(n, α) \rightarrow ²²Ne, with the α particles and neutrons coming from α decay and fission respectively. Production rates are quite small, for example about 4.2 × 10⁻²² cc/g of ²¹Ne in the mantle. The production rate of ²¹Ne is about an order of magnitude higher than that of ²⁰Ne and ²²Ne, and ²¹Ne is the least abundant of the Ne isotopes (0.27%), so the effect of these reactions is to increase the ²¹Ne/²²Ne ratio without changing the ²⁰Ne/²²Ne ratio. In addition, nucleosynthetic processes are apparently responsible for some of the Ne isotopic variations in meteorites (see Chapter 10).

Noble gases trapped in basalts from mid-ocean ridges and basalts and mantle xenoliths from oceanic islands form arrays that diverge from atmospheric composition to compositions with higher 20 Ne/ 22 Ne and 21 Ne/ 22 Ne (e.g., Sarda, et al., 1988; Poreda and Farley, 1992). Accroding to Farley and Poreda (1993), the array reflect contamination of mantle Ne by atmospheric Ne; in essence, they are mixing arrays between mantle components with high 20 Ne/ 22 Ne and variable 21 Ne/ 22 Ne ratios. The very low concentrations of He in the atmosphere minimizes the problem of contamination in basalts; but atmospheric concentrations of the heavier.

centrations of the heavier noble gases, including Ne, are higher, so this is much more of a problem in the case of these gases.

Even though the arrays themselves are the products of contamination, the variation in the slope of the arrays is significant and indicates variation of ²¹Ne/²²Ne in the mantle, with the MORB source having higher ²¹Ne/²²Ne and OIB sources. This would be consistent with the MORB source being more degassed and hence having lower Ne concentrations. Addition of nuclogenic Ne would have a larger effect on ²¹Ne/²²Ne in that case.

Two hypotheses have been proposed to explain the observed isotopic variations in the Earth. In



Figure 8.25. Ne isotope ratios in terrestrial materials. The line marked "MDF" is the mass dependent fractionation line. The line has a slope of 2 because the change in ${}^{20}\text{Ne}/{}^{22}\text{Ne}$ resulting from fractionation should be twice that of ${}^{21}\text{Ne}/{}^{22}\text{Ne}$ because the mass difference between ${}^{20}\text{Ne}$ and ${}^{22}\text{Ne}$ is twice that between ${}^{21}\text{Ne}$ and ${}^{22}\text{Ne}$. OIB arrays are data fields for basalts and mantle xenoliths from Samoa and Hawaii.

the model advocated by Japanese workers (e.g., Honda et al., 1991; Ozima and Zahnle, 1993), the Earth initially formed with an Ne isotopic composition similar to that of solar wind. During the Earth's very early history, much of Earth's atmophile element inventory escaped from the interior, perhaps during the existence of a magma ocean, to form a hot, massive atmosphere. Much of the hydrogen and helium and a significant fraction of the neon escaped the atmosphere during this time. The efficiency of escape will depend on the mass of the molecule, so that the proportional loss of ²⁰Ne is greatest and that of ²²Ne the least. Thus the isotopic composition of the residual atmosphere has lower ²⁰Ne/²²Ne and ²¹Ne/²²Ne than the original one.

Geochemistry

Subsequently, the ²¹Ne/²²Ne in the degassed mantle increased due to nucleogenic production of ²¹Ne. Portions of the mantle that experienced greater degassing should have higher ²¹Ne/²²Ne than less degassed portions. Thus MORB apparently sample the most degassed mantle, presumably the upper mantle, while oceanic island basalts and xenoliths are apparently derived from a less degassed part of the mantle, presumably mantle plumes originating in the deep mantle.

Allègre et al. (1993) proposed a somewhat different model, one inspired by the observation that slowly accumulating pelagic sediments have Ne isotope ratios different from atmospheric, apparently due to the presence of cosmic dust and micrometeorites in the sediment. They suggest that the Earth original formed with a Ne isotopic composition similar to the present mantle. According to them, subduction of deep sea sediments has increased the ²⁰Ne/²²Ne ratio of the degassed mantle. Nucleogenic production in the mantle then increases the ²¹Ne/²²Ne ratio, and this effect has been greatest in the most degassed portion of the mantle, which is presumably sampled by MORB.

8.4.8.2 K-AR-CA

There are several aspects of the K-Ar system that make it particularly advantageous for some geochronological applications. First ⁴⁰K has one of the shortest half-lives of the decay systems in Table 8.2, making it more useful for dating young rocks than the other systems. In favorable cases, K-Ar ages on rocks as young as 30,000 yrs can be obtained with useful precision. Second, Ar is a rare gas and in most lavas it is virtually totally lost upon eruption. With no initial component, an age can be obtained from a single K-Ar determination. Being a rare gas, Ar diffuses somewhat more readily than other elements. Further, K is concentrated in sheet silicates such as micas, which have comparatively open crystal structures. As a result, the closure temperature for the K-Ar system is low, making it a useful system for study of relatively low temperature phenomena (such as petroleum genesis). This of course, can be a disadvantage, as it means the K-Ar system is reset rather easily.

From a geochemical viewpoint, the K-Ar system is of most interest with respect to the degassing of the Earthand evolution of the Earth's atmosphere. The most important questions in this regard are: what proportion of the Earth has been degassed, how extreme has this outgassing been, and when did this degassing occur? To briefly illustrate how K-Ar systematics can address these problems, consider the second question. We can think of two extreme possibilities: outgassing occurred yesterday; outgassing occurred simultaneously with the formation of the Earth. If outgassing occurred yesterday, then the ${}^{40}\text{Ar}/{}^{36}\text{Ar}$ ratio of the Earth's interior will be the same as that of the atmosphere. If outgassing occurred early in Earth's history then the 40 Ar/ 36 Ar ratio of the atmosphere should be close to the initial ratio for the solar system and the ⁴⁰Ar/³⁶Ar of the Earth's interior should be much higher because of the decay of ⁴⁰K over the history of the Earth. In actuality, neither of these models match observation. The ⁴⁰Ar/³⁶Ar ratio measured in MORB, in which noble gases are trapped by the pressure prevailing on the ocean floor, is very much higher than in the atmosphere. Ratios in MORB can be as high as 40,000, whereas that of the atmosphere is uniform at 295.5. So the Earth did not outgas yesterday. On the other hand, the atmospheric ratio is much higher than the solar system initial ratio. Indeed, the solar system initial ratio is less than 1, so most of the Ar in the atmosphere is actually radiogenic. Consequently, we can conclude that substantial degassing occurred after the Earth formed.

Now let's consider the first question: how much of the Earth has degassed. As we mentioned essentially all the Ar in the atmosphere is radiogenic. Since both the concentration of Ar and its isotopic composition are uniform in the atmosphere, it is fairly easy to estimate the amount of ⁴⁰Ar in the atmosphere, which works out to be 1.65×10^{18} moles. The K concentration is the silicate part of the Earth is about 250 ± 50 ppm, which corresponds to $2.56\pm0.5 \times 10^{22}$ moles of K. Over the 4.56 Ga of Earth's his-

Geochemistry

CHAPTER O: RADIOGENIC ISOTOPE GEOCHEMISTRY

W. M. White

tory, this amount of K should have produced about $3.5\pm0.7 \times 10^{18}$ moles of Ar. Thus about half (±10%) of the Ar that we expect has been produced by radioactive decay in now in the atmosphere; the remainder must still be in the solid Earth. It is fairly easy to show that most of this must be in the mantle rather than the crust (Allegre et al., 1996).

Combining Ar systematics with isotope systematics of other noble gases, there seems to be consensus on these points: (1) the Earth probably lost its early, primitive atmosphere, (2) the present atmosphere is a product of degassing of the Earth's interior, (3)

NuclideHalf-life (y) λ (y-1) ^{10}Be 1.5 x 10 ⁶ 0.462 x 10 ⁻⁶ ^{14}C 57300.1209 x 10 ⁻³ ^{26}Al 0.716 x 10 ⁶ 0.968 x 10 ⁻⁶ ^{36}Cl 0.308 x 10 ⁶ 2.25 x 10 ⁻⁶ ^{39}Ar 2690.257 x 10 ⁻²			
	Nuclide	Half-life (y)	λ (y ⁻¹)
	¹⁰ Be ¹⁴ C ²⁶ Al ³⁶ Cl ³⁹ Ar	$\begin{array}{c} 1.5 \times 10^6 \\ 5730 \\ 0.716 \times 10^6 \\ 0.308 \times 10^6 \\ 269 \end{array}$	$\begin{array}{c} 0.462 \ x \ 10^{-6} \\ 0.1209 \ x \ 10^{-3} \\ 0.968 \ x \ 10^{-6} \\ 2.25 \ x \ 10^{-6} \\ 0.257 \ x \ 10^{-2} \end{array}$

 Table 8.5.
 Some Cosmogenic Nuclides

 of Geologic Interest

much, perhaps half or more, of the mantle has been degassed rather extensively, and (4) most of the degassing occurred in Earth's history, probably shortly after it formed. Beyond these points, there is considerable disagreement.

Most of ⁴⁰K (90%) decays to ⁴⁰Ca. But this decay scheme has not yielded much geochemical or geochronological information thus far. The reason is the low ratio of ⁴⁰K to ⁴⁰Ca in most geological materials, which in turn means the radiogenic contribution to ⁴⁰Ca is small and ⁴⁰Ca/⁴²Ca variations are very small. Also since the isotopes are rather light, they are comparatively easily fractionated. Indeed, it has been shown that they can be fractionated by the chemical techniques used to separate Ca. The variation of ⁴⁰Ca/⁴²Ca is so small that in one good geochronological age determination, the initial ⁴⁰Ca/⁴²Ca ratio of the Pike's Peak batholith (age 1 Ga) was found to be indistinguishable from the initial ratio in meteorites. Work by Marshall and DePaolo (1987) showed that MORB have ε_{Ca} values (ε_{Ca} is defined in exactly the same sense and ε_{Nd} and ε_{Hf}) around 0. Thus although the source of mid-ocean ridge basalts has been depleted in K relative to Ca, this has not produced Ca isotope ratios significantly different from that of the bulk Earth. As a result of high K/Ca ratios, crustal rocks have slightly positive ε_{Ca} ratios. Interestingly, a number of island arc volcanics have ε_{Ca} in the range of 1-2, which is consistent with the hypothesis that their source contains a crustal component, a hypothesis based on the isotopic compositions of Pb, Sr and Nd in them.

8.5 Cosmogenic and Fossil Isotopes

The Earth is constantly bombarded by 'cosmic rays' (we shouldn't feel picked on, the entire cosmos is). These are atomic nuclei (mostly H) stripped of their electrons and traveling at relativistic velocities. Some originate in the Sun, but most originate in high-energy regions of the cosmos such as supernovae. For the most part their origin is not well understood. What is understood is that they have sufficient energy to shatter a nucleus when they collide with one. The nucleus breaks into several parts in a process called *spallation*.

Having mass and charge, cosmic rays don't have much penetrating power. Thus the intensity of cosmic radiation increases with increasing altitude in the atmosphere (indeed, this is how they were shown to be of cosmic origin). Most are stopped in the atmosphere, their interactions creating a cascade of lesser energy particles, or are slowed considerably. Even if they don't score a direct hit, they lose energy through electromagnetic interaction with matter (ionizing the atoms they pass by). Thus cosmic rays have their greatest effect in the atmosphere and somewhat less of an effect in the uppermost few centimeters or so of rock.

The nuclear effects of cosmic radiation are on the whole pretty trivial. Nitrogen and oxygen, being the principal constituents of the atmosphere are the most common targets, yet their is no change in the isotopic abundances of these elements. Cosmic radiation is most interesting because of the production of nuclides whose half-lives are so short they would not otherwise exist. The nuclides of greatest interest are listed in Table 8.5 along with their half-lives and their decay constants. These nuclides are created either directly through spallation (e.g., ¹⁰Be), or by nuclear reactions with particles produced by spallation, e.g., ¹⁴C: ¹⁴N (n,p) \rightarrow ¹⁴C. We have space only to briefly consider a few of the many applications of cosmogenic nuclides.

Geochemistry

8.5.1 ¹⁴C

The best known of the cosmogenic nuclide dating methods is ¹⁴C, of course. It is useful in dating archaeological, climatological, volcanological, seismological, paleontological, etc., 'events'. Present technology utilizes accelerator mass spectrometry rather than traditional β -counting and useful information on samples as old as 40,000 years can be produced. The principle of this method is quite simple. One assumes a constant production of ¹⁴C in the atmosphere. The atmosphere is well mixed and has a uniform ¹⁴C/¹²C ratio. The isotope ratio of carbon isolated from the atmosphere by plants will decrease with time according to:

$$\frac{{}^{14}C}{{}^{12}C} = \left(\frac{{}^{14}C}{{}^{12}C}\right)_0 e^{-\lambda_{14}t}$$
8.56

If the production rate is constant, the 'naught' value is the present day atmospheric value. In practice, ¹⁴C is generally reported in unites of activity of ¹⁴C per gram carbon – which is defined as the *specific ac-tivity*.

Research over the last 50 years has demonstrated that the production rate of ¹⁴C has certainly not been constant; and it can be inferred from this that the cosmic ray flux has not been constant. The primary cause of this variation is variation in solar activity and the amount of solar wind. Solar wind consists of charged particles and the electromagnetic field associated it with deflects some fraction of cosmic rays that would otherwise enter the solar system and strike the Earth. In addition to this, the ¹⁴C/¹²C ratio in the present atmosphere is somewhat lower than it has been in the recent past due to dilution with CO₂ released by extensive burning of fossil fuels.

To get around the problem of the non-constant production rate, ¹⁴C ages have been calibrated by comparison with absolute ages derived from tree rings. This calibration has now been extended back about 12,000 years. Thus, accurate ages ¹⁴C ages can be obtained despite production rate variations. Currently, an effort is under way to calibrate ¹⁴C ages back 50,000 years by comparing then with ^{230Th}/²³⁸U ages of corals. The discrepancy between ¹⁴C ages and other chronometers such as tree rings and ^{230Th}/²³⁸U ages provides a measure of the ¹⁴C production rate and cosmic ray flux. That provides an indirect measure past solar activity. Solar activity in turn has a direct effect on terrestrial climate, and hence must be constrained if past climatic variations are to be understood.

Other interesting applications include determining the 'age' of deep water in the oceans, which we will discuss in Chapter 15. By comparing $^{14}C/^{12}C$ ratios of benthic and planktonic foraminifera in a single sedimentary horizon, it is possible to determine bottom water ages in the past, and hence constrain 'paleocirculation' rates.

Other cosmogenic isotopes have been used to determine sedimentation rates, growth rate of Mn nodules where ages are too old for U-series disequilibrium techniques, and the ages of ice in Greenland and Antarctic ice cores.

8.5.2 ³⁶Cl in Hydrology

³⁶Cl has been applied to hydrological problems for some time. The general principle is that ³⁶Cl is produced at a constant rate in the atmosphere, carried to the surface of the Earth in rain and incorporated into hydrological systems. Cl is an excellent candidate element for such studies because it should remain in solution in most instances, and hence the only loss should be through radioactive decay. Imagine a simple system in which rainwater is incorporated into an aquifer at some unknown time in the past. In this case, if we can specify the number of ³⁶Cl atoms per liter in rain (and if we can assume this value is time-invariant), then we can determine the age of water in the aquifer by measuring the number of ³⁶Cl atoms per liter since:

$${}^{36}\mathrm{Cl} = {}^{36}\mathrm{Cl}^{\circ} \ e^{-\lambda t}$$

Determining the age of water in underground aquifers is an important problem because of the increasing demands placed in many parts of the world on limited water resources. A prudent policy for water resource management is to withdraw from a reservoir at a rate no greater than the recharge rate. Determination of recharge rate is thus prerequisite to wise management.

Dealing with just the number, or concentration, of ³⁶Cl atoms can have disadvantages, and can be misleading. Evaporation, for example, would increase the number of ³⁶Cl atoms. Thus the ³⁶Cl/Cl ratio (Cl has two stable isotopes: ³⁵Cl and ³⁷Cl) is often used. This also can have problems since chlorine can be leached from rocks. This chlorine will be nearly free of ³⁶Cl (some ³⁶Cl will be produced naturally by neu-

Table 8.6. ³⁶Cl measurements in the Dead Sea system

Site	³⁶ Cl/Cl	Cl	³⁶ Cl
	(10^{-15}) mg/l	$10^6 \text{ atoms}/1$	
		4 = 0	(a -
Mt. Hermon Snow	1580 ± 120	1.50	40 ± 5
Banias River	400 ± 60	11.9	80±15
Snir River	430±125	11.0	80±20
Dan River	725±140	10.5	129±25
Lake Kinneret	49±15	252	210±65
Jordan River	121±19	646	1,320±210
Dead Sea	17±2	2.30×10^{5}	6.6×10^{4}
Ashlag Spring (saline spr	ing) 4±2	2.6×10^{5}	
	0,		

tron capture), and hence this process will decrease the ³⁶Cl/Cl ratio. Further complications arise from the nuclear bomb-produced ³⁶Cl. Particularly large amounts were produced by nuclear bomb testing at sea, where bomb-produced neutrons were captured by ³⁵Cl in seawater.

In a somewhat different application, Paul et al. (1986) used ³⁶Cl to determine the accumulation time of dissolved salt in the Dead Sea. The Dead Sea is a particularly simply hydrologic system because it has no outlet. In such a simple system, we can describe the variation of the number of ³⁶Cl atoms with time as the rate of input less the rate of decay:

$$\frac{dN}{dt} = I - \lambda t \tag{8.58}$$

where I is the input rate (precipitation of chloride is assumed negligible). Integration of this equation yields:

$$N = \frac{I}{\lambda} (1 - e^{-\lambda t})$$
8.59

Paul et al. measured ³⁶Cl/Cl in Mt. Hermon snow, in various rivers in the Dead Sea system, and in saline springs in the Dead Sea basin. These results are summarized in Table 8.6. Using equation 8.58, they estimated an accumulation time of 20,000 years for the salt in the Dead Sea. The Dead Sea basin has been estimated to be 15,000 years old based on ¹⁴C. The difference suggests some of the Cl now in the Dead Sea was inherited from its less saline Pleistocene predecessor, Lake Lisan. The data in Table also illustrates how a combination of Cl and ³⁶Cl data can distinguish between addition of Cl from rock leaching and evaporation, both of which are processes that will increase the concentration of Cl. Evaporation should not significantly change the ³⁶Cl/Cl ratio, while addition of Cl derived from rock leaching should decrease this ratio. There is a general southward (downstream) increase in Cl concentration in the Jordan River—Dead Sea system. It is apparent from the data in Table 8.6 that this is due to both addition of rock-derived Cl and evaporation.

8.5.3 ¹⁰Be in Subduction Zone Studies

One of the usual uses of cosmogenic nuclides is the use of ¹⁰Be to trace sediment subduction (Tera et al., 1986). Since ¹⁰Be does not exist in the Earth's interior, its presence there could result only from subduction of sediment (which concentrates cosmogenic ¹⁰Be). ¹⁰Be has been identified in some island arc volcanics, but not in other volcanic rocks (at least not above the background level of 10⁶ atoms per gram, i.e. 1 atom in 10¹⁸). This is strong evidence that subducted sediment plays a role in island arc magma genesis, something suspected on the basis of other geochemical evidence. We'll examine this application of cosmogenic isotopes in more detail in Chapter 12.

Geochemistry

8.5.3 Cosmic-Ray Exposure Ages of Meteorites

The surfaces of meteorites in space are subject to a fairly high cosmic ray flux because there is no atmosphere to protect them. This leads to another interesting application of cosmogenic isotopes: 'cosmic ray exposure ages'. Here rare stable isotopes are used rather than radioactive ones because of the long times involved. For example, potassium is not present naturally in iron meteorites, but is produced by cosmic ray interactions. Knowing the production rate of ⁴¹K and its abundance, it is possible to calculate how long a meteorite has been exposed to cosmic rays. Two important results of such studies are worth mentioning: (1) exposure ages are much younger than formation ages and (2) meteorites that are compositionally and petrologically similar tend to have similar exposure ages (Figure 8.26). This means meteorites now colliding with the Earth have not existed as small bodies since the solar system formed. Instead, they are probably more or less continually (on a time scale of 10⁹ yr) produced by breakup of larger bodies through collisions. Also, similar meteorites probably come from the same parent body.

8.5.3 Fossil Nuclides

Finally, there is evidence short-lived isotopes existed in abundance when the solar system formed.

They must have been created shortly before, implying a nucleosynthetic process, such as a red giant star or a supernova event shortly before the formation of the Earth. Some of these isotopes (e.g., ²⁶Al) may have been sufficiently abundant that they could have been significant heat sources early in the history of the Earth and other planetary bodies. Evidence of their existence comes from non-uniform distribution of their daughter products. For example, if ²⁶Mg, the daughter of ²⁶Al, were found be more abundant in an Al-rich phase such as plagioclase or



Figure 8.26. Cosmic-Ray exposure age distribution in iron meteorites. Seventeen of 18 IIIAB irons fall in a cluster at 650±100 Ma (after Voshage, 1967).

spinel than in olivine, one might conclude the excess ²⁶Mg was produced by the decay of ²⁶Al. There is only one example of non-homogeneous distribution of the daughter of a fossil radionuclide in the Earth: ¹²⁹Xe, the daughter of ¹²⁹I. Excess ¹²⁹Xe has been identified in CO_2 well gas and in some oceanic basalts. The implication is first, the Earth formed before ¹²⁹I decayed away, and second, there was probably a major outgassing event very early in which most of the Xe went into the atmosphere. ¹²⁹I remained in the solid Earth, decaying to ¹²⁹Xe. As a result, the Xe that remained in the Earth was richer in ¹²⁹Xe.

References and Suggestions for Further Reading:

- Allégre, C. J., and J. M. Luck. 1980. Osmium isotopes as petrogenic and geologic tracers, *Earth Planet. Sci. Lett.*, 48: 148-154.
- Allegre, C. J., A. W. Hofmann and K. O'Nions, 1996. The argon constraints on mantle structure, *Geophysical Research Letters*, 23: 3555-3557.

Allègre, C. J., P. Sarda and T. Staudacher. 1993. Speculations about the cosmic origin of He and Ne in the interior of the Earth, *Earth Planet. Sci. Lett.*, 117, 229-233, 1993.

- Clark, W. B, M. A. Beg, and H. Craig. 1969. Excess ³He in the sea: evidence for terrestrial primordial helium, *Earth Planet. Sci. Lett.*, 6: 213-230, 1969.
- Creaser, R. A., D. A. Papanastassiou and G. J. Wasserburg. 1991. Negative thermal ion mass spectrometry of osmium, rhenium, and iridium. *Geochim. Cosmochim. Acta*. 55: 397-401.

DePaolo, D. J. 1988. Neodymium Isotope Geochemistry, an Introduction, Berlin: Springer-Verlag.

DePaolo, D. J. 1981. Neodymium isotopes in the Colorado Front Range and crust mantle evolution in the Proterozoic, *Nature*, 291, 193-196.

Dickin, A. P., Cerium isotope geochemistry of ocean island basalts, Nature, 326: 283-284, 1987.

Dickin, A. 1995. Radiogenic Isotope Geochemistry. Cambridge: Cambridge University Press.

Esser, B. K. and K. K. Turekian. 1993. The osmium isotopic composition of the continental crust. *Geochimica et Cosmochimica Acta*. 57: 3093-3104.

Geochemistry

- Farley, K. A. and R. J. Poreda. 1993., Mantle neon and atmospheric contamination, *Earth Planet. Sci. Lett.*, 114, 325-339.
- Faure, G. 1986. Principles of Isotope Geology, 2nd ed., New York: Wiley & Sons.
- Gast, P. W. 1960. Limitations on the composition of the upper mantle, J. Geophys. Res., 65: 1287-1297.
- Hauri, E. H. and S. R. Hart. 1993. Re-Os isotope systematics of HIMU and EMII oceanic island basalts from the south Pacific Ocean. *Earth Planet. Sci. Lett.* 114: 353-371.
- Hirt, B., W. Herr and W. Hoffmester. 1963. Age determinations by the rhenium-osmium method. in *Radioactive Dating*, ed. 35-44. Vienna: Internation. Atom Energy Agency.
- Honda, M., M. I., D. Patterson, A. Doulgeris and D. Clague, Possible solar noble-gas component in Hawaiian basalts, *Nature*, 349, 149-151, 1991.
- Huh, C.-A. and T.-L. Ku. 1984. Radiochemical observations on manganese nodules from three sedimentary environments in the north Pacific, *Geochim. Cosmochim. Acta*, 48. 951-964.
- Kurz, M. D. 1987. In situ production of terrestrial cosmogenic helium and some applications to gechronology, *Geochim. et Cosmochim. Acta*, 50: 2855–2862.
- Kurz, M. D., W. J. Jenkins, and S. R. Hart. 1982. Helium isotopic systematics of oceanic islands and mantle heterogeneity, *Nature*, 297: 43-47.
- Luck, J.-M., J.-L. Birck and C. J. Allegre. 1980. ¹⁸⁷Re-¹⁸⁷Os systematics in meteorites: Early chronology of the solar system and age of the Galaxy. *Nature*. 283: 256-259.
- Lupton J. E. and H. Craig. 1981. A major ³He source on the East Pacific Rise, *Science*, 214: 13-18.
- Lupton, J. E. 1983. Terrestrial inert gases: isotopic tracer studies and clues to primordial components in the mantle, *Ann. Rev. Earth. Planet. Sci.*, 11: 371-414.
- Lupton, J. G. and H. Craig. 1975 Excess ³He in oceanic basalts: evidence for terrestrial primordial helium, *Earth. Planet. Sci. Lett.*, 26: 133-.
- Marshall, B. D. and D. J. DePaolo, Initial Ca isotope variations in igneous rocks: implications for petrogenetic models and the composition of the lower crust, *EOS*, 68; 465, 1987.
- Martin, C. E. 1991. Osmium isotopic characteristics of mantle-derived rocks, *Geochim. Cosmochim. Acta*, 55, 1421-1434.
- Nyquist, L. E., D. D. Bogard, H. Wiesmann, B. M. Bansal., C.-Y. Shih and R. M. Morris. 1990. Age of a eucrite clase from the Bholghati howardite. *Geochim. Cosmochim. Acta*. 54: 2195-2206.
- Ozima, M. and K. Zahnle. 1993. Mantle degassing and atmospheric evolution: noble gas view, *Geochem*. *J.*, 27, 185-200.
- Patchett, P. J. 1983. Importance of the Lu-Hf isotopic system in studies of planetary chronology and chemical evolution, *Geochim. Cosmochim. Acta*, 47: 81-91.
- Patchett, P. J., W. M. White, H. Feldmann, S. Kielinczuk and A. W. Hofmann. 1984. Hafnium/rare earth element fractionation in the sedimentary system and crustal recycling into the Earth's mantle. *Earth. Planet. Sci. Lett.* 69: 365-378.
- Paul, M., A. Kaufman, M. Magaritz, D. Fink, W. Henning, et al. 1986. A new ³⁶Cl hydrological model and ³⁶Cl systematics in the Jordan River/Dead Sea system, *Nature*, 29, 511-515.
- Peucker-Ehrenbrink, B., G. Ravissa and A. W. Hoffmann. 1995. The marine ¹⁸⁷Os/¹⁸⁶Os record of the past 80 million years. *Earth Planet. Sci. Lett.* 130: 155-167.
- Poreda, R. J. and K. A. Farley, Rare gases in Samoan xenoliths. 1992. Earth Planet. Sci. Lett., 113, 129-144.
- Reisberg, L., A. Zindler, F. Marcantonio, W. M. White, D. Wyman, and B. Weaver. 1993. Os isotope systematics in ocean island basalts. *Earth Planet. Sci. Lett.* 120: 149-167.
- Sarda, P., T. Staudacher and C. J. Allègre. 1988. Neon isotopes in submarine basalts, *Earth Planet. Sci. Lett.*, 91, 73-88.
- Tera, F., L. Brown, J. Morris, I. S. Sacks, J. Klein, R. Klein and R. Middleton. 1986. sediment incorporation in island-arc magmas: inferences from ¹⁰Be, *Geochim. Cosmochim. Acta*, 50: 535-550.
- Volpe, A. M. and P. E. Hammond. 1991. ²³⁸U-²³⁰Th-²²⁶Ra disequilibria in young Mount St. Helens rocks: time constraint for magma formation and crystallization, *Earth Planet. Sci. Lett.*, 107, 475-486.
- Voshage, H. 1967. Bestrahlungsalter und Herkunft der Eisenmeteorite, Z. Naturforschung, 22a: 477-506.

W. M. White Geochemistry Chapter &: Radiogenic Isotope Geochemistry

- Walker, R. J., R. W. Carlson, S. B. Shirey and F. R. Boyd. 1989. Os, Sr, Nd, and Pb isotope systematics of southern African peridotite xenoliths: implications for the chemical evolution of the subcontinental mantle. *Geochim. Cosmochim. Acta*. 53: 1583-1595.
- White, W. M., P. J. Patchett and D. BenOthman. 1986. Hf isotope ratios of marine sediments and Mn nodules: evidence for a mantle source of Hf in seawater. *Earth Planet. Sci. Lett.* 79: 46-54.
- York, D. 1969. Least squares fitting of a straight line with correlated errors, *Earth Planet. Sci. Lett.*, 5, 320-324.

Zindler, A. and S. R. Hart. 1986. Helium: problematic primordial signals, *Earth. Planet Sci. Lett.*, 79: 1-8.

Problems

1. A few years ago, chemists at the University of Utah claimed to have succeeded with a 'cold' fusion experiment in which deuterium nuclei fuse to produce a ³He nucleus and a neutron, i.e.:

 $^{2}\text{H} + ^{2}\text{H} \rightarrow ^{3}\text{He} + n$

If the mass of ²H is 2.0141077 u, the mass of ³He is 3.01602929 u, the mass of a neutron is 1.00866489 u, and normal water has ${}^{2}H/{}^{1}H= 1.4 \times 10^{-4}$, what is the energy yield (in kJ) of this reaction per mole of normal water? (Don't forget that each reaction requires two deuterons).

2. What are the binding energies per nucleon of 147 Sm (mass = 146.914907 u) and 143 Nd (mass = 142.909823)?

3. Calculate the maximum β^- energy in the decay of ⁸⁷Rb to ⁸⁷Sr. The mass of ⁸⁷Rb is 86.9091836 u; the mass of ⁸⁷Sr is 86.9088902 u.

4. What is the decay constant (λ) of ¹⁵²Gd if its half-life is 1.1×10^{14} yr?

5. The following data were measured on whole rock gneiss samples from the Bighorn Mountains of Wyoming. Use linear regression to calculate the age and initial ⁸⁷Sr/⁸⁶Sr for this gneiss.

Sample	⁸⁷ Rb/ ⁸⁶ Sr	⁸⁷ Sr/ ⁸⁶ Sr
4173	0.1475	0.7073
3400	0.2231	0.7106
7112	0.8096	0.7344
3432	1.1084	0.7456
3422	1.4995	0.7607
83	1.8825	0.7793

6. The following were measured on a komatiite flow in Canada. Use simple linear regression to calculate slope. Plot the data on isochron diagrams.

	$^{147}Sm/^{144}Nd$	$^{143}Nd/^{144}Nd$
M654	0.2427	0.513586
M656	0.2402	0.513548
M663	0.2567	0.513853
M657	0.2381	0.513511
AX14	0.2250	0.513280
AX25	0.2189	0.513174
M666	0.2563	0.513842
M668	0.2380	0.513522

a. Calculate the Sm-Nd age and the error on the age.

b. Calculate the initial ε_{Nd} (i.e., the ε_{Nd} at the age you calculate in a.) and the error on the initial ε_{Nd} .
W. M. White Chapter &: Radiogenic Isotope Geochemistry

Geochemistry

7. A sample of granite has $^{143}Nd/^{144}Nd$ and $^{147}Sm/^{144}Nd$ of 0.51215 and 0.1342, respectively. The present chondritic $^{143}Nd/^{144}Nd$ and $^{147}Sm/^{144}Nd$ are 0.512638 and 0.1967, respectively. The decay constant of ^{147}Sm is 6.54 \times 10⁻¹² Ga⁻¹. Calculate the τ_{CHUR} , i.e., crustal residence time relative to a chondritic mantle, for this granite.

8. Imagine that an initially uniform silicate Earth underwent melting at some time in the past to form continental crust (melt) and mantle (melting residue). Calculate the present day Sr and Nd isotopic composition of 1%, 2%, and 5% partial melts and respective melting residues assuming the bulk partition coefficients given in Example 8.4. Assume that the present day 87 Rb/ 86 Sr, 87 Sr/ 86 Sr, 147 Sm/ 144 Nd, 143 Nd/ 144 Nd ratios of the bulk silicate Earth are 0.085, 0.705, and 0.1967, and 0.512638 respectively. Perform the calculation assuming the melting occurred at 4.0 Gam 3.0 Ga, and 2.0 Ga. Plot your results on a Sr-Nd isotope diagram (i.e., 143 Nd/ 144 Nd vs. 87 Sr/ 86 Sr).

9. Given the data on a series of whole rocks below, use linear regression to calculate:

(a.) the age of the rocks and the error on the age

(b.) their initial 143 Nd/ 144 Nd, and the error on the initial ratio

(c.) From b, calculate the initial ϵ_{Nd} , that is, ϵ_{Nd} at the time calculated in a. Take the present-day chondritic ¹⁴³Nd/¹⁴⁴Nd to be 0.512638 and the (present-day) chondritic ¹⁴⁷Sm/¹⁴⁴Nd to be 0.1967 (you need to calculate the chondritic value at the time the rock formed to calculate initial ϵ_{Nd}).

(d.) Calculate the depleted mantle model age τ_{DM} . Assume that the present ¹⁴³Nd/¹⁴⁴Nd of the depleted mantle is 0.51310, and that the depleted mantle has evolved from the chondritic initial with a constant ¹⁴⁷Sm/¹⁴⁴Nd since 4.55 Ga. How does it compare to the age you calculated in (a.)?

sample	147 Sm $/^{144}$ Nd	¹⁴³ Nd/ ¹⁴⁴ Nd	
whole rock	0.1886	0.512360	
garnet	0.6419	0.513401	
clinopyroxene	0.1146	0.512245	

Linear regression functions are available on some scientific calculators and in statistical packages for microcomputers, and in Microsoft Excel. Alternatively, use the equations in Appendix 3.

10. The following were measured on a komatiite flow in Canada. Use simple linear regression to calculate slopes. Plot the data on isochron diagrams.

	²⁰⁶ Pb/ ²⁰⁴ Pb	²⁰⁷ Pb/ ²⁰⁴ Pb	²⁰⁸ Pb/ ²⁰⁴ Pb	
M665	15.718	14.920	35.504	
M654	15.970	14.976	35.920	
M656	22.563	16.213	41.225	
M663	16.329	15.132	35.569	
M657	29.995	17.565	48.690	
AX14	32.477	17.730	49.996	
AX25	15.869	14.963	35.465	
M667	14.219	17.717	33.786	
M666	16.770	15.110	35.848	
M668	16.351	15.047	36.060	
M658	20.122	15.700	39.390	

a. Calculate the Pb-Pb age and the error on the age.

b. Calculate the Th/U ratio of the samples.

Geochemistry

W. M. White Chapter &: Radiogenic Isotope Geochemistry

11. Calculate the 207 Pb/ 204 Pb— 206 Pb/ 204 Pb age for the rocks below. (Hint, first calculate the slope using linear regression, then use Excel's Solver to calculate the age).

sample	²⁰⁶ Pb/ ²⁰⁴ Pb	²⁰⁷ Pb/ ²⁰⁴ Pb	
NPA5	15.968	14.823	
NPA12	17.110	15.043	
NPA15	17.334	15.090	
NPA15HF	17.455	15.104	

12. a.) Calculate the isotopic evolution ($^{207}Pb/^{204}Pb$ and $^{206}Pb/^{204}Pb$ only) of Pb in a reservoir having a $^{238}U/^{204}Pb$ of 8. Do the calculation at 0.5 Ga intervals from 4.5 Ga to present. Assume the reservoir started with Canyon Diablo initial Pb isotopic composition. Plot your results on a $^{207}Pb/^{204}Pb$ vs $^{206}Pb/^{204}Pb$ graph.

b.) Calculate the isotopic evolution $({}^{207}\text{Pb}/{}^{204}\text{Pb}$ and ${}^{206}\text{Pb}/{}^{204}\text{Pb}$ only) of Pb in a reservoir having a ${}^{238}\text{U}/{}^{204}\text{Pb}$ of 7 from 4.5 to 2.5 Ga and ${}^{238}\text{U}/{}^{204}\text{Pb}$ of 9 from 2.5 Ga to present. Do the calculation at 0.5 Ga intervals from 4.5 Ga to present. Assume the reservoir started with Canyon Diablo initial Pb isotopic composition. Plot your results on a ${}^{207}\text{Pb}/{}^{204}\text{Pb}$ vs ${}^{206}\text{Pb}/{}^{204}\text{Pb}$ graph.

HINT: the equation:
$${}^{206}Pb / {}^{204}Pb = {}^{206}Pb / {}^{204}Pb_0 + \mu(e^{\lambda_{238}t} - 1)$$

is valid for calculating the growth of ²⁰⁶Pb/²⁰⁴Pb only between the present and the initial time; i.e., the time when ²⁰⁶Pb/²⁰⁴Pb = (²⁰⁶Pb/²⁰⁴Pb)₀ (because the ²³⁸U/²⁰⁴Pb ratio used is the present ratio). The growth of ²⁰⁶Pb/²⁰⁴Pb between two other times, t_1 and t_2 , where t_1 is older than t_2 , may be calculated by calculating the growth of ²⁰⁶Pb/²⁰⁴Pb between t_2 and the present, and between t_1 and the present and subtracting the latter from the former. If μ_1 is the value of μ between t_1 and t_2 , and μ_2 is the value of μ between t_2 and the present, the relevant equation is then:

$${}^{206}Pb / {}^{204}Pb = ({}^{206}Pb / {}^{204}Pb)_{t_1} + \mu_1(e^{\lambda_{238}t_1} - 1) + \mu_2(e^{\lambda_{238}t_2} - 1)$$

if $\mu_1 = \mu_2$ *then the equation simplifies to:*

²⁰⁶*Pb* / ²⁰⁴*Pb* = (²⁰⁶*Pb* / ²⁰⁴*Pb*)_{t₁} +
$$\mu(e^{\lambda_{238}t_1} - e^{\lambda_{238}t_2})$$

13. A basalt from Réunion has a (230 Th/ 232 Th) ratio of 0.93 and a (238 U/ 232 Th) ratio of 0.75. Assuming the age of the basalt is 0:

a.) what is the $(^{238}U/^{232}Th)$ ratio of the source of the basalt?

b.) what is the 232 Th/ 238 U *atomic* ratio of the source of the basalt?

c.) assuming bulk distribution coefficients of 0.01 for U and 0.005 for Th and equilibrium melting, what is the percent melting involved in generating this basalt?

(remember the parentheses denote *activity* ratios).

14. Given the following data on the Cheire de Mazaye Flow in the Massif Central, France, calculate the age of the flow and the initial 232 Th/ 238 U (atomic). Use simple linear regression in obtaining your solution.

	(²³⁸ U/ ²³² Th)	(²³⁰ Th/ ²³² Th)
whole rock	0.744	0.780 ± 0.012
magnetite M1 (80-23 μ)	0.970	$0.864{\pm}0.017$
magnetite M2 (23-7 μ)	1.142	$0.904{\pm}0.017$
clinopyroxene	0.750	$0.791{\pm}0.019$
plagioclase	0.685	$0.783 {\pm} 0.018$

CHAPTER 9: STABLE ISOTOPES

Chapter 9: Stable Isotope Geochemistry

9.1 INTRODUCTION

S table isotope geochemistry is concerned with variations of the isotopic compositions of elements arising from physicochemical processes rather than nuclear processes. Fractionation^{*} of the isotopes of an element might at first seem to be an oxymoron. After all, in the last chapter we saw that the value of radiogenic isotopes was that the various isotopes of an element had identical chemical properties and therefore that isotope ratios such as ⁸⁷Sr/⁸⁶Sr are not changed measurably by chemical processes. In this chapter we will find that this is not quite true, and that the very small differences in the chemical behavior of different isotopes of an element can provide a very large amount of useful information about chemical (both geochemical and biochemical) processes.

The origins of stable isotope geochemistry are closely tied to the development of modern physics in the first half of the 20th century. The discovery of the neutron in 1932 by H. Urey and the demonstration of variations in the isotopic composition of light elements by A. Nier in the 1930's and 1940's were the precursors to the development of stable isotope geochemistry. The real history of stable isotope geochemistry begins in 1947 with the Harold Urey's publication of a paper entitled *"The Thermodynamic Properties of Isotopic Substances"*. Urey not only showed why, on theoretical grounds, isotope fractionations could be expected, but also suggested that these fractionations could provide useful geological information. Urey then set up a laboratory to determine the isotopic compositions of natural substances and experimentally determine the temperature dependence of isotopic fractionations, in the process establishing the field of stable isotope geochemistry.

What has been learned in the fifty years since that paper was published would undoubtedly astonish even Urey. Stable isotope geochemistry, like radiogenic isotope geochemistry, has become an essential part of not only geochemistry, but the earth sciences as a whole. In this chapter, we will attempt to gain an understanding of the principles underlying stable isotope geochemistry and then briefly survey its many-fold applications in the earth sciences. In doing so, we add the final tool to our geochemical toolbox.

9.1.1 Scope of Stable Isotope Geochemistry

The principal elements of interest in stable isotope geochemistry are H, Li, B, C, N, O, Si, S, and Cl. Of these, O, H, C and S are of the greatest interest. Most of these elements have several common characteristics:

(1) They have low atomic mass.

- (2) The relative mass difference between their isotopes is large.
- (3) They form bonds with a high degree of covalent character.

(4) The elements exist in more than one oxidation state (C, N, and S), form a wide variety of compounds (O), or are important constituents of naturally occurring solids and fluids.

(5) The abundance of the rare isotope is sufficiently high (generally at least tenths of a percent) to facilitate analysis.

It was once thought that elements not meeting these criteria would not show measurable variation in isotopic composition. However, as new techniques offering greater sensitivity and higher precision have become available (particularly use of the MC-ICP-MS), geochemists have begun to explore isotopic variations of metals such as Mg, Ca, Ti, Cr, Fe, Zn, Cu, Ge, Se, Mo, and Tl. The isotopic variations observed in these metals have generally been quite small. Nevertheless, some geologically useful information has been obtained from isotopic study of these metals and exploration of their isotope geochemistry continues. We do not have space to consider them here, but their application is reviewed in a recent book (Johnson et al., 2004).

The elements of interest in radiogenic isotope geochemistry are heavy (Sr, Nd, Hf, Os, Pb), most form dominantly ionic bonds, generally exist in only one oxidation state, and there is only small rela-

Fractionation refers to the change in an isotope ratio that arises as a result of some chemical or physical process.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

TABLE 9.1. Isotope Ratios of Stable Isotopes					
Element	Notation	Ratio	Standard	Absolute Ratio	
Hydrogen	δD	$D/H (^{2}H/^{1})$	H) SMOW	1.557×10^{4}	
Lithium	δ ⁷ Li	⁷ Li/ ⁶ Li	NIST 8545 (L-SVE	EC) 12.285	
Boron	$\delta^{11}B$	$^{11}B/^{10}B$	NIST 951	4.044	
Carbon	$\delta^{13}C$	$^{13}C/^{12}C$	PDB	1.122×10^{-2}	
Nitrogen	$\delta^{15}N$	$^{15}N/^{14}N$	atmosphere	3.613×10^{-3}	
Oxygen	$\delta^{18}O$	$^{18}O/^{16}O$	SMOŴ, PDB	2.0052×10^{-3}	
	$\delta^{17}O$	$^{17}O/^{16}O$	SMOW	3.76×10^{-4}	
Sulfur	$\delta^{34}S$	$^{34}S/^{32}S$	CDT	4.43×10^{-2}	

tive mass differences between the isotopes of interest. Thus isotopic fractionation of these elements is quite small and can generally be ignored. Furthermore, any natural fractionation is corrected for in the process of correcting much larger fractionations that typically occur during analysis (with the exception

of Pb). Thus is it that one group of isotope geochemists make their living by measuring isotope fractionations while the other group makes their living by ignoring them!

Stable isotope geochemistry has been be applied to a great variety of problems, and we will see a number of examples in this chapter. One of the most common is geothermometry. Another is process identification. For instance, plants that produce ' C_4 ' hydrocarbon chains (that is, hydrocarbon chains 4 carbons long) as their primary photosynthetic product fractionate carbon differently than to plants that produce ' C_3 ' chains. This fractionation is retained up the food chain. This allows us, for example, to draw some inferences about the diet of fossil mammals from the stable isotope ratios in their bones. Sometimes stable isotope ratios are used as 'tracers' much as radiogenic isotopes are. So, for example, we can use oxygen isotope ratios in igneous rocks to determine whether they have assimilated crustal material, as crust generally has different O isotope ratios than does the mantle.

9.1.2 Some Definitions

9.1.2.1 The δ Notation

As we shall see, variations in stable isotope ratios are typically in the parts per thousand to parts per hundred range and are most conveniently and commonly reported as *permil* deviations, δ , from some standard. For example, O isotope ratios are often reported as permil deviations from SMOW (standard mean ocean water) as:

$$\delta^{18}O = \left[\frac{({}^{18}O / {}^{16}O)_{sam} - ({}^{18}O / {}^{16}O)_{SMOW}}{({}^{18}O / {}^{16}O)_{SMOW}}\right] \times 10^3$$
9.1

Unfortunately, a dual standard developed for reporting O isotopes. While O isotopes in most substances are reported relative to SMOW, the oxygen isotope composition of carbonates is reported relative to the Pee Dee Belemite (PDB) carbonate standard[‡]. This value is related to SMOW by:

$$\delta^{18}O_{PDB} = 1.03086\delta^{18}O_{SMOW} + 30.86$$
 9.2

A similar δ notation is used to report other stable isotope ratios. Hydrogen isotope ratios, δD , are also reported relative to SMOW, carbon isotope ratios relative to PDB, nitrogen isotope ratios relative to atmospheric nitrogen (denoted ATM), and sulfur isotope ratios relative to troilite in the Canyon Diablo iron meteorite (denoted CDT). Table 9.1 lists the isotopic composition of these standards.

9.1.2.2 The Fractionation Factor

The *fractionation factor*, α , is the ratio of isotope ratios in two phases:

$$\alpha_{A-B} \equiv \frac{R_A}{R_B} \qquad 9.3$$

⁺ There is, however, a good historical reason for this: analytical difficulties in the measurement of carbonate oxygen prior to 1965 required it be measured against a carbonate standard.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

The fractionation of isotopes between two phases is also often reported as $\Delta_{A-B} = \delta_A - \delta_B$. The relationship between Δ and α is:

$$\Delta \approx (\alpha - 1)10^3$$
 or $\Delta \approx 10^3 \ln \alpha$ 9.4[†]

As we will see, at equilibrium, α may be related to the equilibrium constant of thermodynamics by

$$\alpha_{\text{A-B}} = (\text{K/K}_{\text{oo}})^{1/n} \qquad 9.5$$

where *n* is the number of atoms exchanged, K_{∞} is the equilibrium constant at infinite temperature, and K is the equilibrium constant written in the usual way (except that concentrations are used rather than activities because the ratios of the activity coefficients are equal to 1, i.e., there are no isotopic effects on the activity coefficient).

9.2 Theoretical Considerations

Isotope fractionation can originate from both *kinetic* effects and *equilibrium* effects. The former may be intuitively expected, but the latter may at first seem somewhat surprising. After all, we have been taught that the chemical properties of an element are dictated by its electronic structure, and that the nucleus plays no real role in chemical interactions. In the following sections, we will see that quantum mechanics predicts that the mass of an atom affect its vibrational motion therefore the strength of chemical bonds. It also affects rotational and translational motions. From an understanding of these effects of atomic mass, it is possible to predict the small differences in the chemical properties of isotopes quite accurately.

9.2.1 Equilibrium Isotope Fractionations

Most isotope fractionations arise from *equilibrium* effects. *Equilibrium fractionations arise from translational, rotational and vibrational motions of molecules in gases and liquids and atoms in crystals because the energies associated with these motions are mass dependent.* Systems tend to adjust themselves so as to minimize energy. Thus isotopes will be distributed so as to minimize the vibrational, rotational, and translational energy of a system. Of the three types of energies, vibrational energy makes by far the most important contribution to isotopic fractionation. Vibrational motion is the only mode of motion available to atoms in a solid. These effects are, as one might expect, small. For example, the equilibrium constant for the reaction

$$\frac{1}{2}C^{16}O_2 + H_2^{18}O \rightleftharpoons \frac{1}{2}C^{18}O_2 + H_2^{16}O \qquad 9.6$$

is only about 1.04 at 25°C and the ΔG of the reaction, given by $-RT \ln K$, is only -100 J/mol (you'll recall most ΔG 's for reactions are in typically in the kJ/mol range).

9.2.1.1 The Quantum Mechanical Origin of Isotopic Fractionations

It is fairly easy to understand, at a qualitative level at least, how some isotope fractionations can arise from vibrational motion. Consider the two hydrogen atoms of a hydrogen molecule: they do not remain at a fixed distance from one and other, rather they continually oscillate toward and away from each other, even at absolute zero. The frequency of this oscillation is quantized, that is, only discrete

⁺ To derive this, we rearrange equation 8.1 to obtain: $R_A = (\delta_A + 10^3)R_{STD}/10^3$ Note that Δ is sometimes defined as $\Delta \equiv 10^3 \ln \alpha$, in which case $\Delta_{AB} \approx \delta_A - \delta_B$.

So that α may be expressed as: $\alpha = \frac{(\delta_{A} + 10^{3})R_{STD} / 10^{3}}{(\delta_{B} + 10^{3})R_{STD} / 10^{3}} \approx \frac{(\delta_{A} + 10^{3})}{(\delta_{B} + 10^{3})}$

Subtracting 1 from each side and rearranging, we obtain:

$$\alpha - 1 = \alpha - 1 = \frac{(\delta_{A} - \delta_{B})}{(\delta_{B} + 10^{3})} \approx \frac{(\delta_{A} - \delta_{B})}{10^{3}} \approx \Delta \times 10^{-3}$$

since δ is small. The second equation in 9.4 results from the approximation that for $x \approx 1$, ln $x \approx x$ -1.

CHAPTER 9: STABLE ISOTOPES



Interatomic Distance

Figure 9.1. Energy-level diagram for the hydrogen molecule. Fundamental vibration frequencies are 4405 cm⁻¹ for H₂, 3817 cm⁻¹ for HD, and 3119 cm⁻¹ for D₂. The zero-point energy of H₂ is greater than that of HD which is greater than that of D₂. Arrows show the energy, in kJ/mol, required to dissociate the 3 species. After O'Niel (1986).

frequency values are possible. Figure 9.1 is a schematic diagram of energy as a function of interatomic distance in the hydrogen molecule. As the atoms vibrate back and forth, their potential energy varies as shown by the curved line. The zero point energy (ZPE) is the energy level at which the molecule will vibrate in its ground state, which is the state in which the molecule will be in at low temperature. The zero point energy is always some finite amount above the minimum potential energy of an analogous harmonic oscillator.

The potential energy curves for various isotopic combinations of an element are identical, but as the figure shows, the zero point vibrational energies differ, as do vibration energies at higher quantum levels, being lower for the heavier isotopes. Vibration energies and frequencies are directly related to bond strength. Because of this, the

energy required for the molecule to dissociate will differ for different isotopic combinations. For example, 440.6 kJ/mole is necessary to dissociate a D_2 (²H₂) molecule, but only 431.8 kJ/mole is required to dissociate the ¹H₂ molecule. Thus the bond formed by the two deuterium atoms is 9.8 kJ/mole stronger than the H–H bond. The differences in bond strength can also lead to kinetic fractionations, since molecules that dissociate easier will react faster. We will discuss kinetic fractionations a bit later.

9.2.1.2 Predicting Isotopic Fractionations from Statistical Mechanics

Now let's attempt to understand the origin of isotopic fractionations on a more quantitative level. We have already been introduced to several forms of the *Boltzmann distribution law* (e.g., equation 2.84), which describes the distribution of energy states. It states that the probability of a molecule having internal energy E_i is:

$$P_{i} = \frac{g_{i}e^{-E_{i}/kT}}{\sum_{n}g_{n}e^{-E_{n}/kT}}$$
9.7

where g is a statistical weight factor[‡], k is Boltzmann's constant, and the sum in the denominator is taken over all possible states. The average energy (per molecule) is:

^{\pm} This factor comes into play where more than one state corresponds to an energy level E_i (the states are said to be 'degenerate'). In that case g_i is equal to the number of states having energy level E_i.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

$$\overline{E} = \sum_{i} E_{i} P_{i} = \frac{\sum_{i} g_{i} E_{i} e^{-E_{i}/kT}}{\sum_{i} g_{i} e^{-E_{i}/kT}}$$

As we saw in Chapter 2, the *partition function*, *Q*, is the denominator of this equation:

$$Q = \sum_{n} g_n e^{-E_n/kT}$$
 9.9

9.8

The partition function is related to thermodynamic quantities U and S (equations 2.90 and 2.91). Since there is no volume change associated with isotope exchange reactions, we can use the relationship:

$$\Delta G = \Delta U - T \Delta S \qquad 9.10$$

and equations 2.119 and 2.123 to derive:

$$\Delta G_r = -RT \ln \prod_i Q_i^{\nu_i} \qquad 9.11$$

and comparing to equation 3.86, that the equilibrium constant is related to the partition function as:

$$\mathbf{K} = \prod_{i} Q_{i}^{\mathbf{v}_{i}} \qquad 9.12$$

for isotope exchange reactions. In the exchange reaction above (9.6) this is simply:

$$\mathbf{K} = \frac{Q_{C^{18}O_2}^{0.5} Q_{H_2^{16}O}}{Q_{C^{16}O_2}^{0.5} Q_{H_2^{18}O}}$$
9.13

The usefulness of the partition function is that it can be calculated

from quantum mechanics, and from it we can calculate equilibrium fractionations of isotopes.

There are three modes of motion available to gaseous molecules: vibrational, rotational, and translational (Figure 9.2). The partition function can be written as the product of the translational, rotational, vibrational, and electronic partition functions:

$$Q_{total} = Q_{vib}Q_{trans}Q_{rot}Q_{elec}$$
9.14

The electronic configurations and energies of atoms are unaffected by the isotopic differences, so the last term can be ignored in the present context.

The vibrational motion is the most important contributor to isotopic fractionations, and it is the only mode of motion available to atoms in solids. So we begin by calculating the vibrational partition function. We can approximate the vibration of atoms in a simple diatomic molecule such as CO or O_2 by that of a harmonic oscillator. The energy of a 'quantum' oscillator is:

$$E_{vib} = (n + 1/2)hv_0$$
 9.15

where v_0 is the ground state vibrational frequency, *h* is Plank's constant and *n* is the vibrational quantum number. The partition function for vibrational motion of a diatomic molecule is given by:

$$Q_{vib} = \frac{e^{-hv/2kT}}{1 - e^{-hv/2kT}} \qquad 9.16^{\ddagger}$$



Figure 9.2. The three modes of motion, illustrated for a diatomic molecule. Rotations can occur about both the y and x axes; only the rotation about the y axis is illustrated. Since radial symmetry exists about the z axis, rotations about that axis are not possible according to quantum mechanics. Three modes of translational motion are possible: in the x, y, and zdirections. Possible vibrational and rotational modes of motion of polyatomic molecules are more complex.

[‡] Polyatomic molecules have many modes of vibrational motion available to them. The partition function is calculated by summing the energy over all available modes of motion; since energy occurs in the exponent, the result is a product:

Geochemistry

CHAPTER 9: STABLE ISOTOPES

Vibrations of atoms in molecules and crystals approximates that of harmonic oscillators. For an ideal harmonic oscillator, the relation between frequency and reduced mass is:

$$v = \frac{1}{2}\sqrt{\frac{k}{\mu}}$$
9.17

where *k* is the forcing constant, which depends on the electronic configuration of the molecule, but not the isotope involved, and μ is reduced mass:

$$\mu = \frac{1}{1/m_1 + 1/m_2} = \frac{m_1 m_2}{m_1 + m_2}$$
9.18

Hence we see that the vibrational frequency, and therefore also vibrational energy and partition function, depends on the mass of the atoms involved in the bond.

Rotational motion of molecules is also quantized. We can approximate a diatomic molecule as a dumbbell rotating about the center of mass. The rotational energy for a quantum rigid rotator is:

$$E_{rot} = \frac{j(j+1)h^2}{8\pi^2 I}$$
 9.19

where *j* is the rotational quantum number and *I* is the moment of inertia, $I = \mu r^2$, where *r* is the interatomic distance. The statistical weight factor, *g*, in this case is equal to (2j + 1) since there 2 axes about which rotations are possible. For example, if *j* = 1, then there are *j*(*j* + 1) = 2 quanta available to the molecule, and (2j + 1) ways of distributing these two quanta: all to the x-axis, all to the y-axis or one to each. Thus:

$$Q_{rot} = \sum (2j+1)e^{-j(j+1)h^2/8\pi^2 lkT}$$
9.20

Because the rotational energy spacings are small, equation 9.20 can be integrated to yield:

$$Q_{rot} = \frac{8\pi^2 I \mathrm{k}T}{\sigma h^2} \qquad 9.21^*$$

where σ is a symmetry factor whose value is 1 for heteronuclear molecules such as ¹⁶O-¹⁸O and 2 for homonuclear molecules such as ¹⁶O-¹⁶O. This arises because in homonuclear molecules, the quanta must be equally distributed between the rotational axes, i.e., the j's must be all even or all odd. This restriction does not apply to hetergeneous molecules, hence the symmetry factor.

Finally, translational energy associated with each of the three possible translational motions (x, y, and z) is given by the solution to the Schrödinger equation for a particle in a box:

$$E_{trans} = \frac{n^2 h^2}{8 m a^2}$$
 9.22

where n is the translational quantum number, m is mass, and a is the dimension of the box. This expression can be inserted into equation 9.8. Above about 2 K, spacings between translational energies levels are small, so equ. 9.8 can also be integrated. Recalling that there are 3 degrees of freedom, the result is:

$$Q_{vib} = \prod_{i} \frac{e^{-hv_{i}/2kT}}{1 - e^{-hv_{i}/2kT}}$$
 9.16a

^{*} Equation 9.21 also holds for linear polyatomic molecules such as CO_2 . The symmetry factor is 1 if it has plane symmetry, and 2 if it does not. For non-linear polyatomic molecules, 9.21 is replaced by:

$$Q_{rot} = \frac{8\pi^2 (8\pi^3 ABC)^{1/2} (kT)^{3/2}}{\sigma h^3}$$
 9.21a

where *A*, *B*, and *C* are the three principle moments of inertia.

November 30, 2005

Geochemistry

CHAPTER 9: STABLE SOTOPES

$$Q_{trans} = \frac{(2\pi m kT)^{3/2}}{h^3} V$$
 9.23

where V is the volume of the box (a^3) . Thus the full partition function is:

 $l_{\rm m}$ /21/T

$$Q_{total} = Q_{vib}Q_{trans}Q_{rot} = \frac{e^{-hv/2kT}}{1 - e^{-hv/2kT}} \frac{8\pi^2 I kT}{\sigma h^2} \frac{(2\pi m kT)^{3/2}}{h^3} V \qquad 9.24$$

It is the ratio of partition functions that occurs in the equilibrium constant expression, so that many of the terms in 9.24 eventually cancel. Thus the partition function ratio for 2 different isotopic species of the same diatomic molecule, A and B, reduces to:

$$\frac{Q_{A}}{Q_{B}} = \frac{\frac{e^{-hv_{A}/2kT}}{1 - e^{-hv_{B}/2kT}} \frac{I_{A}}{\sigma_{A}} m_{A}^{3/2}}{\frac{e^{-hv_{B}/2kT}}{1 - e^{-hv_{B}/2kT}} \frac{I_{B}}{\sigma_{B}} m_{B}^{3/2}} = \frac{e^{-hv_{A}/2kT} (1 - e^{-hv_{B}/2kT}) I_{A} \sigma_{B} m_{A}^{3/2}}{e^{-hv_{B}/2kT} (1 - e^{-hv_{A}/2kT}) I_{B} \sigma_{A} m_{B}^{3/2}}$$
9.25

Since bond lengths are essentially independent of the isotopic composition, this further reduces to:

$$\frac{Q_A}{Q_B} = \frac{e^{-hv_A/2kT}(1 - e^{-hv_B/2kT})\mu_A\sigma_B m_A^{3/2}}{e^{-hv_B/2kT}(1 - e^{-hv_A/2kT})\mu_B\sigma_A m_B^{3/2}} = \frac{e^{-h(v_A - v_B)/2kT}(1 - e^{-hv_B/2kT})\mu_A\sigma_B m_A^{3/2}}{(1 - e^{-hv_A/2kT})\mu_B\sigma_A m_B^{3/2}}$$
9.26

Notice that all the temperature terms cancel except for those in the vibrational contribution. Thus vibrational motion alone is responsible for the temperature dependency of isotopic fractionations.

To calculate the fractionation factor α from the equilibrium constant, we need to calculate K_{∞} . For a reaction such as:

$$aA_1 + bB_2 \rightleftharpoons aA_2 + bB_1$$

where A_1 and A_2 are two molecules of the same substance differing only in their isotopic composition, and a and b are the stoichiometric coefficients, the equilibrium constant is:

$$K_{\infty} = \frac{(\sigma_{A_2} / \sigma_{A_1})^a}{(\sigma_{B_1} / \sigma_{B_2})^b}$$
 9.27

Thus for a reaction where only a single isotope is exchanged, K_{∞} is simply the ratio of the symmetry factors.

9.2.1.3 Temperature Dependence of the Fractionation Factor

As we noted above, the temperature dependence of the fractionation factor depends only on the vibrational contribution. At temperatures where T<< hv/k, the 1 - $e^{-hv/kT}$ terms in 9.16 and 9.26 tend to 1 and can therefore be neglected, so the vibrational partition function becomes:

$$Q_{\rm wib} \cong e^{-h\nu/2kT} \qquad 9.28$$

In a further simplification, since Δv is small, we can use the approximation $e^{x} \approx 1 + x$ (valid for x<<1), so that the ratio of vibrational energy partition functions becomes

$$Q_{vib}^{A} / Q_{vib}^{B} \cong 1 - h\Delta v / 2kT$$

Since the translational and rotational contributions are temperature independent, we expect a relationship of the form:

$$\alpha \cong A + \frac{B}{T} \tag{9.29}$$

In other words, α should vary inversely with temperature at low temperature.

At higher temperature, the 1 - $\exp(-hv/kT)$ term differs significantly from 1. Furthermore, at higher vibrational frequencies, the harmonic oscillator approximation breaks down (as suggested in Figure

Geochemistry

CHAPTER 9: STABLE ISOTOPES

9.1), as do several of the other simplifying assumptions we have made, so that the relation between the fractionation factor and temperature approximates:

$$\ln \alpha \propto \frac{1}{T^2} \qquad 9.30$$

Since α is generally small, $\ln \alpha \approx 1 + \alpha$, so that $\alpha \propto 1 + 1/T^2$. At infinite temperature, the fractionation is unity, since $\ln \alpha = 0$. This illustrated in Figure 9.3 for distribution of ¹⁸O and ¹⁶O between CO₂ and H₂O. The $\alpha \propto 1/T$ relationship breaks down around 200°C; above that temperature the relationship follows $\alpha \propto 1/T^2$.

It must be emphasized that the simple calculations performed in Example 9.1 are applicable only to a gas whose vibrations can be approximated by a simple harmonic oscillator. Real gases often show fractionations that are complex functions of temperature, with minima, maxima, inflections, and crossovers. Vibrational modes of silicates, on the other hand, are relatively well behaved.

9.2.1.4 Composition and Pressure Dependence

The nature of the chemical bonds in the phases involved is most important in determining isotopic fractionations. A general rule of thumb is that the heavy isotope goes into the phase in which it is most strongly bound. Bonds to ions with a high ionic potential and low atomic mass are associated with high vibrational frequencies and have a tendency to incorporate the heavy isotope preferentially. For example, quartz, SiO₂ is typically the most ¹⁸O rich mineral and magnetite the least dominantly ionically bonded in magnetite so the former is typically enriched in ¹⁸O.

T, °C 1000 300 200 100 0 500 50 1.06 1.04 $\alpha_{_{CO_2-H_2O}}$ Q = 0.969 + 0.243/1 1.02 1.00 0.98 1 3 0 2 1000/T, K T, ℃ 1000 300 200 100 50 $d = 0.998 + 0.005/T^{2}$ 1.06 $\alpha_{_{CO_2}-H_2O}$ 1.04 1.02 1.00 0.98 5 10 0 $10^{6}/T^{2}, K$

Figure 9.3. Calculated value of α^{18} O for CO₂–H₂O, shown vs. 1/T and $1/T^2$. Dashed lines show that up to ~200° C, $\alpha \approx 0.969 + 0.0243/T$. At higher temperatures, $\alpha \approx 0.9983 + 0.0049/T^2$. Calculated from the data of Richet et al. (1977).

¹⁸O rich mineral and magnetite the least. Oxygen is dominantly covalently bonded in quartz, but dominantly ionically bonded in magnetite. The O is bound more strongly in quartz than in magnetite, so the former is typically enriched in ¹⁸O.

Substitution of cations in a dominantly ionic site (typically the octahedral sites) in silicates has only a secondary effect on the O bonding, so that isotopic fractionations of O isotopes between similar silicates are generally small. Substitutions of cations in sites that have a strong covalent character (generally tetrahedral sites) result in greater O isotope fractionations. Thus, for example, we would expect the fractionation between the end-members of the alkali feldspar series and water to be similar, since only the substitution of K^+ for Na⁺ is involved. We would expect the fractionation factors between end-members of the plagioclase series and water to be greater, since this series involves the substitution of Al for Si as well as Ca for Na, and the bonding of O to Si and Al in tetrahedral sites has a large covalent component.

CHAPTER 9: STABLE ISOTOPES

Example 9.1. Predicting Isotopic Fractionations

Consider the exchange of ¹⁸O and ¹⁶O between carbon monoxide and oxygen:

$$\mathbf{C}^{16}\mathbf{O} + {}^{16}\mathbf{O}^{18}\mathbf{O} \rightleftharpoons \mathbf{C}^{18}\mathbf{O} + {}^{16}\mathbf{O}_2$$

The frequency for the C-¹⁶O vibration is 6.505×10^{13} sec⁻¹, and the frequency of the ¹⁶O₂ vibration is 4.738×10^{13} sec⁻¹. How will the fractionation factor, $\alpha = ({}^{18}O/{}^{16}O)_{CO}/{}^{18}O/{}^{16}O)_{O2}$, vary as a function of temperature?

Answer: The equilibrium constant for this reaction is:

$$=\frac{Q_{C^{18}o}Q_{1^{6}o_{2}}}{Q_{C^{16}o}Q_{1^{8}o^{16}o}}$$
 9.31

The rotational and translational contributions are independent of temperature, so we calculate them first. The rotational contribution is:

$$\mathbf{K}_{rot} = \left(\frac{Q_{C^{18}O}Q_{1^{6}O_{2}}}{Q_{C^{16}O}Q_{1^{8}O^{16}O}}\right)_{rot} = \frac{\mu_{C^{18}O}\sigma_{C^{16}O}}{\mu_{C^{16}O}\sigma_{C^{18}O}}\frac{\mu_{1^{6}O_{2}}\sigma_{1^{8}O^{16}O}}{\mu_{1^{8}O^{16}O}\sigma_{1^{6}O_{2}}} = \frac{1}{2}\frac{\mu_{C^{18}O}\mu_{1^{6}O_{2}}}{\mu_{C^{16}O}\mu_{1^{8}O^{16}O}} \qquad 9.32$$

Κ

(the symmetry factor, σ , is 2 for ${}^{16}O_2$ and 1 for the other molecules). Substituting values $\mu_{C^{16}O} = 6.857$, $\mu_{C^{18}O} = 7.20$, $\mu_{18}O_{16}O = 8.471$, $\mu_{16}O_2 = 8$, we find: $K_{rot} = 0.9917/2$.

The translational contribution is:

$$\mathbf{K}_{trans} = \frac{m_{C^{18}O}^{3/2}}{m_{C^{16}O}^{3/2}} \frac{m_{^{16}O_2}^{3/2}}{m_{^{18}O^{16}O}^{3/2}} \qquad 9.33$$

Substituting $m_{C^{16}O} = 28$, $m_{C^{18}O} = 30$, $m_{18O^{16}O} = 34$, $m_{16O_2} = 32$ into 9.33, we find K_{trans} = 1.0126. The vibrational contribution to the equilibrium constant is:

$$K_{vib} = \frac{e^{-h(v_{c^{18}o} - v_{c^{16}o} + v_{16}o_{2} - v_{18}o^{16}o)/2kT} (1 - e^{-hv_{c^{16}o}/kT})(1 - e^{-hv_{18}o^{16}o/kT})}{(1 - e^{-hv_{c^{18}o}/kT})(1 - e^{-hv_{16}o_{2}/kT})} \qquad 9.34$$

To obtain the vibrational contribution, we can a ssume the atoms vibrate as harmonic oscillators and, using experimentally determined vibrational frequencies for the ¹⁶O molecules, solve equation 9.17 for the forcing constant, k, and calculate the vibrational frequencies for the ¹⁸O-bearing molecules. These turn out to be 6.348×10^{13} sec⁻¹ for C-¹⁸O and 4.605×10^{13} sec⁻¹ for ¹⁸O¹⁶O, so that 8.33a becomes:

$$K_{vib} = \frac{e^{-5.580/T} (1 - e^{-3119/T})(1 - e^{-2208/T})}{(1 - e^{-3044/T})(1 - e^{-2272/T})}$$

If we carry the calculation out at T = 300 K,

If we carry the calculation out at T = 300 K, we find:

$$\mathbf{K} = \mathbf{K}_{\rm rot} \mathbf{K}_{\rm trans} \mathbf{K}_{\rm vib} =$$

$$\frac{1}{2}$$
 0.9917 × 1.0126 × 1.1087 = $\frac{1.0229}{2}$

To calculate the fractionation factor α from the equilibrium constant, we need to calculate K_{∞}

$$K_{\infty} = \frac{(1/1)^1}{(2/1)^1} = \frac{1}{2}$$







Geochemistry

CHAPTER 9: STABLE ISOTOPES

Carbonates tend to be very ¹⁸O rich because O is bonded to a small, highly charged atom, C⁴⁺. The fractionation, $\Delta^{18}O_{cal-water}$, between calcite and water is about 30 per mil at 25° C. The cation in the carbonate has a secondary role (due to the effect of the mass of the cation on vibrational frequency). The $\Delta^{18}O_{carb-H_2O}$ decreases to about 25 when Ba replaces Ca (Ba has about 3 times the mass of Ca).

Crystal structure plays a secondary role. The Δ^{18} O between aragonite and calcite is of the order of 0.5 permil. However, is there apparently a large fractionation (10 permil) of C between graphite and diamond.

Pressure effects on fractionation factors turn out to be small, no more than 0.1 permil over 0.2 GPa. We can understand the reason for this by recalling that $\partial \Delta G / \partial P = \Delta V$. The volume of an atom is entirely determined by its electronic structure, which does not depend on the mass of the nucleus. Thus the volume change of an isotope exchange reaction will be small, and hence there will be little pressure dependence. There will be a minor effect because vibrational frequency and bond length change as crystals are compressed. The compressibility of silicates is of the order of 1 part in 10⁴, so we can expect effects on the order of 10^{-4} or less, which are generally insignificant.

9.2.2 Kinetic Isotope Fractionations

Kinetic isotope fractionations are normally associated with fast, incomplete, or unidirectional processes like evaporation, diffusion, dissociation reactions, and biologically mediated reactions. As an example, recall that temperature is related to the average kinetic energy. In an ideal gas, the average kinetic energy of all molecules is the same. The kinetic energy is given by:

$$E = \frac{1}{2}mv^2 \qquad 9.35$$

Consider two molecules of carbon dioxide, ${}^{12}C^{16}O_2$ and ${}^{13}C^{16}O_2$, in such a gas. If their energies are equal, the ratio of their velocities is $(45/44)^{1/2}$, or 1.011. Thus ${}^{12}C^{16}O_2$ can diffuse 1.1% further in a given amount of time than ${}^{13}C^{16}O_2$. This result, however, is largely limited to ideal gases, i.e., low pressures where collisions between molecules are infrequent and intermolecular forces negligible. For the case where molecular collisions are important, the ratio of their diffusion coefficients is the ratio of the square roots of the reduced masses of CO_2 and air (mean molecular weight 28.8):

$$\frac{D_{^{12}CO_2}}{D_{^{13}CO_2}} = \frac{\sqrt{\mu_{^{12}CO_2 - air}}}{\sqrt{\mu_{^{13}CO_2 - air}}} = \frac{17.561}{17.406} = 1.0044$$
9.36

Hence we would predict that gaseous diffusion will lead to a 4.4% rather than 11% fractionation.

Molecules containing the heavy isotope are more stable and have higher dissociation energies than those containing the light isotope. This can be readily seen in Figure 9.1. The energy required to raise the D₂ molecule to the energy where the atoms dissociate is 441.6 kJ/mole, whereas the energy required to dissociate the H₂ molecule is 431.8 kJ/mole. Therefore it is easier to break the H-H than D-D bond. Where reactions attain equilibrium, isotopic fractionations will be governed by the considerations of equilibrium discussed above. *Where reactions do not achieve equilibrium the lighter isotope will usually be preferentially concentrated in the reaction products,* because of this effect of the bonds involving light isotopes in the reactants being more easily broken. Large kinetic effects are associated with biologically mediated reactions (e.g., photosynthesis, bacterial reduction), because such reactions generally do not achieve equilibrium and do not go to completion (e.g., plants don't convert all CO₂ to organic carbon). Thus ¹²C is enriched in the products of photosynthesis in plants (hydrocarbons) relative to atmospheric CO₂, and ³²S is enriched in H₂S produced by bacterial reduction of sulfate.

We can express this in a more quantitative sense. In Chapter 5, we found the reaction rate constant could be expressed as:

$$k = A e^{-E_b/kT} \tag{5.24}$$

where k is the rate constant, A the frequency factor, and E_b is the barrier energy. For example, in a dissociation reaction, the barrier energy is the difference between the dissociation energy and the zeropoint energy when the molecule is in the ground state, or some higher vibrational frequency when it is

Geochemistry

CHAPTER 9: STABLE ISOTOPES

not (Fig. 9.1). The frequency factor is independent of isotopic composition, thus the ratio of reaction rates between the HD molecule and the H_2 molecule is:

$$\frac{k_D}{k_H} = \frac{e^{-(\varepsilon - 1/2 h v_D)/kT}}{e^{-(\varepsilon - 1/2 h v_H)/kT}}$$
9.37

$$\frac{k_D}{k_H} = e^{(v_H - v_D)\hbar/2kT}$$
9.38

or

Substituting for the various constants, and using the wavenumbers given in the caption to Figure 9.1 (remembering that $\omega = cv$, where *c* is the speed of light) the ratio is calculated as 0.24; in other words we expect the H₂ molecule to react four times faster than the HD molecule, a very large difference. For heavier elements, the rate differences are smaller. For example, the same ratio calculated for ¹⁶O₂ and ¹⁸O¹⁶O shows that the ¹⁶O₂ will react about 15% faster than the ¹⁸O¹⁶O molecule.

The greater translational velocities of lighter molecules also allow them to break through a liquid surface more readily and hence evaporate more quickly than a heavy molecule of the same composition. Thus water vapor above the ocean is typically around $\delta^{18}O = -13$ per mil, whereas at equilibrium the vapor should only be about 9 per mil lighter than the liquid.

Let's explore this a bit further. An interesting example of a kinetic effect is the fractionation of O isotopes between water and water vapor. This is an example of Rayleigh distillation (or condensation), and is similar to fractional crystallization. Let *A* be the amount of the species containing the major isotope, e.g., $H_2^{16}O$, and *B* be the amount of the species containing the minor isotope, e.g., $H_2^{18}O$. The rate at which these species evaporate is proportional to the amount present:

$$dA = k_A A \qquad 9.39a$$

$$dB = k_B B \qquad 9.39b$$

Since the isotopic composition affects the reaction, or evaporation, rate, $k_A \neq k_B$. Earlier we saw that for equilibrium fractionations, the fractionation factor is related to the equilibrium constant. For kinetic fractionations, the fractionation factor is simply the ratio of the rate constants, so that:

$$\frac{k_B}{k_A} = \alpha \qquad 9.40$$

and

and

Rearranging and integrating, we have:

 $\ln\frac{B}{B^{\circ}} = \alpha \ln\frac{A}{A^{\circ}}$

 $\frac{B}{B^{\circ}} = \left(\frac{A}{A^{\circ}}\right)^{\alpha}$

9.42

or

where A° and B° are the amount of A and B originally present. Dividing both sides by A/A° :

$$\frac{B/A}{B^{\circ}/A^{\circ}} = \left(\frac{A}{A^{\circ}}\right)^{\alpha-1} \qquad 9.43$$

Since the amount of B makes up only a trace of the total amount of H_2O present, *A* is essentially equal to the total water present, and A/A° is essentially identical to *f*, the fraction of the origi-



Figure 9.5. Fractionation of isotope ratios during Rayleigh and equilibrium condensation. δ is the per mil difference between the isotopic composition of original vapor and the isotopic composition when fraction *f* of the vapor remains.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

nal water remaining. Hence:

$$\frac{B/A}{B^{\circ}/A^{\circ}} = f^{\alpha - 1}$$
9.44

Subtracting 1 from both sides, we have:

$$\frac{B / A - B^{\circ} / A^{\circ}}{B^{\circ} / A^{\circ}} = f^{\alpha - 1} - 1$$
 9.45

The left side of 9.45 is the relative deviation from the initial ratio. The permil relative deviation is simply:

$$\Delta = 1000(f^{\alpha - 1} - 1)$$
 9.46

Of course, the same principle applies when water condenses from vapor. Assuming a value of α of 1.01, δ will vary with *f*, the fraction of vapor remaining, as shown in Figure 9.5.

Even if the vapor and liquid remain in equilibrium through the condensation process, the isotopic composition of the remaining vapor will change continuously. The relevant equation is:

$$\Delta = \left(1 - \frac{1}{(1-f)/\alpha + f}\right) \times 1000$$
9.47

The effect of equilibrium condensation is also shown in Figure 9.5.

9.2.3 Mass Dependent and Mass Independent Fractionations

Oxygen and sulfur have both have more than two isotopes, as do, of course, many other elements. How do we expect the fractionations between various isotope ratios to be related? For example, if a 4‰ fractionation of δ^{18} O is observed in a particular sample, what value of δ^{17} O do we predict? A first guess might be that we would expect the difference in the fractionation to be proportional to the mass difference. In other words, if δ^{18} O is increased by 4‰ in some process, we would expect δ^{17} O to be increased by about 2‰. If we think about this in a little more detail, we realize that fractionation factors are more complex functions of mass. Looking at equation 9.26, for example, we see that mass occurs in a variety of ways, as $m^{3/2}$, as reduced mass, and complexly in the exponential term. Consequently, the ratio of fractionation of 17 O/ 16 O to that of 18 O/ 16 O won't be exactly ½. In fact, in almost all cases the ratio is, on average, about 0.52. Thus the fractionation between isotopes does seem to be proportional to the difference in mass – this is referred to as *mass dependent fractionation*.

There are, however, some exceptions where the ratio of fractionation of ${}^{17}\text{O}/{}^{16}\text{O}$ to that of ${}^{18}\text{O}/{}^{16}\text{O}$ is close to 1. Since the extent of fractionation in these cases seems independent of the mass difference, this is called *mass independent fractionation*. Mass independent fractionation is rare. It was first observed oxygen isotope ratios in meteorites (see Chapter 10) and has subsequently been observed in oxygen isotope ratios of atmospheric gases, most dramatically in stratospheric ozone (Figure 9.6), and most recently in sulfur isotope ratios of Archean sediments and modern sulfur-bearing aerosols in ice. The causes of mass independent fractionation are poorly understood and it seems likely there may be more than one cause.

There is at least a partial theoretical explanation in the case of atmospheric ozone (Gao and Marcus, 2001). Their theory can be roughly explained as follows. Formation of ozone in the stratosphere typically involves the energetic collision of monatomic and molecular oxygen, i.e.:

$O + O_2 \rightarrow O_3$

The ozone molecule thus formed is in an vibrationally excited state and, consequently, subject to dissociation if it cannot loose this excess energy. The excess vibrational energy can be lost either by collisions with other molecules, or by partitioning to rotational energy. In the stratosphere, collisions infrequent compared hence repartitioning of vibrational energy represents an important pathway to stability. Because there are more possible energy transitions for asymmetric species such as ¹⁶O¹⁶O¹⁸O and ¹⁶O¹⁶O¹⁷O than symmetric ones such as ¹⁶O¹⁶O¹⁶O, the latter is repartition its excess energy and form a stable molecule. At higher pressures, such as prevail in the troposphere, the symmetric molecule can loose energy through far more frequent collisions, lessening the importance of the vibrational to rotational

CHAPTER 9: STABLE ISOTOPES



Figure 9.6. Oxygen isotopic composition in the stratosphere and troposphere show the effects of mass independent fractionation. A few other atmospheric trace gases show similar effects. Essentially all other material from the Earth and Moon plot on the *terrestrial fractionation line*.

energy conversion. Gao and Marcus (2001) were able to closely match observed experimental fractionations, but their approach was in part empirical because a fully quantum mechanical treatment is not yet possible.

Theoretical understanding of mass independent sulfur isotope fractionations is less advanced. Mass independent fractionations similar to those observed (discussed below) have been produced in the laboratory by photo dissociation (photolysis) of SO_2 and SO using deep ultraviolet radiation (wavelengths <220 nm). Photolysis at longer wavelengths does not produce mass independent fractionations. Current explanations therefore focus on ultraviolet photolysis. However, there as yet is no theoretical explanation of this effect and alternative explanations, including ones that involve the role in symmetry in a manner analogous to ozone, cannot be entirely ruled out.

9.3 Isotope Geothermometry

One of the principal uses of stable isotopes is geothermometry. Like "conventional" chemical geothermometers, stable isotope geothermometers are based on the temperature dependence of the equilibrium constant. As we have seen, this dependence may be expressed as:

$$\ln K = \ln \alpha = A + \frac{B}{T^2} \qquad 9.48$$

In actuality, the constants *A* and *B* are slowly varying functions of temperature, such that K tends to zero at absolute 0, corresponding to complete separation, and to 1 at infinite temperature, corresponding to no isotope separation. We can obtain a qualitative understanding of why this as so by recalling

FADLE	9.2.	Coefficient	s for	Oxygen	ISOTOPE
RACT	IONAT	ion at Low	Темре	RATURES:	

$\Delta_{QZ-\phi} = \Delta_{QZ-\phi}$	A + B	$\times 10^{\circ}$	T^2
---------------------------------------	-------	---------------------	-------

φ	А	В	
Feldspar	0	$0.97 + 1.04b^*$	
Pyroxene	0	2.75	
Garnet	0	2.88	
Olivine	0	3.91	
Muscovite	-0.60	2.2	
Amphibole	-0.30	3.15	
Biotite	-0.60	3.69	
Chlorite	-1.63	5.44	
Ilmenite	0	5.29	
Magnetite	0	5.27	

* *b* is the mole fraction of anorthite in the feldspar. This term therefore accounts for the compositional dependence discussed above. From Javoy (1976).

CHAPTER 9: STABLE SOTOPES

that the entropy of a system increases with temperature. At infinite temperature, there is complete disorder, hence isotopes would be mixed randomly between phases (ignoring for the moment the slight problem that at infinite temperature there would be neither phases nor isotopes). At absolute 0, there is perfect order, hence no mixing of isotopes between phases. A and B are, however, sufficiently invariant over a limited range temperatures that they can be viewed as constants. We have also noted that at low temperatures, the form of equation 9.48 changes to K $\propto 1/T$.

In principal, a temperature may be calculated from the isotopic fractiona-

Table 9.3 Coefficients for Oxygen Isotope Fractionations at Elevated Temperatures (600° - 1300°C)

	,					
	Cc	Ab	An	Di	Fo	Mt
Qz	0.38	0.94	1.99	2.75	3.67	6.29
Cc		0.56	1.61	2.37	3.29	5.91
Ab			1.05	1.81	2.73	5.35
An				0.76	1.68	4.30
Di					0.92	3.54
Fo						2.62

Coefficients are for mineral pair fractionations expressed as $1000\alpha = B \times 10^6/T^2$ where B is given in the Table. Qz: quartz, Cc: calcite, Ab: albite, An: anorthite, Di: diopside, Fo: forsterite, Mt: magnetite. For example, the fractionation between albite and diopside is $1000\alpha_{An-Di} = 1.81 \times 10^6/T^2$ (T in kelvins). From Chiba et al. (1989).

tion between any two phases provided the phases equilibrated and the temperature dependence of the fractionation factor is known. And indeed, there are too many isotope geothermometers for all of them to be mentioned here. Figure 9.7 shows some fractionation factors between quartz and other minerals



Figure 9.7. Oxygen isotope fractionation for several mineral pairs as a function of temperature.

as a function of temperature. Figure 9.8 shows sulfur isotope fractionation factors between various sulfur-bearing species and H_2S . Table 9.2 lists coefficients *A* and *B* for equation 9.48 for the oxygen isotope fractionation factor between quartz and other oxides and silicates.

Because of the dependence of the equilibrium constant on the inverse square of temperature, stable isotope geothermometry is employed primarily at low temperatures, that is, non-magmatic temperatures. At temperatures in excess of 800°C or so, the fractionations are generally small, making accurate temperatures difficult to determine from them. However, even at temperatures of the upper mantle (1000°C or more), fractionations, although small, remain significant. Experimentally determined fractionation factors for minerals in the temperature range of 600° to 1300°C, which agree well with theoretical calculations, are given in Table 9.3.

Table 9.4 lists similar coefficients for the sulfur fractionation between H_2S and sulfurbearing compounds. Recall that if phases α and γ and α and β are in equilibrium with each other, then γ is also in equilibrium with β . Thus these tables may be used to obtain the fractionation between any two of the phases listed.

All geothermometers are based on the ap-

Geochemistry

W. M. White

CHAPTER 9: STABLE ISOTOPES

parently contradictory assumptions that complete equilibrium was achieved between phases during, or perhaps after, formation of the phases, but that the phases did not re-equilibrate when they subsequently cooled. The reason these assumptions can be made and geothermometry works at all is the exponential dependence of reaction rates on temperature that we discussed in Chapter 5. Isotope geothermometers have the same implicit assumptions about the achievement of equilibrium as other geothermometers.

The importance of the equilibrium basis of geothermometry must be emphasized. Because most stable isotope geothermometers (though not all) are applied to relatively low temperature situations, violation of the assumption that complete equilibrium was achieved is not uncommon. We have seen that isotopic fractionations may arise from kinetic as well as equilibTable 9.4. Coefficients for Sulfur Isotope Fractionation:

$\Delta \phi - H_2 S = A \times 10^6 / T^2 + B \times 10^3 / T \text{ (T in kelvins)}$					
ϕ	В	Α	T°C		
			Range		
CaSO ₄	6.0±0.5	5.26	200-350		
SO_2	-5 ± 0.5	4.7	350-1050		
FeS ₂		$0.4{\pm}0.08$	200-700		
ZnS		$0.10{\pm}0.05$	50-705		
CuS		-0.4 ± 0.1			
Cu ₂ S		-0.75 ± 0.1			
SnS		-0.45 ± 0.1			
MoS_2		0.45 ± 0.1			
Ag ₂ S		-0.8 ± 0.1			
PbS		-0.63 ± 0.05	50-700		

From Ohmoto and Rye (1979)

rium effects. If reactions do not run to completion, the isotopic differences may reflect kinetic effects as much as equilibrium effects. There are other problems that can result in incorrect temperature as well; for example, the system may partially re-equilibration at some lower temperature during cooling. A further problem with isotope geothermometry is that free energies of the exchange reactions are rela-

tively low; meaning there is little chemical energy available to drive the reaction. Indeed, isotopic equilibrium probably often depends on other reactions occurring that mobilize the element involved in the exchange. Solid-state exchange reactions will be particularly slow at temperatures well below the melting point. Equilibrium between solid phases will thus generally depend on reaction of these phases with a fluid. This latter point is true of 'conventional' geothermometers as well, and metamorphism, one of the important areas of application of isotope geothermometry, generally occurs in the presence of a fluid.

Isotope geothermometers do have several advantages over conventional chemical ones. First, as we have noted, there is no volume change associated with isotopic exchange reactions and hence no pressure dependence of the equilibrium constant. However, Rumble has suggested an indirect pressure dependence, wherein the fractionation factor depends on fluid composition, which in turn depends on Second, whereas conventional pressure. chemical geothermometers are generally based on solid solution, isotope geothermometers can make use of pure phases such as SiO₂, etc. Generally, any dependence on the composition of phases involved is of relatively second order



Figure 9.8. Relationship of S isotope fractionation between H₂S and other sulfur-bearing species and temperature.

CHAPTER 9: STABLE SOTOPES

Example 9.2: Oxygen Isotope Geothermometry

A granite-gneiss contains coexisting quartz, muscovite and magnetite of with the following δ^{18} O: quartz: 11.1, magnetite: 1.9. Find the temperature of equilibration.

Answer: According to Bottinga and Javoy (1973), the fraction factors for these minerals can be expressed as:

$$1000 \ln \alpha_{Qz-H_2O} = -3.70 + \frac{4.10 \times 10^6}{T^2} \quad \text{and} \quad 1000 \ln \alpha_{Mt-H_2O} = -3.10 + \frac{1.9 \times 10^6}{T^2}$$

The fractionation factor α_{Qz-Mt} can be found as: $\alpha_{Qz-H_2O}/\alpha_{Mt-H_2O}$, and

$$1000 \ln \alpha_{Qz-Mt} = 1000 (\ln \alpha_{Qz-H_2O} - \ln \alpha_{Mt-H_2O})$$

$$= -3.70 + \frac{4.10}{T^2} + 3.10 - \frac{1.9 \times 10^6}{T^2} = 0.6 + \frac{2.20 \times 10^6}{T^2}$$

Substituting Δ_{Qz-Mag} for 1000 ln α_{Qz-Mag} and solving for T:

$$T = \sqrt{\frac{2.20 \times 10^6}{\Delta_{Qz-Mt} - 0.6}}$$
9.49

We calculate Δ_{Oz-Mag} as 9.2‰. Substituting this into 9.49, we find T = 505 K, = 232 ° C.

importance (there are, however, exceptions). For example, isotopic exchange between calcite and water is independent of the concentration of CO_2 in the water. Compositional effects can be expected only where it effects bonds formed by the element involved in the exchange. For example, we noted substitution of Al for Si in plagioclase affects O isotope fractionation factors because the nature of the bond with oxygen.

9.4 Isotope Fractionation in the Hydrologic System

As we noted above, isotopically light water has a higher vapor pressure, and hence lower boiling point, than isotopically heavy water. Let's consider this in a bit more detail. Raoult's Law (equ. 3.8) states that the partial pressure of a species above a solution is equal to its molar concentration in the solution times the partial pressure exerted by the pure solution. So for the two isotopic species of water (we will restrict ourselves to O isotopes for the moment), Raoult's Law is:

$$p_{H_2^{16}O} = p_{H_2^{16}O}^o[\mathrm{H}_2^{16}\mathrm{O}]$$
 9.50a

and

$$p_{H_2^{18}O} = p_{H_2^{18}O}^o [H_2^{18}O]$$
 9.49b

where *p* is the partial pressure and, as usual, the brackets indicate the aqueous concentration. Since the partial pressure of a species is proportional to the number of atoms of that species in a gas, we can define the fractionation factor, α , between liquid water and vapor as:

$$\alpha_{\nu/l} = \frac{p_{H_2^{18}O} / p_{H_2^{16}O}}{[H_2^{18}O] / [H_2^{16}O]}$$
9.51

Substituting 9.50a and 9.50b into 9.51, we arrive at the relationship:

$$\alpha_{v/l} = \frac{p_{H_2^{18}O}^o}{p_{H_2^{16}O}^o}$$
9.52

Thus interestingly enough, the fractionation factor for oxygen between water vapor and liquid turns out to be just the ratio of the standard state partial pressures. The next question is how the partial pressures vary with temperature. Thermodynamics provides the answer. The temperature dependence of the partial pressure of a species may be expressed as:

CHAPTER 9: STABLE ISOTOPES

$$\frac{d\ln p}{dT} = \frac{\Delta H}{RT^2} \qquad 9.53$$

where *T* is temperature, ΔH is the enthalpy or latent heat of evaporation, and R is the gas constant. Over a sufficiently small range of temperature, we can assume that ΔH is independent of temperature. Rearranging and integrating, we obtain:

$$\ln p = \frac{\Delta H}{RT} + const \quad 9.54$$

We can write two such equations, one for $[H_2^{16}O]$ and [one for $H_2^{18}O$]. Dividing one by the other we obtain:

$$\ln \frac{p_{H_2^{18}O}^o}{p_{H_2^{16}O}^o} = a - \frac{B}{RT} \qquad 9.55$$

where *a* and *B* are constants. This can be rewritten as:

$$\alpha = a e^{B/RT} \qquad 9.56$$

Over a larger range of temperature, ΔH is not constant. The log of the fractionation factor in that case depends on the inverse square of temperature, so that the temperature dependence of the fractionation factor can be represented as:

$$\ln \alpha = a - \frac{B}{T^2} \qquad 9.57$$

Given the fractionation between water and vapor, we might predict that there will be considerable variation in the isotopic composition of water in the hydrologic cycle, and indeed there is. Figure 9.9 shows the global variation in δ^{18} O in *precipitation*.



Figure 9.9. Variation of δ^{18} O in precipitation as a function of mean annual temperature.



Figure 9.10. Cartoon illustrating the process of Rayleigh fractionation and the decreasing δ^{18} O in rain as it moves inland.

Precipitation of rain and snow from clouds is a good example of Rayleigh condensation. Isotopic fractionations will therefore follow equation 9.46. Thus in addition to the temperature dependence of α , the isotopic composition of precipitation will also depend on f, the fraction of water vapor remaining in the air. The further air moves from the site of evaporation, the more water is likely to have condensed and fallen as rain, and therefore, the smaller the value of f in equation 9.46. Thus fractionation will increase with distance from the region of evaporation (principally tropical and temperature oceans). Topography also plays a role as mountains force air up, causing it to cool and water vapor to

Geochemistry

CHAPTER 9: STABLE ISOTOPES

condense, hence decreasing f. Thus precipitation from air that has passed over a mountain range will be isotopically lighter than precipitation on the ocean side of a mountain range. These factors are illustrated in the cartoon in Figure 9.10.

Hydrogen as well as oxygen isotopes will be fractionated in the hydrologic cycle. Indeed, δ^{18} O and δ D are reasonably well correlated in precipitation, as is shown in Figure 9.11. The fractionation of hy-

drogen isotopes, however, is greater because the mass difference is greater.

9.5 ISOTOPE FRACTIONATION IN Biological Systems

Biological processes often inlarge isotopic volve fractionations. Indeed, for carbon, nitrogen, and sulfur, biological processes are the most important cause of isotope fractionations. For the most part, the largest fractionations occur during the initial production of organic matter by the so-called primary producers, or autotrophs. These include all plants and many kinds of bacteria. The most important means of production of organic matter is photosynthesis, but organic matter may also be produced by chemosynthesis, for



Figure 9.11. Northern hemisphere variation in δD and $\delta^{18}O$ in precipitation and meteoric waters. The relationship between δD and $\delta^{18}O$ is approximately $\delta D = 8 \times \delta^{18}O + 10$. After Dansgaard (1964).

example at mid-ocean ridge hydrothermal vents. Large fractionations of both carbon and nitrogen isotopes occur during primary production. Additional fractionations also occur in subsequent reactions and up through the food chain as *hetrotrophs* consume primary producers, but these are generally smaller.

9.5.1 Carbon Isotope Fractionation During Photosynthesis

Biological processes are the principal cause of variations in carbon isotope ratios. The most important of these processes is photosynthesis (a discussion of photosynthesis may be found in Chapter 14). As we earlier noted, photosynthetic fractionation of carbon isotopes is primarily kinetic. The early work of Park and Epstein (1960) suggested fractionation occurred in several steps. Subsequent work has elucidated the fractionations involved in these steps, which we will consider in slightly more detail.

For terrestrial plants (those utilizing atmospheric CO_2), the first step is diffusion of CO_2 into the boundary layer surrounding the leaf, through the stomata, and internally in the leaf. On theoretical grounds, a fractionation of -4.4% is expected, as we found in equation 9.36. Marine algae and aquatic plants can utilize either dissolved CO_2 or HCO_3^- for photosynthesis:

$$CO_{2(g)} \rightarrow CO_{2(aq)} + H_2O \rightarrow H_2CO_3 \rightarrow H^+ + HCO_3^-$$

An equilibrium fractionation of +0.9 per mil is associated with dissolution ($^{13}CO_2$ will dissolve more readily), and an equilibrium +7.0 to +8% fractionation occurs during hydration and dissociation of CO₂ (i.e., steps 2 and 3 in the reaction above).

At this point, there is a divergence in the chemical pathways. Most plants use an enzyme called *ribulose bisphosphate carboxylase oxygenase* (RUBISCO) to catalyze a reaction in which *ribulose bisphosphate carboxylase* reacts with one molecule of CO_2 to produce 3 molecules of 3-phosphoglyceric acid, a com-

Geochemistry

CHAPTER 9: STABLE ISOTOPES

pound containing 3 carbon atoms, in a process called *carboxylation* (Figure 9.12). The carbon is subsequently reduced, carbohydrate formed, and the ribulose bisphosphate regenerated. Such plants are called C_3 plants, and this process is called the Benson-Calvin Cycle, or simply the Calvin Cycle. C_3 plants constitute about 90% of all plants today and include algae and autotrophic bacteria and comprise the majority of cultivated plants, including wheat, rice, and nuts. There is a kinetic fractionation associated with carboxylation of ribulose bisphosphate that has been determined by several methods to be – 29.4‰ in higher terrestrial plants. Bacterial carboxylation has different reaction mechanisms and a smaller fractionation of about -20‰. Thus for terrestrial plants a fractionation of about -34‰ is expected from the sum of the individual fractionations. The actual observed total fractionation is in the range of -20 to -30‰. The disparity between the observed total fractionation and that expected from the sum of the steps presented something of a conundrum. This solution appears to be a model that assumes the amount of carbon isotope fractionation expressed in the tissues of plants depends on the ratio of the concentration of CO_2 inside plants to that in the external environment: the more photosynthesis depletes the CO_2 in the plant interior, the less the fractionation that occurs.

The other photosynthetic pathway is the Hatch-Slack cycle, used by the C_4 plants that include hotregion grasses and related crops such as maize and sugarcane. These plants use *phosphenol pyruvate carboxylase* (PEP) to fix the carbon initially and form oxaloacetate, a compound that contains 4 carbons (Fig. 9.13). A much smaller fractionation, about -2.0 to -2.5‰, occurs during this step. In phospho-

phoenol pyruvate carboxylation, the CO_2 is fixed in outer mesophyll cells as oxaloacetate and carried as part of a C_4 acid, either malate or asparatate, to inner bundle sheath cells where it is decarboxylated and refixed by RuBP (Fig. 9.14). The environment in the bundle sheath cells is almost a closed system, so that virtually all the carbon carried there is refixed by RuBP, so there is little fractionation during this step. Thus C_4 plants have average $\delta^{13}C$ of -13‰, much less than C_3 plants. As in the case of RuBP photosynthesis, the fractionation ar

photosynthesis, the fractionation appears to depend on the ambient concentration of CO_2 .

A third group of plants, the CAM plants, has a unique metabolism called the *Crassulacean acid metabolism*. These plants generally use the C_4 pathway, but can use the C_3 pathway under certain conditions. These plants are generally succulents adapted to arid environments and include pineapple and many cacti; they have $\delta^{13}C$ intermediate between C_3 and C_4 plants.

Terrestrial plants, which utilize CO_2 from the atmosphere, generally produce greater fractionations than marine and aquatic plants, which utilize dissolved CO_2 and HCO_3^- , together referred to as *dissolved inorganic carbon* or DIC. As we noted above, there is about a +8‰ equilibrium fractionation between dissolved CO_2 and HCO_3^- . Since HCO_3^- is about 2 orders of



Figure 9.12. Ribulose bisphosphate (RuBP) carboxylation, the reaction by which C_3 plants fix carbon during photosynthesis.



Figure 9.13. Phosphoenolpyruvate carboxylation, the reaction by which C_4 plants fix CO_2 during photosynthesis.



Figure 9.14. Chemical pathways in C₄ photosynthesis.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

magnitude more abundant in seawater than dissolved CO_2 , marine algae utilize this species, and hence tend to show a lower net fractionation between dissolved carbonate and organic carbon during photosynthesis. Diffusion is slower in water than in air, so diffusion is often the rate-limiting step. Most aquatic plants have some membrane-bound mechanism to pump DIC, which can be turned on when DIC is low. When DIC concentrations are high, fractionation in aquatic and marine plants is generally similar to that in terrestrial plants. When it is low and the plants are actively pumping DIC, the fractionation is less because most of the carbon pumped into cells is fixed. Thus carbon isotope fractionations between dissolved inorganic carbon and organic carbon can be as low as 5‰ in algae.

Not surprisingly, the carbon isotope fractionation in C fixation is also temperature dependent. Thus higher fractionations are observed in cold-water phytoplankton than in warm water species. However, this observation also reflects a kinetic effect: there is generally less dissolved CO₂ available in warm waters because of the decreasing solubility with temperature. As a result, a larger fraction of the CO₂ is utilized and there is consequently less fractionation. Surface waters of the ocean are generally enriched in ¹³C (relative to ¹³C) because of uptake of ¹²C during photosynthesis (Figure 9.15). The degree of enrichment depends on the productivity: biologically productive areas show greater depletion. Deep water, on the other hand, is enriched in ¹²C. Organic matter falls through the water column and is decomposed and "remineralized", i.e., converted to inorganic carbon, by the action of bacteria, enriching deep water in ¹²C and total DIC. Thus biological activity acts to "pump" carbon, and particularly ¹²C from surface to deep waters.

Nearly all organic matter originates through photosynthesis. Subsequent reactions convert the photosynthetically produced carbohydrates to the variety of other organic compounds utilized by organisms. Further fractionations occur in these reactions. These fractionations are thought to be kinetic in origin and may partly arise from organic C-



Figure 9.15. Depth profile of total dissolved inorganic carbon and δ^{13} C in the North Atlantic.

H bonds being enriched in ¹²C and organic C-O bonds are enriched in ¹³C. ¹²C is preferentially consumed in respiration (again, because bonds are weaker and it reacts faster), which enriches residual organic matter in ¹³C. Thus the carbon isotopic composition of organisms becomes slightly more positive moving up the food chain.

The principal exception to the creation of organic matter through photosynthesis is *chemosynthesis*. In chemosynthesis, chemical reactions rather than light provides the energy to "fix" CO_2 . Regardless of the energy source, however, fixation of CO_2 involves the Calvin-Benson cycle and RUBSICO. Not surprisingly, then, chemosynthesis typically results in carbon isotope fractionations similar to those of photosynthesis. Thus large carbon isotope fractionations are the signature of both photosynthesis and chemosynthesis.

9.5.2 Nitrogen Isotope Fractionation in Biological Processes

Nitrogen is another important element in biological processes, being an essential component of all amino acids, proteins, and other key compounds such as RNA and DNA. The understanding of isotopic fractionations of nitrogen is much less advanced than that of carbon. There are five important forms of inorganic nitrogen (N_2 , NO_3^- , NO_2^- , NH_3 and NH_4^+). Equilibrium isotope fractionations occur between these five forms, and kinetic fractionations occur during biological assimilation of nitrogen. Ammonia is the form of nitrogen that is ultimately incorporated into organic matter by growing plants. Most terrestrial plants depend on symbiotic bacteria for *fixation* (i.e., reduction) of N_2 and other

Geochemistry

W. M. White

CHAPTER 9: STABLE SOTOPES

forms of nitrogen to ammonia. Many plants, including many marine algae, can utilize oxidized nitrogen, NO_3^- and NO_2^- , and a few (blue-green algae and legumes, for example) are able to utilize N₂ directly. In these cases, nitrogen must first be reduced by the action of reductase enzymes. As with carbon, fractionation may occur in each of the several steps that occur in the nitrogen assimilation process. Denitrifying bacteria use nitrates as electron donors (i.e., as an oxidant) and reduce it without assimilating it. In this dissimilatory denitrification, there is a significant kinetic fractionation with the light isotope, ¹⁴N, being preferentially reduced leaving residual nitrate enriched in 15 N by 6-7‰.

While isotope fractionations during assimilation of ammonium are still poorly understood, it appears there is a strong depen-



Figure 9.16. Dependence of nitrogen isotope fractionation by bacteria and diatoms on dissolved ammonium concentration.

dence on the concentration of the ammonium ion. Such dependence has been observed, as for example in Figure 9.16. The complex dependence in Figure 9.16 is interpreted as follows. The increase in fractionation from highest to moderate concentrations of ammonium reflects the switching on of active ammonium transport by cells. At the lowest concentrations, essentially all available nitrogen is transported into the cell and assimilated, so there is little fractionation observed.

The isotopic compositions of marine particulate nitrogen and non-nitrogen-fixing plankton are typically -3% to +12% δ^{15} N. Non-nitrogen fixing terrestrial plants unaffected by artificial fertilizers generally have a narrower range of +6% to +13 per mil, but are isotopically lighter on average. Marine blue-green algae range from -4 to +2, with most in the range of -4 to -2%. Most nitrogen-fixing terrestrial plants fall in the range of -2 to +4%, and hence are typically heavier than non-nitrogen fixing plants.

9.5.3 Oxygen and Hydrogen Isotope Fractionation by Plants

Oxygen is incorporated into biological material from CO_2 , H_2O , and O_2 . However, both CO_2 and O_2 are in oxygen isotopic equilibrium with water during photosynthesis, and water is the dominant source of O. Therefore, the isotopic composition of plant water determines the oxygen isotopic composition of

plant material. The oxygen isotopic composition of plant material seems to be controlled by exchange reactions between water and carbonyl oxygens (oxygens doubly bound to carbon):

$$C = {}^{16}O + H_2 {}^{18}O \rightarrow C = {}^{18}O + H_2 {}^{16}O$$

Fractionations of +16 to +27‰ (i.e., the organically bound oxygen is heavier) have been measured for these reactions. Consistent with this, cellulose from most plants has δ^{18} O of +27±3‰. Other factors, however, play a role in the oxygen isotopic composition of plant material. First, the isotopic composition of water varies from δ^{18} O \approx -55‰ in Arctic regions to δ^{18} O \approx 0‰ in the oceans. Second, less than complete equilibrium may be achieved if photosynthesis is occurring at a rapid pace, result-



Figure 9.17. Isotopic Fractionations of hydrogen during primary production in terrestrial plants. After Fogel and Cifuentes (1993).

Geochemistry

CHAPTER 9: STABLE ISOTOPES

ing in less fractionation. Finally, some fractionation of water may occur during transpiration, with residual water in the plant becoming heavier.

Hydrogen isotope fractionation during photosynthesis occurs such that the light isotope is enriched in organic material. In marine algae, isotope fractionations of -100 to -150% have been observed, which is a little more than that observed in terrestrial plants (-86 to -120%). Among terrestrial plants, there appears to be a difference between C₃ and C₄ plants. The former show fractionations of -117 to -121%, while fractionations of -86 to -109% have been observed in C₄ plants. However, little is known in detail about the exact mechanisms of fractionation.

As for oxygen, variations in the isotopic composition of available water and fractionation during transpiration are important in controlling the hydrogen isotopic composition of plants. This is illustrated in Figure 9.17.

9.5.4 Biological Fractionation of Sulfur Isotopes

Though essential to life, sulfur is a minor component in living tissue (C:S atomic ratio is about 200). Plants take up sulfur as sulfate and subsequently reduce it to sulfide and incorporate it into cysteine, an amino acid. There is apparently no fractionation of sulfur isotopes in transport across cell membranes and incorporation, but there is a fractionation of +0.5 to -4.5% in the reduction process, referred to as *assimilatory sulfate reduction*. This is substantially less than the expected fractionation of about -20%, suggesting most sulfur taken up by primary producers is reduced and incorporated into tissue.

Sulfur, however, plays two other important roles in biological processes. First, sulfur, in the form of sulfate, can act as an electron acceptor or oxidant, and is utilized as such by sulfur-reducing bacteria. This process, in which H₂S is liberated, is called *dissimilatory sulfate reduction* and plays an important role in biogeochemical cycles, both as a sink for sulfur and source for atmospheric oxygen. A large fractionation of +5 to -46% is associated with this process. This process produces by far the most significant fractionation of sulfur isotopes, and thus governs the isotopic composition of sulfur in the exogene. Sedimentary sulfate typically has δ^{34} S of about +17, which is similar to the isotopic composition of sulfate in the oceans (+20), while sedimentary sulfide has a δ^{34} S of -18. The living biomass has a δ^{34} S of ≈ 0 .

The final important role of sulfur is a reductant. Sulfide is an electron acceptor used by some types of photosynthetic bacteria as well as other bacteria in the reduction of CO_2 to organic carbon. Unique among these perhaps are the chemosynthetic bacteria of submarine hydrothermal vents. They utilize H_2S emanating from the vents as an energy source and form the base of the food chain in these unique ecosystems. A fractionation of +2 to -18% is associated with this process.



Figure 9.18. Relationship between δ^{13} C and δ^{15} N among the principal classes of autotrophs.

9.5.5 Isotopes and Diet: You Are What You Eat

As we have seen, the two main photosynthetic pathways, C₃ and C₄, lead to organic carbon with different carbon isotopic compositions. Terrestrial C₃ plants have δ^{13} C values that average of about -27‰, C_4 plants an average $\delta^{13}C$ of about -13‰. Marine plants (which are all C₃) utilized dissolved bicarbonate rather than atmospheric CO₂. Seawater bicarbonate is about 8.5% heavier than atmospheric CO₂, and marine plants average about 7.5% heavier than terrestrial C₃ plants. In addition, because the source of the carbon they fix is isotopically more variable, the isotopic composition of marine plants is also more variable. Finally, marine cyanobacteria (blue-green algae) tend to fractionate carbon isotopes less during photosynthesis than do true marine plants, so they tend to average 2 to 3‰ higher in δ^{13} C.

Geochemistry

CHAPTER 9: STABLE ISOTOPES



Figure 9.19. Values of δ^{13} C and δ^{15} N in various marine and terrestrial organisms. From Schoeninger and DeNiro (1984).

Plants may also be divided into two types based on their source of nitrogen: those that can utilized N₂ directly, and those that utilize only "fixed" nitrogen in ammonia and nitrate. The former include the legumes (e.g., beans, peas, etc.) and marine cyanobacteria. The legumes, which are exclusively C₃ plants, utilize both N₂ and fixed nitrogen, and have an average δ^{15} N of +1‰, whereas modern non-leguminous plants average about +3‰. Prehistoric nonleguminous plants were more positive, averaging perhaps +9‰, because the isotopic composition of present soil nitrogen has been affected by the use of chemical fertilizers. For both groups, there was probably a range in δ^{15} N of ±4 or 5‰, because the isotopic composition of soil nitrogen varies and there is some fractionation involved in uptake. Marine plants have δ^{15} N of +7±5‰, whereas marine cyanobacteria have δ^{15} N of -1±3‰. Thus based on their δ^{13} C and δ^{15} N values, autotrophs can be divided into several groups, which are summarized in Figure 9.18.

DeNiro and Epstein (1978) studied the relationship between the carbon isotopic composition of animals and their diet. They found that there is only slight further fractionation of carbon by animals and that the carbon isotopic composition of animal tissue closely reflects that of the animal's diet. Typically, carbon in animal tissue is about 1‰ heavier than their diet. The small fractionation between animal tissue and diet is a result of the slightly weaker bond formed by ¹²C compared to ¹³C. The weaker bonds are more readily broken during respiration, and, not surprisingly, the CO₂ respired by most animals investigated was slightly lighter than their diet. Thus only a small fractionation in carbon isotopes occurs as organic carbon passes up the food web. Terrestrial food chains usually do not have more than 3 trophic levels, implying a maximum further fractionation of +3‰; marine food chains can have up to 7 trophic levels, implying a maximum carbon isotope difference between primary producers and top predators of 7‰. These differences are smaller than the range observed in primary producers.

In another study, DeNiro and Epstein (1981) found that $\delta^{15}N$ of animal tissue reflects the $\delta^{15}N$ of the animal's diet, but is typically 3 to 4‰ higher than that of the diet. Thus in contrast to carbon, significant fractionation of nitrogen isotopes will occur as nitrogen passes up the food chain. These relationships are summarized in Figure 9.19. *The significance of these results is that it is possible to infer the diet of an animal from its carbon and nitrogen isotopic composition.*

Schoeninger and DeNiro (1984) found that the carbon and nitrogen isotopic composition of bone collagen in animals was similar to that of body tissue as a whole. Apatite in bone appears to undergo isotopic exchange with meteoric water once it is buried, but bone collagen and tooth enamel appear to be

Geochemistry

CHAPTER 9: STABLE ISOTOPES



Figure 9.20. δ^{13} C in carbonates from paleosols of the Potwar Plateau in Pakistan. The change in δ^{13} C may reflect the evolution of C₄ plants. From Quade et al. (1989).

global cause, while the monsoons are a regional phenomenon. C_4 photosynthesis is more efficient at low concentrations of CO_2 that is the C_3 pathway. Thus the evolution of C_4 plants 20 may have occurred in response to a decrease in atmospheric CO_2 suspected on other grounds (Cerling et al., 1993). Delemas

9.5.5.1 Isotopes in Archaeology

The differences in nitrogen and carbon isotopic composition of various foodstuffs and the preservation of these isotope ratios in bone collagen provides a means of determining what ancient peoples ate. In the first investigation of bone collagen from human remains, DeNiro and Epstein (1981) concluded that Indians of the Tehuacan Valley in Mexico probably depended heavily on maize (a C₄ plant) as early as 4000 BC, whereas archeological investigations had concluded maize did not become important in their diet until perhaps 1500 BC. In addition, there seemed to be steady increase in the dependence on legumes (probably beans) from 6000 BC to 1000 AD and a more marked increase in legumes in the diet after 1000 AD.

Mashed grain and vegetable charred onto potsherds during cooking provides an additional record of the diets of ancient peoples. DeNiro and Hasdorf (1985) found that vegetable matter subjected to conditions similar to burial in

robust and retain their original isotopic compositions. This means that the nitrogen and carbon isotopic composition of fossil bone collagen and teeth can be used to reconstruct the diet of fossil animals.

Plant photosynthesis can also influence the isotopic composition of soil carbonate: when the plant dies, its organic carbon is incorporated into the soil and oxidized to CO_2 by bacteria and other soil organisms. In arid regions, some of this CO_2 precipitates as soil calcium carbonate. In an area of Pakistan presently dominated by C_4 grasses, Quade et al. (1989) found that a sharp shift in $\delta^{13}C$ in soil carbonate occurred about 7 million years ago (Figure 9.20). Quade et al. concluded that the shift marks the transition from C_3 dominant to C_4 dominant grasslands. They initially interpreted this as a response to the uplift of the Tibetan Plateau and the development of the monsoon. However, other evidence, including oxygen isotope data from Pakistani soil carbonates, suggests the monsoons developed about a million years earlier. Interestingly, other studies indicated that North American grasslands transitioned from C_3 to C_4 dominated about the same time. This synchronicity suggests a



Figure 9.21. δ^{13} C and δ^{15} N of various food stuffs and of diets reconstructed from bone collagen and vegetable matter charred onto pots by DeNiro and colleagues. The Huanca people were from the Upper Mantaro Valley of Peru. Data from pot sherds of the Huanca I period (AD 1000-1200) suggest both C₃ and C₄ plants were cooked in pots, but only C₃ plants during the Huanca II period (AD 1200-1470).

Geochemistry

CHAPTER 9: STABLE ISOTOPES

soil underwent large shifts in δ^{15} N and δ^{13} C but that vegetable matter that was burned or charred did not. The carbonization (charring, burning) process itself produced only small (2 or 3‰) fractionations. Since these fractionations are smaller than the range of isotopic compositions in various plant groups, they are of little significance. Since potsherds are among the most common artifacts recovered in archeological sites, this provides a second valuable means of reconstructing the diets of ancient peoples.

Figure 9.21 summarizes the results obtained in a number of studies of bone collagen and potsherds (DeNiro, 1987). Studies of several modern populations, including Eskimos and the Tlinglit Indians of the Northwest U.S. were made as a control. Judging from the isotope data, the diet of Neolithic Europeans consisted entirely of C_3 plants and herbivores feeding on C_3 plants, in contrast to the Tehuacan Indians, who depended mainly on C_4 plants. Prehistoric peoples of the Bahamas and Denmark depended both on fish and on agriculture. In the case of Mesolithic Denmark, other evidence indicates the crops were C_3 , and the isotope data bear this out. Although there is no corroborating evidence, the isotope data suggest the Bahamians also depended on C_3 rather than C_4 plants. The Bahamians had lower δ^{15} N because the marine component of their diet came mainly from coral reefs. Nitrogen fixation is particularly intense on coral reefs, which leads to 15 N depletion of the water, and consequently, of reef organisms.

9.5.6 Isotopic "Fossils" and the Earliest Life

As noted earlier, large carbon isotope fractionations are common to all autotrophs. Consequently, δ^{13} C values of -20‰ or less are generally interpreted as evidence of biologic origin of those compounds. Schidlowski (1988) first reported δ^{13} C as low as -26‰ in carbonate rocks from West Greenland that are ostensibly older than 3.5 Ga. In 1996, Mojzsis and others reported δ^{13} C between -20 to -50‰ graphite inclusions in grains of apatite in 3.85 Ga in banded-iron formations (BIFs) from the same area. In 1999, Rosing reported δ^{13} C of -19‰ from graphite in tubiditic and pelagic metasedimentary rocks from the Isua greenstone belt in the same area. These rocks are thought to be older than 3.7 Ga. In each case, these negative δ^{13} C values were interpreted as evidence of a biogenic origin of the carbon, and therefore that life existed on Earth at this time. This interpretation remains, however, controversial.

There are several reasons for the controversy, but all ultimately relate to the extremely complex geological history of the area. The geology of the region includes not only the early Archean rocks, but also rocks of middle and late Archean age as well. Most rocks are multiply and highly deformed and metamorphosed and the exact nature, relationships, and structure of the precursor rocks are difficult to decipher. Indeed, Rosing et al. (1996) argued that at least some of the carbonates sampled by Schidlowski (1988) are veins deposited by metamorphic fluid flow rather than metasediments. Others have argued that the graphite in these rocks formed by thermal decomposition of siderite (FeCO₃) and subsequent reduction of some of the carbon. Further effort will be needed to resolve this controversy.

9.6 Paleoclimatology

Perhaps one of the most successful and significant application of stable isotope geothermometry has been paleoclimatology. At least since the work of Louis Agassiz in 1840, geologists have contemplated the question of how the Earth's climate might have varied in the past. Until 1947, they had no means of quantifying paleotemperature changes. In that year, Harold Urey initiated the field of stable isotope geochemistry. In his classic 1947 paper, Urey calculated the temperature dependence of oxygen isotope fractionation between calcium carbonate and water and proposed that the isotopic composition of carbonates could be used as a paleothermometer. Urey's students and post-doctoral fellows empirically determined temperature dependence of the fractionation between calcite and water as:

$$T^{\circ}C = 16.9 - 4.2\Delta_{cal-H_{2}O} + 0.13\Delta_{cal-H_{2}O}^{2}$$
9.58

For example, a change in $\Delta_{cal:H_2O}$ from 30 to 31 permil implies a temperature change from 8° to 12°. Urey suggested that the Earth's climate history could be recovered from oxygen isotope analyses of ancient marine carbonates. Although the problem has turned out to be much more complex than Urey anticipated, this has proved to be an extremely fruitful area of research. Deep-sea carbonate oozes con-

Geochemistry

CHAPTER 9: STABLE SOTOPES

tained an excellent climate record, and, as we shall see, several other paleoclimatic records are available as well.

9.6.1 The Marine Quaternary δ^{18} O Record and Milankovitch Cycles

The principles involved in paleoclimatology are fairly simple. As Urey formulated it, the isotopic composition of calcite secreted by organisms should provide a record of paleo-ocean temperatures because the fractionation of oxygen isotopes between carbonate and water is temperature dependent. In actual practice, the problem is somewhat more complex because the isotopic composition of the shell, or test, of an organism will depend not only on temperature, but also on the isotopic composition of water in which the organism grew, "vital effects" (i.e., different species may fractionate oxygen isotopes somewhat differently), and post-burial isotopic exchange with sediment pore water. As it turns out, the vital effects and post-burial exchange are usually not very important for late Tertiary/Quaternary carbonates, but the isotopic composition of water is.

The first isotopic work on deep-sea sediment cores with the goal of reconstructing the temperature history of Pleistocene glaciations was by Cesare Emiliani (1955), who was then a student of Urey at the University of Chicago. Emiliani analyzed δ^{18} O in foraminifera from piston cores from the world's oceans. Remarkably, many of Emiliani's findings are still valid today, though they have been revised to various degrees. Rather than just the 5 glacial periods that geomorphologists had recognized, Emiliani found 15 glacial–interglacial cycles over the last 600,000 years. He found that these were global events, with notable cooling even in low latitudes and concluded that the fundamental driving force for Quaternary climate cycles was variations in the Earth's orbit and consequent variations in the solar energy flux, *insolation*.

Emiliani had realized that the isotopic composition of the ocean would vary between glacial and in-

terglacial times as isotopically light water was stored in glaciers, thus enriching the oceans in ¹⁸O. He estimated that this factor accounted for about 20% of the observed variations. The remainder he attributed to the effect of temperature on isotope fractionation. Subsequently, Shackleton and Opdyke (1973) concluded that storage of isotopically light water in glacial ice was the main effect causing oxygen isotopic variations in biogenic carbonates, and that the temperature effect was only secondary.

The question of just how much of the variation is deep-sea carbonate sediments in due to ice build-up and how much is due to the effect of temperature on fractionation is an important one. The resolution depends, in part, on the isotopic composition of glacial ice and how much this might vary between glacial and interglacial times. It is fairly clear that the average δ^{18} O of glacial ice is probably less than – 15‰, as Emiliani had assumed. Typical values for Greenland ice are -30 to -35‰ (relative to SMOW) and as low as -50% for Antarctic ice. If the exact isotopic composition of ice and the ice volume were known, it would be a straightforward exercise to calculate the effect of continen-



Figure 9.22. Cartoon illustrating how δ^{18} O of the ocean changes between glacial and interglacial periods.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

tinental ice build-up on ocean isotopic composition. For example, the present volume of continental ice is 27.5×10^6 km³, while the volume of the oceans is 1350×10^6 km³. Assuming glacial ice has a mean δ^{18} O of -30% relative to SMOW, we can calculate the δ^{18} O of the total hydrosphere as -0.6% (neglecting freshwater reservoirs, which are small). At the height of the Wisconsin Glaciation (the most recent one), the volume of glacial ice is thought to have increased by 42×10^6 km², corresponding to a lowering of sea level by 125 m. If the δ^{18} O of ice was the same then as now (-30‰), we can readily calculate that the δ^{18} O of the ocean would have increased by 1.59%. This is illustrated in Figure 9.22.

We can use equation 9.58 to see how much the effect of ice volume on seawater δ^{18} O affects estimated temperature changes. According to this equation, at 20° C, the fractionation between water and calcite should be 33‰. Assuming present water temperature of 20° C, Emiliani would have calculated a temperature change of 6° C for an observed increase in the δ^{18} O of carbonates between glacial times and present δ^{18} O of 2‰, after correction for 0.5‰ change in the isotopic composition of water. In other words, he would have concluded that surface ocean water in the same spot would have had a temperature.



Figure 9.23. A. "Stacking of five cores" selected by Imbrie et al. (1984). Because the absolute value of δ^{18} O varies in from core to core, the variation is shown in standard deviation units. B. Smoothed average of the five cores in A. After Imbrie et al. (1984).

ture of 14° C. If the change in the isotopic composition of water is actually 1.5‰, the calculated temperature difference is only about 2° C. Thus the question of the volume of glacial ice and its isotopic composition must be resolved before δ^{18} O in deep-sea carbonates can be used to calculate paleotemperatures.

Comparison of sealevel curves derived from dating of terraces and coral reefs indicate that each 0.011% variation in δ^{18} O represents a 1 m change in sealevel. Based on this and other observations, it is now generally assumed that the δ^{18} O of the ocean changed by about 1.5% between glacial and interglacial periods, but the exact value is still debated.

By now hundreds, if not thousands, of deep-sea sediment cores have been analyzed for oxygen isotope ratios. Figure 9.23 shows the global δ^{18} O record constructed by averaging analyses from 5 key cores (Imbrie, et al., 1984). A careful examination of the global curve shows a periodicity of approximately 100,000 years. The same periodicity was apparent in Emiliani's initial work and led him to conclude that the glacial-interglacial cycles were due to variations in the Earth's orbital parameters. These are often referred to as the Milankovitch cycles, after Milutin Milankovitch, a Serbian astronomer who argued in the 1920's and 1930's[†] that they caused the ice ages (Milankovitch, 1938).

Let's explore Milankovitch's idea in a bit more detail. The *eccentricity* (i.e., the degree to which the orbit differs from circular) of the Earth's orbit about the Sun and the degree of tilt, or *obliquity*, of the Earth's rotational axis vary slightly. In addition, the direction in which the Earth's rotational axis tilts varies, a phenomenon called *precession*. These variations are illustrated in Figure 9.24. Though variation in these "Milankovitch parameters" has negligible effect on the *total* radiation the Earth receives,

⁺While Milankovitch was a strong and early proponent of the idea that variations in the Earth's orbit caused ice ages, he was not the first to suggest it. J. Croll of Britain first suggested it in 1864, and published several subsequent papers on the subject.

Geochemistry

W. M. White

CHAPTER 9: STABLE ISOTOPES

they do affect the *pattern* of incoming radiation. For example, tilt of the rotational axis determines seasonality and the latitudinal gradient of insolation. The gradient is extremely important because it drives atmospheric and oceanic circulation. If the tilt is small, seasonality will be reduced (cooler summers, warmer winters). Precession relative to the eccentricity of the Earth's orbit also affects seasonality. For example, presently the Earth is closest to the Sun in January. As a result, northern hemisphere winters (and southern hemisphere summers) are somewhat warmer than they would be otherwise. For a given latitude and season, precession will result in a $\pm 5\%$ difference in insolation. While the Earth's orbit is only slightly elliptical, and variations in eccentricity are small, these variations are magnified because insolation varies with the inverse square of the Earth-Sun distance.

Variation in tilt approximates a simple sinusoidal function with a period of 41,000 yrs. Variations in eccentricity can be approximately described with period of 100,000 years. In actuality, however, eccentricity variation is more complex and is more accurately described with



Figure 9.24. Illustration of the *Milankovitch parameters*. The eccentricity is the degree the Earth's orbit departs from circular. Obliquity is the tilt of the Earth's rotation axis with respect to the plane of the ecliptic and varies between 21.5° and 24.5°. Precision is the variation in the direction of tilt at the Earth's closest approach to the Sun (perihelion). The parameter ω is the angle between the Earth's position on June 21 (summer solstice), and perihelion.

complex, and is more accurately described with periods of 123,000 yrs, 85,000 yrs, and 58,000 yrs. Similarly, variation in precession has characteristic periods of 23,000 and 18,000 yrs.

Although Emiliani suggested that δ^{18} O variations were related to variations in these orbital parameters, the first quantitative approach to the problem was that of Hayes et al. (1976). They applied Fourier analysis to the δ^{18} O curve (Fourier analysis is a mathematical tool that transforms a complex variation such as that in Figure 9.23 to the sum of a series of simple sin functions). Hayes et al. then used spectral analysis to show that much of the spectral power of the δ^{18} O curve occurred at frequencies similar to those of the Milankovitch parameters. The most elegant and convincing treatment is that of Imbrie (1985). Imbrie's treatment involved several refinements and extension of the earlier work of Hayes et al. (1976). First, he used improved values for Milankovitch frequencies. Second, he noted these Milankovitch parameters vary with time (the Earth's orbit and tilt are affected by the gravitational field of the Moon and other planets, and other astronomical events, such as bolide impacts), and the climate system's response to them might also vary over time. Thus Imbrie treated the first and second 400,000 years of Figure 9.23 separately.

Imbrie observed that climate does not respond instantaneously to forcing. For example, maximum temperatures are not reached in Ithaca, New York until late July, 4 weeks after the maximum insolation, which occurs on June 21. Thus there is a *phase* lag between the forcing function (insolation) and climatic response (temperature). Imbrie (1985) constructed a model for response of global climate (as measured by the δ^{18} O curve) in which each of the 6 Milankovitch forcing functions was associated with a different gain and phase. The values of gain and phase for each parameter were found statistically by minimizing the residuals of the power spectrum of the δ^{18} O curve in Figure 9.23. The resulting "gain and phase model" is shown in comparison with the data for the past 400,000 years and the next 25,000 years in Figure 9.25. The model has a correlation coefficient, *r*, of 0.88 with the data. Thus about *r*², or 77%, of the variation in δ^{18} O, and therefore presumably in ice volume, can be explained by Imbrie's Milankovitch model. The correlation for the period 400,000–782,000 yrs is somewhat poorer, around 0.80, but nevertheless impressive.

CHAPTER 9: STABLE ISOTOPES



Figure 9.25. Gain and phase model of Imbrie relating variations in eccentricity, tilt, and precession to the oxygen isotope curve. Top shows the variation in these parameters over the past 400,000 and next 25,000 years. Bottom shows the sum of these functions with appropriated gains and phases applied and compares them with the observed data. After Imbrie (1985).

Since variations in the Earth's orbital parameters do not affect the average total annual insolation the Earth receives, but only its pattern in space and time, one might ask how this could cause glaciation. The key factor seems to be the insolation received during summer at high northern latitudes. This is, of course, the area where large continental ice sheets develop. The southern hemisphere, except for Antarctica, is largely ocean, and therefore not subject to glaciation. Glaciers apparently develop when summers are not warm enough to melt the winter's accumulation of snow in high northern latitudes.

Nevertheless, even at a given latitude the total variation in insolation is small, and not enough by itself to cause the climatic variations observed. Apparently, there are feedback mechanisms at work that serve to amplify the fundamental Milankovitch forcing function. One of these feedback mechanisms was identified by Agassiz, and that is ice albedo, or reflectance. Snow and ice reflect much of the incoming sunlight back into space. Thus as glaciers advance, they will cause further cooling. Any additional accumulation of ice in Antarctica, however, does not result in increased albedo, because the continent is fully ice covered even in interglacial times. Hence the dominant role of northern hemisphere insolation in driving climate cycles. Other possible feedback mechanisms include carbon dioxide and ocean circulation. The role of atmospheric CO₂ in controlling global climate is a particularly important issue because of burning of fossil fuels has resulted in a significant increase in atmospheric CO₂ concentration over the last 150 years. We will return to this issue in a subsequent chapter.

9.6.2 The Record in Glacial Ice

Climatologists recognized early on that continental ice preserves a stratigraphic record of climate change. Some of the first ice cores recovered for the purpose of examining the climatic record and analyzed for stable isotopes were taken from Greenland in the 1960's (e.g., Camp Century Ice Core). Subsequent cores have been taken from Greenland, Antarctica, and various alpine glaciers. The most remarkable and useful of these cores has been the 2000 m core recovered by the Russians from the Vostok station in Antarctica (Jouzel, et al., 1987, 1993, 1996), and is compared with the marine δ^{18} O record in Figure 9.25. The marine record shown is a smoothed version SPECMAP record, which is a composite record based on that of Imbrie et al. (1984), but with further modification of the chronology. The core provides a 240,000 year record of δD and $\delta^{18}O_{icer}$ as well as CO_2 , and $\delta^{18}O_{O_2}$ in bubbles (the latter, which was subsequently published and not shown in Figure 9.26, provides a measure of δ^{18} O in the atmosphere and, indirectly, the ocean), which corresponds to a full glacial cycle.

Jouzel et al. (1987) converted δD to temperature variations after subtracting the effect of changing ice volume on δD of the oceans. The conversion is based on a 6‰/°C relationship between δD and temperature in modern Antarctic snow (they found a similar relationship using circulation models). The hydrogen isotopic fractionation of water is a more sensitive function of temperature than is oxygen fractionation. Their results, taken at face value, show dramatic 10° C temperature variations between glacial and interglacial times.

CHAPTER 9: STABLE ISOTOPES



Figure 9.26. The lower curve shows smoothed δ^{18} O for marine carbonates (SPECMAP), the middle curve shows δ D of ice in the Vostok ice core, and the upper curve the temperature calculated from δ D at Vostok at the time of deposition of the ice relative to the present mean annual temperature. Vostok data is from Jouzel, et al., (1987, 1993, 1996).

Dating of the Vostok ice core was based only on an ice flow and accumulation model. Nevertheless, the overall pattern observed is in good agreement the marine δ^{18} O record, particular from 110,000 years to the present. The record of the last deglaciation is particularly similar to that of the marine δ^{18} O record, and even shows evidence of a slight return trend toward glacial conditions from 12 kyr to 11 kyr BP, which corresponds to the well-documented Younger Dryas event of the North Atlantic region. It is also very significant that spectral analysis of the Vostok isotope record shows strong peaks in variance at 41 kyr (the obliquity frequency), and at 25 kyr, which agree with the 23 kyr precessional frequency when the age errors are considered. Thus the Vostok ice core data appear to confirm the importance of Milankovitch climatic forcing. It is interesting and significant that even in this core, taken at 78° S, it is primarily insolation at 65° N that is the controlling influence. There are, however, significant differences between the Vostok record and the marine record. Some of these are probably due to inaccuracies in dating; others may reflect differences in northern and southern hemisphere response to orbital forcing.

9.6.3 Soils and Paleosols

As we found in Chapter 6, the concentration of CO_2 in soils is very much higher than in the atmosphere, reaching 1% by volume. As a result, soil water can become supersaturated with respect to carbonates. In soils where evaporation exceeds precipitation, soil carbonates form. The carbonates form

Geochemistry

CHAPTER 9: STABLE ISOTOPES

in equilibrium with soil water, and there is a strong correlation between δ^{18} O in soil carbonate and local meteoric water, though soil carbonates tend to be about 5‰ more enriched than expected from the calcite-water fractionation. There are 2 reasons why soil carbonates are heavier. First, soil water is enriched in ¹⁸O relative to meteoric water due to preferential evaporation of isotopically light water molecules. Second, rain (or snow) falling in wetter, cooler seasons in more likely to run off than that falling during warm seasons. Taking these factors into consideration, the isotopic composition of soil carbonates may be used as a paleoclimatic indicator.

Figure 9.27 shows one example of δ^{18} O in paleosol carbonates used in this way. The same Pakistani paleosol samples analyzed by Quade et al. (1989) for δ^{13} C (Figure 9.19) were also analyzed for δ^{18} O. The δ^{13} C values recorded a shift toward more positive values at 7 Ma that apparently reflect the appearance of C₄ grasslands. The δ^{18} O shows a shift to more positive values at around 8 Ma, or a million years before the δ^{13} C shift. Quade et al. interpreted this as due to an intensification of the Monsoon system at that time, an interpretation consistent with marine paleontological evidence.



Figure 9.27. δ^{18} O in paleosol carbonate nodules from the Potwar Plateau in northern Pakistan. Different symbols correspond to different, overlapping sections that were sampled. After Quade et al. (1989).

Clays, such as kaolinites, are another important constituent of soil. Lawrence and Taylor (1972) showed that during soil formation, kaolinite and montmorillonite form in approximate equilibrium with meteoric water so that their δ^{18} O values are systematically shifted by +27 ‰ relative the local meteoric water, while δ D are shifted by about 30‰. Thus kaolinites and montmorillonites define a line parallel to the meteoric water line (Figure 9.28), the so-called *kaolinite line*. From this observation, Sheppard et al. (1969) and Lawrence and Taylor (1972) reasoned that one should be able to deduce the isotopic composition of rain at the time ancient kaolinites formed from their δ D values. Since the isotopic composition of precipitation is climate dependent, ancient kaolinites provide another continental paleoclimatic record.

There is some question as to the value of hydrogen isotopes in clays for paleoclimatic studies. Lawrence and Meaux (1993) concluded that most ancient kaolinites have exchanged hydrogen subsequent



Figure 9.28. Relationship between δD and $\delta^{18}O$ in modern meteoric water and kaolinites. Kaolinites are enriched in ¹⁸O by about 27‰ and ²H by about 30‰. After Lawrence and Taylor (1972).

to their formation, and therefore are not a good paleoclimatic indicator. Others argue that clays to do not generally exchange hydrogen unless recrystallization occurs or the clay experiences temperatures in excess of 60-80° C for extended periods (A. Gilg, pers. comm., 1995).

Lawrence and Meaux (1993) concluded, however, that oxygen in kaolinite does preserve the original δ^{18} O, and that can, with some caution, be used as a paleoclimatic indicator. Figure 9.29 compares the δ^{18} O of ancient Cretaceous North American kaolinites with the isotopic composition of modern precipitation. If the Cretaceous climate were the same as the present one, the kaolinites should be systematically 27‰ heavier than modern precipitation. For the southeastern US, this is

Geochemistry

W. M. White

CHAPTER 9: STABLE ISOTOPES

approximately true, but the difference is generally less than 27‰ for other kaolinites, and the difference decreases northward. This indicates these kaolinites formed in a warmer environment than the present one. Overall, the picture provided by Cretaceous kaolinites confirm what has otherwise be deduced about Cretaceous climate: the Cretaceous climate was generally warmer, and the equator to pole temperature gradient was lower. A drawback with this approach, however, is the lack of good ages on paleosols, which have proven to be very difficult to date precisely. The ages of many paleosols used by Lawrence and Meaux (1993) were simply inferred from stratigraphy.

9.7 Hydrothermal Systems and Ore Deposits

When large igneous bodies are intruded into high levels of the crust, they heat the surrounding rock and the water in the cracks and pore in this rock, setting up convection systems. The water in these hydrothermal systems reacts with the hot rock and undergoes isotopic exchange; the net result is that both the water and the rock change their isotopic compositions.

One of the first of many important contributions of stable isotope geochemistry to understanding hydrothermal systems was the demonstration by Harmon Craig (another student of Harold Urey) that water in these systems was meteoric, not magmatic (Craig, 1963). The argument is based upon the data shown in Figure 9.30. For each geothermal system, the δD of the "chloride" type geothermal waters is the same as the local precipitation and groundwater, but the δ^{18} O is shifted to higher values. The shift in δ^{18} O results from high temperature ($\leq 300^{\circ}$ C) reaction of the local meteoric water with hot rock. However, because the rocks contain virtually no hydrogen, there is little change in the hydrogen isotopic composition of the water. If the water involved in these systems were magmatic, it would not have the same isotopic composition as local meteoric water (it is possible that these systems contain a few percent magmatic water).

Acidic, sulfur-rich waters from hydrothermal systems can have δD that is different from local meteoric water. This shift occurs when hydrogen isotopes are fractionated during boiling of geo-



Figure 9.29. δ^{18} O in Cretaceous kaolinites from North American (black) compared with contours of δ^{18} O (in red) of present-day meteoric water. After Lawrence and Meaux (1993).



Figure 9.30. δD and $\delta^{18}O$ in meteoric hydrothermal systems. Closed circles show the composition of meteoric water in Yellowstone, Steamboat Springs, Mt. Lassen, Iceland, Larderello, and The Geysers, and open circles show the isotopic composition of chloride-type geothermal waters at those locations. Open triangles show the location of acidic, sulfiderich geothermal waters at those locations. Solid lines connect the meteoric and chloride waters, dashed lines connect the meteoric and acidic waters. The "Meteoric Water Line" is the correlation between δD and $\delta^{18}O$ observed in precipitation in Figure 9.10.

thermal waters. The steam produced is enriched in sulfide. The steam mixes with cooler meteoric water, condenses, and the sulfide is oxidized to sulfate, resulting in their acidic character. The mixing

Geochemistry

CHAPTER 9: STABLE ISOTOPES

lines observed reflect mixing of the steam with meteoric water as well as the fractionation during boiling.

Estimating temperatures at which ancient hydrothermal systems operated is a fairly straight–forward application of isotope geothermometry, which we have already discussed. If we can measure the oxygen (or carbon or sulfur) isotopic composition of any two phases that were in equilibrium, and if we know the fractionation factor as a function of temperature for those phases, we can estimate the temperature of equilibration. We will focus now on water-rock ratios, which may also be estimated using stable isotope ratios.

9.7.1 WATER-ROCK RATIOS

For a closed hydrothermal system, we can write two fundamental equations. The first simply describes isotopic equilibrium between water and rock:

$$\Delta = \delta_w^f - \delta_r^f \qquad 9.59$$

where we use the subscript w to indicate water, and r to indicate rock. The superscript f indicates the final value. So 9.59 just says that the difference between the final isotopic composition of water and rock is equal to the fractionation factor (we implicitly assume equilibrium). The second equation is the mass balance for a closed system:

$$c_w W \delta_w^i + c_r R \delta_r^i = c_w W \delta_w^f + c_r R \delta_r^f$$
9.60

where *c* indicates concentration (we assume concentrations do not change, which is valid for oxygen, but perhaps not valid for other elements), *W* indicates the mass of water involved, *R* the mass of rock involved, the superscript *i* indicates the initial value and *f* again denotes the final isotope ratio. This equation just says the amount of an isotope present before reaction must be the same as after reaction. We can combine these equations to produce the following expression for the water-rock ratio:

$$\frac{W}{R} = \frac{\delta_r^f - \delta_r^i}{\delta_w^f - \delta_r^i} \frac{c_r}{c_w} \qquad 9.61$$

The initial δ^{18} O can generally be inferred from unaltered samples and from the δ D– δ^{18} O meteoric water line and the final isotopic composition of the rock can be measured. The fractionation factor can be estimated if we know the temperature and the phases in the rock. For oxygen, the ratio of concentration in the rock to water will be close to 0.5 in all cases.

Equation 9.61 is not very geologically realistic because it applies to a closed system. A completely open system, where water makes one pass through hot rock, would be more realistic. In this case, we might suppose that a small parcel of water, dW, passes through the system and induces an incremental change in the isotopic composition of the rock, $d\delta_r$. In this case, we can write:

$$Rc_r d\delta_r = \delta_w^i - [\Delta + \delta_r]c_w dW \qquad 9.62$$

This equation states that the mass of an isotope exchanged by the rock is equal to the mass of that isotope exchanged by the water (we have substituted $\Delta + \delta_r$ for δ_w^f). Rearranging and integrating, we have:



Figure 9.31. $\delta^{18}O_r^f$ as a function of *W*/*R* and Δ computed from equation 9.63.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

$$\frac{W}{R} = \ln\left(\frac{\delta_r^f - \delta_r^i}{-\delta_r^f + \delta_w^i - \Delta} + 1\right)\frac{c_w}{c_r}$$
9.63

Thus it is possible to deduce the water rock ratio for an open system as well as a closed one.

Oxygen isotope studies can be a valuable tool in mineral exploration. Mineralization is often (though not exclusively) associated with the region of greatest water flux, such as areas of upward moving hot water above intrusions. Such areas are likely to have the lowest values of δ^{18} O. To understand this, let's solve equ. 9.63 for δ_r^f , the final value of δ^{18} O in the rock:

$$\delta_r^f = \left(\delta_r^i - \delta_w^i + \Delta\right) e^{-\frac{wc_w}{Rc_r}} + \delta_w^i + \Delta$$
9.64

Assuming a uniform initial isotopic composition of the rocks and the water, all the terms on the r.h.s. are constants except W/R and Δ , which is a function of temperature. Thus the final values of δ^{18} O are functions of the temperature of equilibration, and an exponential function of the W/R ratio. Figure 9.31 shows $\delta^{18}O_r^f$ plotted as a function of W/R and Δ , where $\delta^{18}O_r^f$ is assumed to be +6 and $\delta^{18}O_W^1$ is assumed to be –13. In a few cases, ores apparently have precipitated by fluids derived from the magma itself. In those cases, this kind of approach does not apply.

Figure 9.32 shows an example of the δ^{18} O imprint of an ancient hydrothermal system: the Bohemia Mining District in Lane Country, Oregon, where Tertiary volcanic rocks of the western Cascades have been intruded by a series of dioritic plutons. Approximately \$1,000,000 worth of gold was removed from the region between 1870 and 1940. δ^{18} O contours form a bull's eye pattern, and the region of low δ^{18} O corresponds roughly with the area of propylitic (i.e., greenstone) alteration. Notice that this region is broader than the contact metamorphic aureole. The primary area of mineralization occurs within the δ^{18} O < 0 contour. In this area, relatively large volumes of gold-bearing hydrothermal solution cooled, perhaps due to mixing with groundwater, and precipitated gold. This is an excellent example of the

value of oxygen isotope studies to mineral exploration. Similar bull's eye patterns are found around many other hydrothermal ore deposits.

9.7.2 Sulfur Isotopes and Ore Deposits

A substantial fraction of base and noble metal ores are sulfides. These have formed in a great variety of environments and under a great variety of conditions. Sulfur isotope studies have been very valuable in sorting out the genesis of these deposits. Of the various stable isotope systems we will consider here, sulfur isotopes are undoubtedly the most complex. This complexity arises in part because there are four common valence states in which sulfur can occur in the Earth, +6 (e.g.,



Figure 9.32. δ^{18} O variations in the Bohemia mining district, Oregon. Note the bull's eye pattern of the δ^{18} O contours. After Taylor (1974).
Geochemistry

CHAPTER 9: STABLE ISOTOPES

 $BaSO_4$), +4 (e.g., SO_2), 0 (e.g., S), -1 (e.g., FeS_2) and -2 (H_2S). Significant equilibrium fractionations occur between these valence states. Each of these valence states forms a variety of compounds, and fractionations can occur between these as well (e.g., Figure 9.7). Finally, sulfur is important in biological processes and fractionations in biologically mediated oxidations and reductions are often different from fractionations in the abiological equivalents.

There are two major reservoirs of sulfur on the Earth that have uniform sulfur isotopic compositions: the mantle, which has $\delta^{34}S$ of ~0 and in which sulfur is primarily present in reduced form, and seawater, which has $\delta^{34}S$ of +20* and in which sulfur is present as SO_4^{2-} . Sulfur in sedimentary, metamorphic, and igneous rocks of the continental crust may have $\delta^{34}S$ that is both greater and smaller than these values (Figure 9.33). All of these can be sources of sulfide in ores, and further fractionation may occur during transport and deposition of sulfides. Thus the sulfur isotope geochemistry of sulfide ores is remarkably complex.

At magmatic temperatures, reactions generally occur rapidly and most systems appear to be close to equilibrium. Below 200°C, however, sulfur isotopic equilibration is slow even on geologic time scales, hence equilibration is rare and kinetic effects often dominate. Isotopic equilibration between two sulfide species or between two sulfate species is achieved more readily than between sulfate and sulfide species. Sulfate-sulfide reaction rates have been shown to depend on pH (reaction is more rapid at low pH). In addition, equilibration is much more rapid when sulfur species of intermediate valence species are present. Presumably, this is because reaction rates between species of adjacent valance states (e.g., sulfate and sulfite) are rapid, but reaction rates between widely differing valence states (e.g., sulfate and sulfide) are much slower.

Mississippi Valley type deposits provide an interesting example of how sulfur isotopes can be used to understand ore genesis. These are carbonate-hosted lead and zinc sulfides formed under relatively low temperature conditions. Figure 9.34 shows the sulfur isotope ratios of some examples. They can be subdivided into Zn-rich and Pb-rich classes. The Pb-rich and most of the Zn rich deposits were formed between 70 and 120° C, while some of the Zn-rich deposits, such as those of the Upper Mississippi Valley, were formed at temperatures up to 200° C, though most were formed in the range of 90° to 150° C.

Figure 9.35 illustrates a generalized model for the genesis of Mississippi Valley type deposits. In most instances, metals and sulfur appear to have been derived distant from sedimentary units, perhaps particularly from evaporites, by hot, deep-circulating formation water. In North America, most of these seem to have formed during or shortly after the late Paleozoic Appalachian-Ouchita-Marathon Orogeny. The direct cause of precipitation of the ore probably may have differed



Figure 9.33. $\delta^{34}S_{CDT}$ in various geologic materials (after Hoefs, 1987).



Figure 9.34. Sulfur isotope ratios in some Mississippi Valley-type Pb and Zn deposits. After Ohmoto and Rye (1979).

^{*}This is the modern value for seawater. As we shall see in a subsequent chapter, though the sulfur isotopic composition of seawater is uniform at any time, this value has varied with time.

CHAPTER 9: STABLE ISOTOPES



Figure 9.35. Cartoon illustrating the essential features of the genesis of Mississippi Valley sulfide ores. After Ohmoto (1986).

in the different localities. Oxygen and hydrogen isotope data suggest that mixing of the hot saline fluids with cooler low salinity ground water was probably the immediate cause of metal precipitation in some instances. In others, such as Pine-Point, local reduction of sulfate in the fluids to sulfide may have cause precipitation (Ohmoto, 1986). Local production of sulfide shortly before ore deposition may help to account for the lack of isotopic equilibrium in this deposit, since time is an element in the attainment of isotopic equilibrium. In other cases, mixing of separate S-bearing and metal-bearing fluids from different aquifers may have been the cause.

The sulfur isotope data suggest there are a number of possible sources of sulfur. In many cases, sulfur was apparently derived sulfate in evaporites, in others sulfur was derived from brines produced by evaporative concentration of seawater. In many cases, the sulfate was reduced by reaction with organic matter at elevated temperature (thermochemical reduction). In other cases, it may have been bacterially reduced. Other sulfur sources include sulfide from petroleum. In at least some deposits, sulfur appears to have been derived from more than one of these sources.

Sulfur isotope ratios in the S. E. Missouri district are quite variable and δ^{34} S is correlated with Pb isotope ratios in galenas. This suggests there was more than one source of the sulfur. Based on a combined Pb and S isotopic study, Goldhaber et al. (1995) concluded that the main stage ores of the S.E. Missori district were produced by mixing of fluids from several aquifers. They concluded that Pb was derived from the Lamont Sandstone (which directly overlies basement). Fluid in this aquifer was maintained at relatively high pc by the presence of hematite, and hence had low sulfide concentrations. This allowed leaching and transport of the Pb. Isotopically heavy S was produced by dissolution of the upper Bonneterre Formation, which overlies the Lamont Sandstone. Both Pb and isotopically light sulfur migrated through the lower part of the Bonneterre Formation. PbS precipitated when the fluids were forced to mix by pinchout of the Sillivan Siltstone.

9.7.3 Mass Independent Sulfur Isotope Fractionation and the Rise of Atmospheric Oxygen

In the modern Earth, variations in sulfur isotope ratios are almost always mass dependent and $\delta^{33}S$ and ^{36}S are related to variations in $\delta^{34}S$ as

$$\delta^{33}S = 0.515 \text{ x } \delta^{34}S$$
 and $\delta^{336}S = 1.90 \text{ x } \delta^{34}S$

Geochemistry

CHAPTER 9: STABLE ISOTOPES

However, Farquhar et al. (2000) found that relationships do not hold for sulfur in many Archean sediments and ore deposits. Let's define the quantities Δ^{33} S and Δ^{36} S as:

$$\Delta^{33}S = \delta^{33}S - 0.515 \times \delta^{34}S$$
 9.65

and
$$\Delta^{36}S = \delta^{36}S - 1.90 \times \delta^{34}S$$
 9.66

Farquhar et al. found that many sulfides (primarily pyrite) in sediments and metasediments formed prior to 2500 Ma have positive Δ^{33} S and negative Δ^{36} S while hydrothermal sulfide ores and sedimentary sulfates (mainly barite) have negative Δ^{33} S and positive Δ^{36} S (Figure 9.36). During the Archean, Δ^{33} S, i.e., deviations from mass dependent fractionation, of over 3‰ occurred. Farquhar et al. found smaller deviations, up to ½ ‰ in rocks from the period 2500-2000 Ma. They found no significant deviations from mass dependent fractionation in rocks younger than 2000 Ma.

This raises the obvious question of what process or processes operated during the Archean that could produce mass independent fractionation? Why does this process no longer operate? As mentioned earlier, mass independent isotope fractionation has been observed during laboratory photolysis of SO₂ and SO using ultraviolet light. This, together with the observation that ozone and molecular oxygen absorb ultraviolet radiation, provides a possible explanation. Modern volcanic eruptions loft vast quantities of SO₂ into the atmosphere and Archean eruptions likely did as well. If the Archean atmosphere lacked O₂ and O₃, ultraviolet radiation could penetrate deeply into it and photodissociate SO₂. According to experiments, these reactions produce sulfate with negative Δ^{33} S and elemental sulfur with positive Δ^{33} S. The sulfate would dissolve in rain and ultimately find its way into the oceans. Some of this would precipitate as barite, BaSO₄. Some sulfate would be reduced in hydrothermal systems and precipitate as metal sulfides. Both these processes occur in modern oceans. The S would form particulate S₈ and also be swept out of the atmosphere by rain and ultimately incorporated into sediments, where it would react to form sedimentary sulfides. The latter also requires an absence of O₂ in the atmosphere – elemental S in the modern atmosphere would be quickly oxidized.



Figure 9.36. Δ^{33} S in sulfur through time. During Stage 1 (prior to 2.45 Ga), extensive mass independent fractionation occurred as evidenced by Δ^{33} S up to 3.5‰, indicating the lack of atmospheric oxygen. During Stage 2 (2.45 to 2.0 Ga), limited amounts of atmospheric oxygen (up to 1% of present levels) absorbed most UV radiation and limited mass independent fractionation to <0.5‰. High levels of atmospheric oxygen in Stage 3 (since 2.0 Ga) effectively eliminate UV radiation and mass independent fractionation. From Farquhar and Wing (2003).

Geochemistry

CHAPTER 9: STABLE ISOTOPES

Mass independent fractionation in Archean sulfur thus provides evidence that the early atmosphere lacked free oxygen. This is entirely consistent with other evidence, such as the oxidation state of paleosols and detrital minerals, that the atmosphere did not contain significant amounts of O_2 until the early Proterozoic, about 2 – 2.3 billion years ago. Farquhar and Wing (2003) and Kasting and Catling (2003) provide excellent reviews on this topic.

9.8 Stable Isotopes in the Mantle and Magmatic Systems 9.8.1 Stable Isotopic Composition of the Mantle

The mantle is the largest reservoir for oxygen, and may be for H, C, N, and S as well. Thus we need to know the isotopic composition of the mantle to know the isotopic composition of the Earth. Variations in stable isotope ratios in mantle and mantle-derived materials also provide important insights on mantle and igneous processes.

9.8.1.1 OXYGEN

Assessing the oxygen isotopic composition of the mantle, and particularly the degree to which its oxygen isotope composition might vary, has proved to be more difficult than expected. One approach has been to use basalts as samples of mantle, as is done for radiogenic isotopes. Isotope fractionation occurring during partial melting are small, so the oxygen isotopic composition of basalt should the similar to that of its mantle source. However, assimilation of crustal rocks by magmas and oxygen isotope exchange during weathering complicate the situation. An alternative is to use direct mantle samples such as xenoliths occasionally found in basalts. However, these are considerably rarer than are basalts.

Figure 9.37 shows the oxygen isotope composition of olivines and clinopyroxenes in 76 peridotite xenoliths analyzed by Mattey et al. (1994) using the laser fluorination technique. The total range of values observed is only about twice that expected from analytical error alone, suggesting the mantle is fairly homogeneous in its isotopic composition. The difference between co-existing olivines and clinopyroxenes



Figure 9.37. Oxygen isotope ratios in olivines and clinopyroxenes from mantle peridotite xenoliths. Data from Mattey et al. (1994).

Geochemistry

CHAPTER 9: STABLE ISOTOPES



averages about 0.5 per mil, which is consistent with the expected fractionation between these minerals at mantle temperatures. Mattey et al. (1994) estimated the bulk composition of these samples to be about +5.5 per mil.

Figure 9.38 shows the distribution of δ^{18} O in basalts from 4 different groups. To avoid weathering problems, Harmon and Hoefs (1995) included only submarine basaltic glasses and basalts having less than 0.75% water or have erupted historically in their compilation. There are several points worth noting in these data.

MORB have a mean $\delta^{18}O_{\text{SMOW}}$ of +5.7‰, with a relatively smaller variation about this mean. Thus the depleted upper mantle appears to be a relatively homogeneous reservoir for oxygen. The homogeneity observed in MORB is consistent with that observed in mantle-derived xenoliths. The small difference between the means of the two groups (0.2‰), may well reflect fractionation occurring during partial melting.

The other groups show significantly greater heterogeneity than either MORB or peridotite xenoliths. Oceanic island basalts, which presumably sample mantle plumes, have a mean identical to that of peridotite xenoliths, but are more variable. Subduction-related basalts (i.e., island arc basalts and their continental equivalents) are shifted to more positive δ^{18} O values, as are continental intraplate volcanics.

Whether the variability in these groups reflects mantle heterogeneity, the melting process, assimilation, or merely weathering remains unclear. Hawaii is the one locality where high quality data exists for both basalts and xenoliths and comparisons can be made. Both Hawaiian basalts and olivines in Hawaiian xenoliths (Figure 9.37) have lower δ^{18} O than their equivalents from elsewhere. This suggests the possibility, at least, that part of the variability in oceanic basalts reflects mantle heterogeneity. However, additional studies will be required to establish this with certainty.

9.8.1.2 Hydrogen

Figure 9.38. δ^{18} O in young, fresh basalts. Dashed line is at the mean of MORB (+5.7). After Harmon and Hoefs (1995).

Estimating the isotopic composition of mantle hydrogen, carbon, nitrogen, and sulfur is even more problematic. These are all volatile elements and are present at low concentrations in mantle materials. They partitioning partially or entirely into the gas phase of magmas upon their eruption. This gas phase is lost, except in deep submarine eruptions. Furthermore, C, N, and S have several oxidation

states, and isotopic fractionations occur between the various compounds these elements form (e.g., CO_2 , CO_1 , CO_2 , CO_4 , H_2 , H_2O_2 , H_2S_2 , SO_2). This presents two problems. First, significant fractionations can occur during degassing, even at magmatic temperatures. Second, because of loss of the gas phase, the concentrations of these elements are low. Among other problems, this means their isotope ratios are subject to disturbance by contamination. Thus the isotopic compositions of these elements in igneous rocks do not necessarily reflect those of the magma or its mantle source.

Hydrogen, which is primarily present as water, but also as $H_2 H_2S$, and CH_4 , can be lost from magmas during degassing. However, basalts erupted beneath a kilometer or more of ocean retain most of their dissolved water. Thus mid-ocean ridge basalts and basalts erupted on seamounts are important sources of information of the abundance and isotopic composition of hydrogen in the mantle.

As Figure 9.39 indicates, MORB have a mean δD_{SMOW} of about -67.5‰ and a standard deviation of ±14‰. How much of this variability reflects fractionation during degassing and contamination is un-

Geochemistry

W. M. White

CHAPTER 9: STABLE ISOTOPES

clear. Kyser (1986) has argued that mantle hydrogen is isotopically homogeneous with δD_{SMOW} of -80‰. The generally heavier isotopic composition of MORB, he argues, reflects H₂O loss and other processes. Others, for example, Poreda et al. (1986), Chaussidon et al. (1991) have observed correlations between δD and Sr and Nd isotope ratios and have argued that these provide clear evidence of that mantle hydrogen is isotopically heterogeneous. Chausidon et al. (1991) have suggested a δD value for the depleted upper mantle of about -55‰.

The first attempt to assess the hydrogen isotopic composition of the mantle materials was that of Sheppard and Epstein (1970), who analyzed hydrous minerals in xenoliths and concluded δD was variable in the mantle. Since then, a number of additional studies of these materials have been carried out. As Figure 9.39 shows, phlogopites (KMg₃AlSi₃O₁₀(OH)₂) have δD that is generally similar to that of MORB, though heavier values also occur. Amphiboles have much more variable δD and have heavier hydrogen on average. Part of this difference probably reflects equilibrium fractiona-



Figure 9.39. δD in MORB and in mantle phlogopites and amphiboles. The MORB and phlogopite data suggest the mantle has δD_{SMOW} of about –60 to –90.

tion. The fractionation between water and phlogopite is close to 0% in the temperature range 800° - 1000° C, whereas the fractionation between water and amphibole is about -15%. However, equilibrium fractionation alone cannot explain either the variability of amphiboles or the difference between the mean δD of phlogopites and amphiboles. Complex processes that might include Rayleigh distillation may be involved in the formation of mantle amphiboles. This would be consistent with the more variable water content of amphiboles compared to phlogopites. There are also clear regional variations in δD in xenoliths that argue for large-scale heterogeneity. For example, Deloule et al. (1991) found that δD in amphiboles from xenoliths in basalts in the Massif Central of France have systematically high δD (-59 to -28‰) while those from Hawaii can be very low (-125‰). At the opposite extreme, Deloule et al. (1991) also observed isotopic heterogeneity within single crystals, which they attribute to incomplete equilibration with magmas or fluids.

9.8.1.3 Carbon

Most carbon in basalts is in the form of CO_2 , which has limited solubility in basaltic liquids. As a result, basalts begin to exsolve CO_2 before they erupt. Thus virtually every basalt, including those erupted at mid-ocean ridges, has lost some carbon, and subareal basalts have lost virtually all carbon (as well as most other volatiles). Therefore only basalts erupted beneath several km of water provide useful samples of mantle carbon, so the basaltic data set is essentially restricted to MORB and samples recovered from seamounts and the submarine part of Kilauea's East Rift Zone. The question of the isotopic composition of mantle carbon is further complicated by fractionation and contamination. There is a roughly 4‰ fractionation between CO_2 dissolved in basaltic melts and the gas phase, with ¹³C enriched in the gas phase. Carbonatites and diamonds provide alternative, and generally superior, samples of mantle carbon, but their occurrence is extremely restricted.

MORB have a mean δ^{13} C of -6.5% (Figure 9.40), but the most CO₂-rich MORB samples have δ^{13} C of about -4%. Since they are the least degassed, they presumably best represent the isotopic composition of the depleted mantle (Javoy and Pineau, 1991). Carbonatites also have a mean δ^{13} C close to -4%. Oceanic island basalts erupted under sufficient water depth to preserve some CO₂ in the vesicles appear to have similar isotopic compositions. Gases released in subduction zone volcanos and back-arc basin basalts (Figure 9.38), which erupt behind subduction zones and are often geochemically similar to island arc basalts, have carbon that can be distinctly lighter (lower δ^{13} C), though most δ^{13} C values are in

CHAPTER 9: STABLE ISOTOPES



Figure 9.40. Carbon isotope ratios in mantle (red) and mantle-derived materials (gray). After Mattey (1987).

the range of -2 to -4%, comparable to the most gas-rich MORB. Carbon in mantle diopsides (clinopyroxenes) appears to be isotopically lighter on average, and the significance of this is unclear.

Diamonds show a large range of carbon isotopic compositions (Figure 9.40). Most diamonds have δ^{13} C within the range of -2 to -8‰, hence similar to MORB. However, some diamonds have much lighter carbon. Based on the inclusions they contain, diamonds can be divided between peridotitic and eclogitic. Most peridotitic diamonds have δ^{13} C close to -5‰, while eclogitic diamonds are much more isotopically variable. Most, though not all, of the diamonds are isotopically zoned, indicating they grew in several stages.

Three hypotheses have been put forward to explain the isotopic heterogeneity in diamonds: primordial heterogeneity, fractionation effects, and recycling of organic carbon from the Earth's surface into the mantle. Primordial heterogeneity seems unlikely for a number of reasons. Among these is the absence of very negative δ^{13} C in other materials, such as MORB, and the absence of any evidence for primordial heterogeneity from the isotopic compositions of other elements. Boyd and Pillinger (1994) have argued that since diamonds are kinetically sluggish (witness their stability at the surface of the Earth, where they are thermodynamically out of equilibrium), isotopic equilibrium might not achieved

during their growth. Large fractionations might therefore occur due to kinetic effects. However, these kinetic fractionations have not been demonstrated, and fractionations of this magnitude (20‰ or so) would be surprising at mantle temperatures.

On the other hand, several lines of evidence support the idea that isotopically light carbon in some diamonds had its origin as organic carbon at the Earth's surface. First, such diamonds are primarily of eclogitic paragenesis, and ecolgite is the high pressure equivalent of basalt. Subduction of oceanic crust continuously carries large amounts of basalt into the mantle. Oxygen isotope heterogeneity observed in some eclogite xenoliths suggests these eclogites do indeed represent subducted oceanic crust. Second, the nitrogen isotopic composition of isotopically light diamonds is anomalous relative to nitrogen in other mantle materials. Finally, Farquhar et al. (2002) found mass independent fractionation (Δ^{33} S up to 0.6‰) in sulfide inclusions in eclogitic diamonds. If, as is believed the mass independent fractionation arises from UV photolysis in the atmosphere, then the sulfur in these diamonds must once have been at the surface of the Earth.

9.8.1.4 Nitrogen

The solubility of N_2 in magmas is very limited, hence of volcanic rocks, once again only submarine basalts provide useful samples of mantle N. There are both contamination and analytical problems with determining nitrogen isotope ratios in basalts, which, combined with its low abundance (generally less than a ppm), mean that accurate measurements are difficult to make. Nitrogen substitutes readily for carbon in diamonds, which can contain up to 2000 ppm N, making diamonds important samples of mantle N. Nitrogen may be present as the ammonium ion in many kinds of rocks. NH_4^+ substitutes readily for K in many minerals; consequently N concentrations in sediments and metasediments may reach several hundred ppm. In all, a fair amount of data is now available for a variety of mantle and crustal materials and these data are summarized in Figure 9.41. Measurements of $\delta^{15}N_{ATM}$ in MORB

Geochemistry

CHAPTER 9: STABLE ISOTOPES

range from about –10 to +8‰, with a mean value of about -3 to -5‰. Peridotitic diamonds show a similar range, but are a little lighter on average. Eclogitic diamonds show considerably greater scatter in $\delta^{15}N_{ATM}$, with the most common values in the range of -12 to -6‰. Interestingly, oceanic island basalts, which presumably sample mantle plumes, have distinctly more positive $\delta^{15}N_{ATM}$, with a mean in the range of +3 to +4‰. The N isotopic composition of mantle plumes thus appears to match well with that of organic matter in post-Archean sediments, metamorphic rocks, and subduction-related volcanic rocks. This observation led Marty and Dauphas (2003) to propose that nitrogen in mantle plumes is largely recycled from the surface of the Earth. They argue that because nitrogen can be bound in minerals as ammonium, it is more readily subducted and recycling into the mantle than other gases.

Figure 9.41 also illustrates an interesting shift in $\delta^{15}N$ in organic matter in sediments in the late Archean. Marty and Dauphas (2003) suggest 2 possible causes. First, the absence of atmospheric oxy-



Figure 9.41. Isotopic composition of nitrogen in rocks and minerals of the crust and mantle. Modified from Marty and Dauphas (2003).

Geochemistry

CHAPTER 9: STABLE ISOTOPES

gen in the Archean meant that there was no nitrate, and hence no dissimilatory denitrification and consequently no fractionation associated with that process. Second, chemosynthetic life may have dominated the Archean. Modern chemosynthetic bacteria from hydrothermal vent ecosystems have a lower δ^{15} N than most plants, meaning the N isotopic composition of organic matter in Archean sediments would have been lighter than in post-Archean sediments.

9.8.1.5 Sulfur

There are also relatively few sulfur isotope measurements on basalts, in part because sulfur is lost during degassing, except for those basalts erupted deeper than 1 km below sealevel. In the mantle, sulfur is probably predominantly or exclusively in the form of sulfide, but in basalts, which tend to be somewhat more oxidized, some of it may be present as SO₂ or sulfate. Equilibrium fractionation should lead to SO₂ being a few per mil lighter than sulfate. If H₂S is lost during degassing the remaining sulfur would become heavier; if SO₂ or SO₄ is lost, the remaining sulfur would become lighter. Total sulfur in MORB has $\delta^{34}S_{CDT}$ in the range of +1.3 to -1%, with most values in the range 0 to +1%. Sakai et al. (1984) found that sulfate in MORB, which constitutes 10-20% of total sulfur, was 3.5 to 9‰ heavier than sulfide. Basalts from Kilauea's East Rift Zone have a very restricted range of δ^{34} S of +0.5 to +0.8 (Sakai, et al., 1984).

Chaussidon et al. (1989) analyzed sulfides present as inclusions in minerals, both in basalts and in xenoliths, and found a wide range of $\delta^{34}S$ (–5 to +8‰). Low Ni sulfides in oceanic island basalts, kimberlites, and pyroxenites had more variable $\delta^{34}S$ that sulfides in peridotites and peridotite minerals. They argued there is a fractionation of +3‰ between sulfide liquid and sulfide dissolved in silicate melt. Carbonatites have $\delta^{34}S$ between +1 and -3‰ (Hoefs, 1986; Kyser, 1986). Overall, it appears the mantle has a mean $\delta^{34}S$ in the range of 0 to +1‰, which is very similar to meteorites, which average about +0.1‰.

Chaussidon, et al. (1987) found that sulfide inclusions in diamonds of peridotitic paragenesis (δ^{13} C ~ -4‰) had δ^{34} S of about +1‰ while eclogitic diamonds had higher, and much more variable δ^{34} S (+2 to +10‰). Eldridge et al. (1991) found that δ^{34} S in diamond inclusions was related to the Ni content of the sulfide. High Ni sulfide inclusions, which they argued were of peridotitic paragenesis, had δ^{34} S between +4‰ and -4‰. Low Ni sulfides, which are presumably of eclogitic paragenesis, had much more variable δ^{34} S (+14‰ to -10). These results are consistent with the idea that eclogitic diamonds are derived from subducted crustal material.

9.8.2 Stable Isotopes in Crystallizing Magmas

The variation in stable isotope composition produced by crystallization of a magma depends on the manner in which crystallization proceeds. The simplest, and most unlikely, case is *equilibrium* crystallization. In this situation, the crystallizing minerals remain in isotopic equilibrium with the melt until crystallization is complete. At any stage during crystallization, the isotopic composition of a mineral and the melt will be related by the fractionation factor, α . Upon complete crystallization, the rock will have precisely the same isotopic composition as the melt initially had. At any time during the crystallization, the isotope ratio in the remaining melt will be related to the original isotope ratio as:

$$\frac{R_{\ell}}{R_{\alpha}} = \frac{1}{f + \alpha(1 - f)}; \ \alpha \equiv \frac{R_s}{R_{\ell}}$$
9.67

where R_l is the ratio in the liquid, R_s is the isotope ratio of the solid, R_0 is the isotope ratio of the original magma, *f* is the fraction of melt remaining. This equation is readily derived from mass balance, the definition of α , and the assumption that the O concentration in the magma is equal to that in the crystals; an assumption valid to within about 10%. It is more convenient to express 9.67, in terms of δ :

$$\Delta = \delta_{melt} - \delta_0 \cong \left[\frac{1}{f + \alpha(1 - f)}\right] * 1000$$
9.68

where δ_{melt} is the value of the magma after a fraction *f*-1 has crystallized and δ_{o} is the value of the original magma. For silicates, α is not likely to much less than 0.998 (i.e., $\Delta = \delta^{18}O_{\text{melt}} > = -2$). For α

CHAPTER 9: STABLE ISOTOPES

= 0.999, even after 99% crystallization, the isotope ratio in the remaining melt will change by only 1 per mil.

The treatment of *fractional crystallization* is analogous to Rayleigh distillation. Indeed, it is governed by the same equation:

The key to the operation of either of these processes is that the product of the reaction (vapor in the case of distillation, crystals in the case of crystallization) is only instantaneously in equilibrium with the original phase. Once it is produced, it is removed from further opportunity to attain equilibrium with original phase. This process is more efficient at producing isotopic variations in igneous rocks, but its effect remains





Figure 9.42. Plot of δ^{18} O versus fraction of magma solidified during Rayleigh fractionation, assuming the original δ^{18} O of the magma was +6. After Taylor and Sheppard (1986).

limited because α is generally not greatly different from 1. Figure 9.42 shows calculated change in the isotopic composition of melt undergoing fractional crystallization for various values of Δ (\approx 1000(α –1)). In reality, Δ will change during crystallization because of (1) changes in temperature (2) changes in the minerals crystallizing, and (3) changes in the liquid composition. The changes will generally mean that the effective Δ will increase as crystallization proceeds. We would expect the greatest isotopic fractionation in melts crystallizing non-silicates such as magnetite and melts crystallizing at low temperature, such as rhyolites, and the least fractionation for melts crystallizing at highest temperature, such as basalts.

Figure 9.43 shows observed δ^{18} O as a function of temperature in two suites: one from a propagating



Figure 9.43. δ^{18} O as a function of SiO₂ in a tholeiitic suite from the Galapagos Spreading Center (GSC) (Muehlenbachs and Byerly, 1982) and an alkaline suite from Ascension Island (Sheppard and Harris, 1985). Dashed line shows model calculation for the Ascension suite.

rift on the Galapagos Spreading Center, the other from the island of Ascension. There is a net change in δ^{18} O between the most and least differentiated rocks in the Galapagos of about 1.3‰; the change in the Ascension suite is only about 0.5‰. These, and other suites, indicate the effective Δ is generally small, on the order of 0.1 to 0.3‰. Consistent with this is the similarity of δ^{18} O in peridotites and MORB, which suggests a typical fractionation during melting of 0.2‰ or less.

We can generalize the temperature dependence of oxygen isotope fractionations by saying that at low temperature (i.e., Earth surface temperatures up to the temperature of hydrothermal systems, 300-400° C), oxygen isotope ratios are changed by chemical processes. The amount of change can be used as an indication of the nature of the process involved, and, under equilibrium conditions, of the temperature at which the process oc-

CHAPTER 9: STABLE ISOTOPES

curred. At high temperatures (temperatures of the interior of the Earth or magmatic temperatures), oxygen isotope ratios are minimally affected by chemical processes and can be used as tracers much as radiogenic isotope ratios are.

These generalizations lead to an axiom: igneous rocks whose oxygen isotopic compositions show significant variations from the primordial value (+6) must either have been affected by low temperature processes, or must contain a component that was at one time at the surface of the Earth (Taylor and Sheppard, 1986).

9.8.3 Combined Fractional Crystallization and Assimilation

Because oxygen isotope ratios of mantlederived magmas are reasonably uniform $(\pm 1\%)$ of 5.6\%) and generally different from rocks that have equilibrated with water at the surface of the Earth, oxygen isotopes are a useful tool in identifying and studying the assimilation of country rock by intruding



Figure 9.44. Variation in δ^{18} O of a magma undergoing AFC vs. amount crystallized. Initial δ^{18} O of the magma is +5.7. After Taylor (1980).

magma. We might think of this process as simple mixing between two components: magma and country rock. In reality, it is always at least a three-component problem, involving country rock, magma, and minerals crystallizing from the magma. Magmas are essentially never superheated; hence the heat required to melt and assimilate surrounding rock can only come from the latent heat of crystallization of the magma. Approximately 1 kJ/g would be required to heat rock from 150° C to 1150°C and another 300 J/g would be required to melt it. If the latent heat of crystallization is 400 J/g, crystallization of 3.25 g of magma would be required to assimilate 1 g of country rock. Since some heat will be lost by simple conduction to the surface, we can conclude that the amount of crystallization will inevitably be greater than the amount of assimilation (the limiting case where mass crystallized equals mass assimilated could occur only at great depth in the crust where the rock is at its melting point to begin with). The change in isotopic composition of a melt undergoing the combined process of assimilation and fractional crystallization (AFC) is given by:

$$\delta_m - \delta_0 = \left(\left[\delta_a - \delta_0 \right] - \Delta \times R \right) \left\{ 1 - f^{1/(R-1)} \right\}$$
9.69

where R is the mass ratio of material crystallized to material assimilated, Δ is the difference in isotope ratio between the magma and the crystals ($\delta_{maggma} - \delta_{crystals}$), f is the fraction of liquid remaining, δ^m is the δ^{18} O of the magma, δ_o is the initial δ^{18} O of the magma, and δ_a is the δ^{18} O of the material being assimilated. The assumption is made that the concentration of oxygen is the same in the crystals, magma and assimilant, which is a reasonable assumption. This equation breaks down at R = 1, but, as discussed above, R is generally significantly greater than 1. Figure 9.44 shows the variation of δ^{18} O of a magma with an initial δ^{18} O = 5.7 as crystallization and assimilation proceed.

Combining stable and radiogenic isotope ratios provides an even more powerful tool to examine assimilation. We shall consider this in Chapter 12.

9.9 ISOTOPES OF BORON AND LITHIUM

Although there are a few earlier works in the literature, there was little interest in the isotopic composition of boron and lithium until about 20 years ago. This is perhaps because both elements have low abundances in the Earth compared to the other elements we have discussed thus far. It is perhaps also because of the analytical problems: neither B nor Li form gaseous species that can be analyzed in the gas-source mass spectrometers used for analysis of other stable isotopes, and the fractionation pro-

Geochemistry

CHAPTER 9: STABLE ISOTOPES

duced in thermal ionization mass spectrometers, which are used for elements such as Sr, Nd, and Pb, would exceed the natural ones. Since the development of new analytical techniques that overcame this latter problem in the 1980's, the field of boron isotope geochemistry has developed rapidly, and a range of about 9% in the ¹¹B/¹⁰B ratio in terrestrial materials has been observed. Boron isotope geochemistry has been used to address a wide variety of geochemical problems. Most notable among these are: hydrothermal processes, the nature and origin of ore-forming fluids, the origin of evaporites and brines, pH of ancient oceans, the origin of subduction-related magmas, and geochemical evolution of the mantle.

Though both lithium and boron can occur as stoichiometric components of minerals, these minerals have limited occurrence, and these elements generally substitute for other elements in silicates. Boron is relatively abundant in seawater, with a concentration of 4.5 ppm. Lithium is somewhat less abundant, with a concentration of 0.17 ppm. In silicate rocks, the concentration of boron ranges from a few tenths of a ppm or less in fresh basalts and peridotites to several tens of ppm in clays. Lithium concentrations in these materials typically range from a few ppm to a few tens of ppm. The ¹¹B/¹⁰B is reported as per mil variations, δ^{11} B, from the NIST SRM^{*} 951 standard. The ⁷Li/⁶Li ratio is reported as per mil variation, δ^{7} Li, from the NIST SRM 8545 Li₂CO₃ (L-SVEC) standard (Table 9.1). Prior to 1996, Li isotope ratios were often reported as δ^{6} Li, i.e., deviations from the ⁶Li/⁷Li ratio. However, the standard used was the same, so that for variations less than about 10, δ^{7} Li $\approx \delta^{6}$ Li; at higher deviations, a more exact conversion is necessary, e.g., -38.5‰ δ^{6} Li = 40‰ δ^{7} Li. The analytical precision for most of the Li isotope now in the literature is about 1‰, but recent advances, particularly the use of multiple-collector inductively coupled plasma mass spectrometers, promises to reduce uncertainty to as little as 0.2‰.

9.9.1 BORON ISOTOPES

In nature, boron has a valence of +3 is almost always bound to oxygen or hydroxyl groups in either trigonal (e.g., BO_3) or tetrahedral (e.g., $B(OH)_4^-$) coordination (the only exception is boron bound to fluorine, e.g., BF_3). Since the bond strengths and vibrational frequencies of trigonal and tetrahedral forms differ, we can expect that isotopic fractionation will occur between these two forms. This is confirmed by experiments which show a roughly 20‰ fractionation between B(OH)₃ and B(OH)₄⁻, with ¹¹B preferentially found in the B(OH)₃ form.

In natural aqueous solutions boron occurs as both boric acid, $B(OH)_3$, and the borate ion, $B(OH)_4^-$, the dominant species being determined by pH. At pH of around 9 and above, $B(OH)_4^-$, dominates, at lower pH $B(OH)_3$ dominates. In seawater which has a pH in the range of 7.6 to 8.1, about 80-90% of boron will be in the $B(OH)_3$ form. Most fresh waters are a little more acidic so $B(OH)_3$ will be more dominant; only in highly alkaline solutions, such as saline lakes, will $B(OH)_4^-$ be dominant. The most common boron mineral in the crust is tourmaline (Na(Mg,Fe,Li,Al)_3Si_6O_{18} (BO_3)_3(OH,F)_4), in which boron is present in BO₃ groups. In clays, boron appears to occur primarily as $B(OH)_4^-$, most likely substituting for silica in tetrahedral layers. The coordination of boron in common igneous minerals is uncertain, possibly substituting for Si in tetrahedral sites. As we noted in Chapter 7, boron is an incompatible element in igneous rocks and is very fluid mobile. It is also readily adsorbed onto the surfaces of clays. There is an isotopic fractionation between dissolved and adsorbed B of -20 to -30‰ (i.e, adsorbed B is ¹¹B poor), depending on pH and temperature (Palmer et al., 1987).

Figure 9.45 illustrates the variation in B isotopic composition in a variety of geologic materials. Spivack and Edmond (1987) found the δ^{11} B of seawater to be +39.5‰, and was uniform within analytical error (±0.25‰). Fresh mid-ocean ridge basalts typically have δ^{11} B of about -4‰. Oceanic island basalts (OIB) have slightly lighter δ^{11} B (e.g., Chaussidon and Jambon, 1994). The average B isotopic composition of the continental crust probably lies between -13‰ and -8‰ (Chaussidon and Abarède, 1992).

Perhaps the most remarkable aspect of B isotope geochemistry is the very large fractionation of B isotopes between the oceans and the silicate Earth. It was recognized very early that this difference re-

^{*} NIST stands for the U.S. National Institute of Standards and Technology and was formerly known as the National Bureau of Standards (NBS). SRM stands for standard reference material.

CHAPTER 9: STABLE ISOTOPES

flected the fractionation that occurred during adsorption of boron on clays (e.g., Schwarcz, et al., 1969). However, as we noted above, this fractionation is only about 30% or less, whereas the difference between the continental crust and seawater is close to 50%. Furthermore, the net effect of hydrothermal exchange between the oceanic crust and seawater is to decrease the δ^{11} B of seawater (Spivack and Edmond, 1977). The discrepancy reflects the fact that the ocean is not a simple equilibrium system, but rather a kinetically controlled open one. Since all processes operating in the ocean appear to preferentially remove ¹⁰B from the ocean, seawater is driven to an extremely ¹¹B-rich composition. In addition, it is possible that additional fractionations of boron isotopes may occur during diagenesis. Ishikawa and Nakamura (1993) noted that ancient limestones and cherts have more negative δ^{11} B than their modern equivalents, calcareous and siliceous oozes, and suggested that further fractionation occurred during diagenesis.

Spivack and Edmond (1987) investigated the exchange of boron between seawater and oceanic crust. Boron is readily incorporated into the alteration products of basalt, so that even slightly altered basalts show a dramatic increase in B concentration and an increase in δ^{11} B, with altered oceanic crust having δ^{11} B in the



Figure 9.45. Boron isotopic composition in crystalline rocks (MORB: mid-ocean ridge basalts; OIB: oceanic island basalts; BABB: back-arc basin basalts, IAV: island arc volcanics), sediments, groundwater, freshwater, salt lakes, seawater, and mid-ocean ridge hydrothermal solutions.

range of 0 to +25‰. Smith et al. (1995) estimated that average altered oceanic crust contains 5 ppm B and δ^{11} B of +3.4‰. During high temperature interaction between seawater and oceanic crust, Spivack and Edmond (1987) concluded that boron was quantitatively extracted from the oceanic crust by hydro-thermal fluids. The isotopic composition of these fluids is slightly lower than that of seawater. They inferred that the B in these fluids is a simple mixture of seawater- and basalt-derived B and that little or no isotopic fractionation was involved. Analysis of hydrothermal altered basalts recovered by the Ocean Drilling Project generally confirm these inferences, as they are boron-poor and have δ^{11} B close to 0 (Ishikawa and Nakamura, 1992).

The more positive δ^{11} B values of island arc volcanics (IAV) relative to MORB must in part reflect the incorporation of subducted marine sediment and altered oceanic crust into the sources of island arc magmas (e.g., Palmer, 1991). This idea is consistent with a host of other data, which we will discuss in Chapter 12. There is, however, some debate as to the extent the more positive δ^{11} B values in IAV might also reflect assimilation of sediment or altered oceanic crust during magma ascent (e.g., Chaussidon and Jambon, 1994).

The differences in δ^{11} B between oceanic island basalts (OIB) and MORB is perhaps more problematic. Though no experimental or theoretical studies have been carried out, it seems unlikely that significant fractionation of boron isotopes will occur during melting of the mantle, both because the temperatures are high, and because the atomic environment of B in silicate melts is probably similar to that in silicate

Geochemistry

CHAPTER 9: STABLE ISOTOPES

solids. Thus, as we found was the case for O isotopes, B isotope fractionation probably occurs only at the surface of the Earth, and the difference between OIB and MORB must somehow reflect surface processes. Chaussidon and Marty (1995) argued that the boron isotopic composition of the mantle is that of OIB (-10‰) and that the higher δ^{11} B of MORB reflects contamination of MORB magmas by altered oceanic crust. This seems unlikely for several reasons. First, although there are still relatively few data available, MORB appear to be relatively homogeneous in their boron isotopic composition. This means assimilation would have to be systematic and pervasive and that all MORB magmas would have to assimilate roughly the same amount of material. Both of these seem highly improbable. Second, there is little or no other evidence for assimilation of oceanic crust by MORB magmas. Third, oceanic island basalts have an opportunity to assimilate not only altered oceanic crust, but also overlying sediment. Yet, according to the Chaussidon and Marty (1995) hypothesis, they are not systematically contaminated. Although they are not *systematically* contaminated, there is evidence of occasional assimilation of oceanic crust and/or sediment by oceanic island basalt magmas from both B and Os isotope geochemistry. This may explain some of the lower δ^{11} B values in OIB seen in Figure 9.45 (Chaussidon and Jambon, 1994).

The alternative explanation for the lower $\delta^{11}B$ in OIB is that they contain a component of material recycled into the mantle, through subduction, from the surface of the Earth. The idea that mantle plumes, and the OIB magmas they produce, contain material recycled from the surface of the Earth has been suggested on other grounds (Hofmann and White, 1982), and we shall discuss it further in Chapter 11. You (1994) points out that significant fractionation of B isotopes will occur during sediment dewatering at moderate temperatures during the initial stages of subduction. The fluid produced will be enriched in ¹¹B, leaving the sediment depleted in ¹¹B. Thus the effect of subduction zone processes will be to lower the $\delta^{11}B$ of oceanic crust and sediment carried into the deep mantle. Deciding between this view and that of Chaussidon and Marty (1995) will require further research.

One of the more interesting applications of boron isotopes has been determining the paleo-*pH* of the oceans. Boron is readily incorporated into carbonates, with modern marine carbonates having B concentrations in the range of 15-60 ppm. In modern foraminfera, δ^{11} B is roughly 20‰ lighter than the seawater in which they grow. This fractionation seems to result from the kinetics of B coprecipitation in CaCO₃, in which incorporation of B in carbonate is preceded by surface adsorption of B(OH)⁻₄ (Vengosh et al., 1991; Heming and Hanson, 1992).

We noted above that boron is present in seawater both as $B(OH)_3$, and . Since the reaction between the two may be written as:

$$B(OH)_3 + H_2O \rightleftharpoons B(OH)_4 + H^+$$
 9.70

The equilibrium constant for this reaction is:

$$-pK^{app} = \ln \frac{B(OH)_4}{B(OH)_3} - pH$$
9.71

The relative abundance of these two species is thus pH dependent. Furthermore, we can easily show that the isotopic composition of these two species must vary if the isotopic composition of seawater is constant. From mass balance we have:

$$\delta^{11} \mathbf{B}_{sw} = \delta^{11} \mathbf{B}_3 f + \delta^{11} \mathbf{B}_4 (1-f)$$
9.72

where *f* is the fraction of B(OH)₃ (and 1 - *f* is therefore the fraction of B(OH)₄⁻), $\delta^{11}B_3$ is the isotopic composition of B(OH)₃, and $\delta^{11}B_4$ is the isotopic composition of B(OH)₄⁻. If the isotopic composition of the two species are related by a constant fractionation factor, Δ_{3-4} , we can write 9.72 as:

$${}^{1}B_{SW} = \delta^{11}B_{3}f + \delta^{11}B_{4} - \delta^{11}B_{4}f = \delta^{11}B_{4} - \Delta_{3.4}f \qquad 9.73$$

Solving for $\delta^{11}B_4$, we have:

$$\delta^{11}B_4 = {}^{11}B_{SW} + \Delta_{3.4}f \qquad 9.74$$

Thus assuming a constant fractionation factor and isotopic composition of seawater, the δ^{11} B of the two B species will depend only on *f*, which, as we can see in equation 9.71, will depend on pH.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

If the mechanism for incorporation of B in carbonate presented above is correct, the fractionation of ${}^{11}B/{}^{10}B$ between calcite and seawater will be pH dependent. There is still some debate as to as to the exact mechanism of boron incorporation in carbonate, in particular whether only borate ion or both boric acid and borate ion can be incorporated. Regardless of the exact mechanism, the isotopic composition of boron in carbonate precipitated from seawater has been shown to be a strong function of pH (Sanyal et al., 2000), allowing in principle the reconstruction of paleo-seawater pH from carbonates. There are a some additional factors that must be considered: (1) different carbonate-secreting species may fractionate B isotopes slightly differently, perhaps because they alter the pH of their micro-environment, or perhaps because B(OH)₃ is also utilized to varying degrees, (2) the fractionation factor is temperature dependent, and (3) the B isotopic composition of seawater may vary with time. Nevertheless, if care is exercised to account for "vital" effects and variation in the isotopic composition and temperature of seawater, the B isotopic composition of marine biogenic carbonate preserved in sediment should reflect the pH of the water from which they were precipitated.



Figure 9.46. Top graph shows the variation pH of surface seawater during Tertiary time as inferred from $\delta^{11}B$ in shells of planktonic foraminifera in ODP cores. Bottom graph shows the concentration of atmospheric CO₂ calculated from seawater pH. Adapted from Pearson and Palmer (2000).

Geochemistry

CHAPTER 9: STABLE ISOTOPES

The pH of seawater, in turn, is largely controlled by the carbonate equilibrium, and depends therefore on the partial pressure of CO₂ in the atmosphere (Example 6.2 explores how pH of a solution will depend on p_{CO_2}). Thus if the pH of ancient seawater can be determined, it should be possible to estimate p_{CO_2} of the ancient atmosphere. Given the concern about the relation of future climate to future p_{CO_2} , it is obviously interesting to know how these factors related in the past.

Pearson and Palmer (2000) measured δ^{11} B in foraminiferal carbonate extracted from Ocean Drilling Program (ODP) cores and from this calculated pH based on equations 9.71 and 9.74. To minimize the effect of temperature on the fractionation factor, they restricted their study to cores from regions that were tropical at the time of deposition. To minimize vital effects, they used only 1 species of planktonic foraminifera for the Neogene period, *G. trilobus* (also known as *G. sacculifer*), which is thought to incorporate B with no vital effect. For the Paleogene, they used 6 species where the vital effect was arguably minimal. They argue that changes in the B isotopic composition of seawater should be slow to occur since the residence time of B in seawater is roughly 20 million years. Nevertheless, they account for a small variation, roughly 1.7‰, in seawater δ^{11} B over Tertiary time. The results suggest dramatically lower seawater pH and dramatically higher p_{CO_2} in the Paleogene. The apparent variation in p_{CO_2} is qualitatively consistent with what is known about Tertiary climate change – namely that a long-term cooling trend began in the early to middle Eocene. In contrast to the Paleogene, Figure 9.46 shows that the Neogene is characterized by atmospheric p_{CO_2} near or slightly below modern pre-industrial levels.

On a more limited time scale, Hönish and Hemming (2005) investigated δ^{11} B over the last two glacial cycles (0-140 and 300-420 ka). In this study, the controlled for temperature by analyzing the Mg/Ca ratio of the carbonate shells, which is know to be strongly temperature dependent. Their calculated pH values ranged from 8.11 to 8.32, which in turn correspond to a p_{CO_2} range of ~180 to ~325 ppm. These calculated p_{CO_2} values are in good agreement with CO₂ concentrations measured in bubbles in the Vostok ice core.

Nevertheless, paleo-seawater pH reconstruction based on boron isotopes remains somewhat controversial. Pagani et al. (2005) and Zeebe (2005) point out that the fractionation factor between $B(OH)_3$ and $B(OH)_4^-$ has not been experimentally determined and remains somewhat uncertain. Attempts to calculate it theoretically are hampered by uncertainties in experimental determination of vibrational frequencies, and attempts to calculate it *ab initio* from molecular orbital theory also result in uncertainties that are too large to be useful. In addition to this, there are additional uncertainties in the mechanism of incorporation of boron into carbonate and the associated fractionation factor, and uncertainties in how $\delta^{11}B$ in seawater has varied with time.

9.9.2 Li Isotopes

Terrestrial lithium isotopic variation is dominated by the strong fractionation that occurs between minerals, particularly silicates, and water. Indeed, this was first demonstrated experimentally by Urey in the 1930's. This fractionation in turn reflects the chemical behavior of Li. Li forms bonds with a high degree of covalent character. The ionic radius of Li¹⁺ is small (78 pm) and Li readily substitutes for Mg²⁺, Fe²⁺, and Al³⁺ in crystal lattices, mainly in octahedral sites. In aqueous solution, it is tetrahedrally coordinated by 4 water molecules (the solvation shell) to which it is strongly bound, judging from the high solvation energy. These differences in atomic environment, differences in binding energies, the partly covalent nature of bonds, and the low mass of Li all lead to strong fractionation of Li isotopes.

Modern study of Li isotope ratios began with the work of Chan and Edmond (1988). They found that the isotopic composition of seawater was uniform within analytical error with a δ^7 Li value of +33‰. Subsequent work suggests that δ^7 Li in seawater might vary by as much as 4‰, but the degree to which this variation reflects analytical error and inter-laboratory biases remains unclear, as the residence time of Li in the ocean (1.5-3 Ma) is much longer than the mixing time.

As is the case for boron, seawater represents one extreme of the spectrum of isotopic compositions in the Earth. During mineral-water reactions, the heavier isotope, ⁷Li, is preferential partitioned into the solution. Thus weathering on the continents results in river water being isotopically heavy, +23.5‰ on average, compared to average continental crust, which has δ^7 Li of ~0, and the suspended load of rivers which have δ^7 Li \approx +2 (Teng et al., 2004). Thus seawater is some 10 per mil heavier than average river

Geochemistry

W. M. White

CHAPTER 9: STABLE ISOTOPES

water, so additional fractionation must occur marine environment. This likely includes adsorption on particles (although Li is less prone to absorption than most other metals), authigenic clay formation, and low temperature alteration of oceanic crust. Chan et al. (1992) estimated the fractionation factor for clay-solution exchange at -19‰ and the fractionation factor for low-temperature alteration of basalt. Reported δ^7 Li values in marine non-carbonate sediments range from -1 to +15‰, but Chan and Frey (2003) suggested, based on new data and reanalysis of older samples that the range is only +1 to +6%. Marine carbonate sediments, which tend to be Li-poor, typically have higher δ^7 Li than non-carbonate sediment.

Fresh MORB has a very limited range of δ^7 Li of +1 to +6‰, with most values falling between +3 and +5, a range not much larger than that expected from analytical error alone. Oceanic island basalts (OIB) range to only slightly higher values, +7‰, than MORB, but only Hawaii has been studied in any detail. There appears to be little the relationship between δ^7 Li and other geochemical parameters, such as MgO concentration, suggesting Li experiences little isotopic fractionation during fractional crystallization, and perhaps also partial melting. Mantle



Figure 9.47. Li isotopic composition of terrestrial materials.

peridotites show a considerable range, from -15% to +10%, but unmetasomatized peridotites have a range of only 0% to +7%, just slightly larger than that of MORB. Alpine eclogites can have distinctly light isotopic compositions. These rocks are thought to be fragments of basaltic oceanic crust deeply subducted and metamorphosed then subsequently rapidly exhumed during the Alpine orogeny. Their light isotopic compositions presumably reflect preferential partitioning of ⁷Li into the fluid produced as the rocks were dehydrated during metamorphism.

 δ^7 Li in oceanic crust altered by seawater at low temperatures takes up Li from solution and has high δ^7 Li compared to fresh basalt. In hydrothermal reactions, however, Li is extracted from basalt into the solution and hydrothermal fluids can Li concentrations up to 50 times greater than seawater. ⁷Li is extracted more efficiently than ⁶Li during this process, so hydrothermally altered basalt can have δ^7 Li as low as -2‰. Serpentinites (hydrothermally altered peridotite) can have even lower δ^7 Li. However, because they extract Li from oceanic crust so completely, hydrothermal solutions have Li isotopic compositions intermediate between MORB despite this fractionation.

Island arc lavas have a range of δ^7 Li of +1 to +11, but all but a few values in the range +2 to +6, which is not very different from the range in MORB. This is somewhat surprising since other isotopic evidence (e.g., ¹⁰Be – Chapter 7) clearly demonstrates island arc magmas contain components derived from subducted oceanic crust and sediment. Furthermore, while δ^7 Li have been shown in some cases to correlate with chemical and isotopic indicators of a subduction component, this is not always the case.

It seems nevertheless likely that the subduction process has profoundly influenced the isotopic composition of the mantle over time. As a consequence of fractionation occurring during weathering, seawater is strongly enriched in ⁷Li. This enrichment is imprinted upon the oceanic crust as it reacts with

Geochemistry

CHAPTER 9: STABLE ISOTOPES

seawater. When the oceanic crust is returned to the mantle during subduction, the mantle becomes progressively enriched in ⁷Li. The continental crust, on the other hand, becomes progressively depleted in ⁷Li over time. Elliot et al. (2004) calculate that this process has increased δ^7 Li in the mantle by 0.5 to 1‰ and decreased δ^7 Li of the continental crust by 3‰ over geologic time.

References and Suggestions for Further Reading

- Bigeleisen, J. and M. G. Mayer. 1947. Calculation of equilibrium constants for isotopic exchange reactions, J. Chem. Phys., 15: 261-267.
- Bowen, R. 1988. Isotopes in the Earth Sciences, Barking (Essex): Elsevier Applied Science Publishers.
- Boyd, S. R. and C. T. Pillinger. 1994. A prelimiary study of ¹⁵N/¹⁴N in octahedral growth form diamonds. *Chem. Geol.* 116: 43-59.
- Broecker, W. and V. Oversby. 1971. *Chemical Equilibria in the Earth*, New York: McGraw-Hill (Chapter 7).
- Cerling, T. E., Y. Wang and J. Quade. 1993. Expansion of C₄ ecosystems as an indicator of global ecological change in the late Miocene, *Nature*, 361, 344-345.
- Chan, L. H., J. M. Edmond, G. Thompson and K. Gillis. 1992. Lithium isotopic composition of submarine basalts: implications for the lithium cycle of the oceans. *Earth Planet. Sci. Lett.* 108: 151-160.
- Chan, L.-H. and J. M. Edmond. 1988. Variation of lithium isotope composition in the marine environment: a preliminary report. *Geochim. Cosmochim. Acta*. 52: 1711-1717.
- Chan, L.-H. and F. A. Frey, Lithium isotope goechemistry of the Hawaiian plume: results from the Hawaii Scientific Drilling Project and Koolau Volcano, *Geochem. Geophys. Geosyst.*, 4: DOI: 10.1029/2002GC000365, 2003.
- Chaussidon, M. and A. Jambon. 1994. Boron content and isotopic composition of oceanic basalts: geochemical and cosmochemical implications. *Earth Planet. Sci. Lett.* 121: 277-291.
- Chaussidon, M. and B. Marty. 1995. Primitive boron isotope composition of the mantle. *Science*. 269: 383-266.
- Chaussidon, M. and F. Albarede. 1992. Secular boron isotope variations in the continental crust: an ion microprobe study. *Earth Planet. Sci. Lett.* 108: 229-241.
- Chaussidon, M., F. Albarède and S. M. F. Sheppard. 1987. Sulphur isotope heterogeneity in the mantle from ion microprobe measurements of sulphide inclusions in diamonds. *Nature*. 330: 242-244.
- Chaussidon, M., F. Albarede and S. M. F. Sheppard. 1989. Sulphur isotope variations in the mantle from ion microprobe analyses of micro-sulphide inclusions. *Earth Planet. Sci. Lett.* 92: 144-156.
- Chaussidon, M., S. M. F. Sheppard and A. Michard. 1991. Hydrogen, sulfur and neodymium isotope variations in the mantle beneath the EPR at 12°50′′N. in *Stable Isotope Geochemistry: A Tribute to Samuel Epstein*, ed. H. P. Taylor, J. R. O'Niel and I. R. Kaplan. 325-338. San Antonio: Geochemical Society.
- Chiba, H., T. Chacko, R. N. Clayton and J. R. Goldsmith. 1989. Oxygen isotope fractionations involving diopside, forsterite, magnetite, and calcite: application to geothermometry. *Geochim. Cosmochim. Acta*. 53: 2985-2995.
- Craig, H. 1963. The isotopic composition of water and carbon in geothermal areas, in *Nuclear Geology on Geothermal Areas*, ed. by E. Tongiorgi, p. 17-53, Pisa: CNR Lab. Geol. Nucl.
- Criss, R. E. and H. P. Taylor. 1986. Meteoric-hydrothermal systems, in *Stable Isotopes in High Temperature Geological Processes, Reviews in Mineralogy* 16, ed. by J. W. Valley, H. P. Taylor and J. R. O'Neil, p. 373-424, Washington: Mineral. Soc. Am., Washington.
- Dansgaard, W. 1964. Stable isotopes in precipitation. *Tellus*, 16, 436-463.
- Deloule, E., F. Albarede and S. M. F. Sheppard. 1991. Hydrogen isotope heterogeneity in the mantle from ion probe analysis of amphiboles from ultramafic rocks. 105: 543-553.
- Elliott, T., A. Jeffcoate and C. Bouman, The terrestrial Li isotope cycle: light-weight constraints on mantle convection, *Earth Planet. Sci. Lett.*, 220: 231-245, 2004.
- DeNiro, M. J. 1987. Stable Isotopy and Archaeology, Am. Scientist, 75: 182-191.
- DeNiro, M. J. and C. A. Hasdorf. 1985. Alteration of ¹⁵N/¹⁴N and ¹³C/¹²C ratios of plant matter during the initial stages of diagenesis: studies utilizing archeological specimens from Peru, *Geochim. Cosmochim. Acta*, 49, 97-115.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

- DeNiro, M. J. and S. Epstein. 1978. Influence of diet on the distribution of carbon isotopes in animals, *Geochim. Cosmochim. Acta*, 42: 495-506.
- DeNiro, M. J. and S. Epstein. 1981. Influence of diet on the distribution of nitrogen isotopes in animals, *Geochim. Cosmochim. Acta*, 45> 341-351.
- Eldridge, C. S., W. Compston, I. S. Williams, J. W. Harris and J. W. Bristow. 1991. Isotope evidence for the involvement of recycled sediment in diamond formation. *Nature*. 353: 649-653.
- Elliott, T., A. Jeffcoate and C. Bouman, The terrestrial Li isotope cycle: light-weight constraints on mantle convection, *Earth Planet. Sci. Lett.*, 220:231-245, 2004.
- Emiliani, C. 1955. Pleistocene temperatures, J. Geol., 63, 538-578.
- Farquhar, J., H. Bao and M. Thiemens, Atmospheric influence of Earth's earliest sulfur cycle, *Science*, 289: 756-758, 2000.
- Farquhar, J., B. A. Wing, K. D. McKeegan, J. W. Harris, P. Cartigny and M. Thiemens, Massindependent fractionation sulfur of inclusions in diamond and sulfur recycling on early Earth, *Science*, 298: 2369-2372, 2002.
- Farquhar, J. and B. A. Wing, Multiple sulfur isotopes and the evolution of atmospheric oxygen, *Earth Planet. Sci. Lett.*, 213: 1-13, 2003.
- Faure, G. 1986. Principles of Isotope Geology, 2nd ed., New York: J. Wiley & Sons.
- Ferronsky, V. I., and V. A. Polyakov. 1982. *Environmental Isotopes in the Hydrosphere*, Chichester: John Wiley and Sons.
- Fogel, M. L., and M. L. Cifuentes. 1993. Isotope Fractionation during Primary Production, in Organic Geochemistry: Principles and Applications edited by M. H. Engel and S. A. Macko, p. 73-98. New York: Plenum.
- Goldhaber, M., S. E. Church, B. R. Doe, C. Taylor, J. C. Brannon, and C. A. Gent. 1995. Sources and transport paths for MVT ore deposit lead and sulfur in the midcontinent of the U.S.A. in *International Field Conference on Carbonate-Hosted Lead-Zinc Deposits Extended Abstracts*. D. L. Leach and M. B. Goldhaber (ed.) p.111-114. Littleton (CO): Society of Economic Geologists.
- Harmon, R. S., and J. Hoefs. 1995. Oxygen isotope heterogeneity of the mantle deduced from a global analysis of basalts from different geotectonic settings, *Contrib. Mineral. Petrol.*, in press.
- Hayes, J. D., J. Imbrie and N. J. Shackleton. 1976. Variations in the Earth's orbit: pacemaker of the ice ages, *Science*, 194, 1121-1132.
- Heming, N. G. and G. N. Hanson. 1992. Boron isotopic composition and concentration in modern marine carbonates. *Geochim. Cosmochim. Acta*. 56: 537-543.
- Hoefs, J. 1988. *Stable Isotope Geochemistry*, 3rd ed., Berlin: Springer-Verlag.
- Hofmann, A. W. and W. M. White. 1982. Mantle plumes from ancient oceanic crust. *Earth Planet. Sci. Lett.* 57: 421-436.
- Imbrie, J. 1985. A theoretical framework for the Pleistocene ice ages, J. Geol. Soc. Lond., 142, 417-432.
- Imbrie, J., J. D. Hayes, D. G. Martinson, A. McIntyre, A. Mix, et al., 1985. The orbital theory of Pleistocene climate: support from a revised chronology of the marine δ¹⁸O record, in *Milankovitch and Climate, Part 1*, edited by A. L. Berger, J. Imbrie, J. Hayes, G. Kukla and B. Saltzman, p. 269-305. Dordrecht: D. Reidel.
- Ishikawa, T. and E. Nakamura. 1992. Boron isotope geochemistry of the oceanic crust from DSDP/ODP Hole 504B. *Geochim. Cosmochim. Acta*. 56: 1633-1639.
- Ishikawa, T. and E. Nakamura. 1993. Boron isotope systematics of marine sediments. *Earth Planet Sci. Lett.* 117: 567-580.
- Javoy, M. 1977. Stable isotopes and geothermometry, J. Geol. Soc. Lond., 133, 609-639.
- Javoy, M. and F. Pineau. 1991. The volatile record of a "popping" rock from the Mid-Atlantic Ridge at 14° N: chemical and isotopic composition of the gas trapped in the vesicles, *Earth Planet. Sci. Lett.*, 107, 598-611.
- Johnson, C. M., B. L. Beard and F. Albarede, *Geochemistry of Non-Traditional Stable Isotopes: Reviews in Mineralogy Volume*, Washington, Mineralogical Society of America, 454, 2004.
- Johnson, D. G., K. W. Jucks, W. A. Traub and K. V. Chance, 2001. Isotopic composition of stratospheric ozone, *J. Geohys. Res.*, 105:9025-9031.

Geochemistry

CHAPTER 9: STABLE SOTOPES

- Jouzel, J., N. I. Barkov, J. M. Barnola, M. Bender, J. Chappelaz, C. Genthon, V. M. Kotlyakov, V. Lipenkov, C. Lorius, R. Petit, *et al.*, 1993. Extending the Vostok ice-core record of paleoclimate to the penultimate glacial period, Nature, 364:407-412.
- Jouzel, C., C. Lorius, J. R. Perfit, C. Genthon, N.I. Barkov, V. M. Kotlyakov, and V. N. Petrov. 1987. Vostok ice core: a continuous isotope temperature record over the last climatic cycle (160,000 years), *Nature*, 329: 403-403.
- Jouzel, J., C. Waelbroeck, B. Malaizé, M. Bender, J. R. Petit, N. I. Barkov, J. M. Barnola, T. King, V. M. Kotlyakov, V. Lipenkov, et al., Climatic interpretation of the recently extended Vostok ice records, *Clim. Dyn.*, 12:513-521, 1996.
- Kasting, J. F. and D. Catling, Evolution of a Habitable Planet, *Ann. Rev. Astron. Astrophys.*, 41:429-463, 2003.
- Kawabe, I. 1979. Calculation of oxygen isotopic fractionation in quartz-water system with special reference to the low temperature fractionation, *Geochim. Cosmochim. Acta*, 42: 613-621.
- Kyser, K. T. 1986. Stable isotope variations in the mantle, in *Stable Isotopes in High Temperature Geologic Processes*, edited by J. W. Valley, H. P. Taylor and J. R. O'Neil, p. 141-164, Washington: Mineral. Soc. Am.
- Lawrence, J. R. and H. P. Taylor. 1972. Hydrogen and oxygen isotope systematics in weathering profiles, *Geochim. Cosmochim. Acta*, 36, 1377-1393.
- Lawrence, J. R. and J. R. Meaux. 1993., The stable isotopic composition of ancient kaolinites of North America, in *Climate Change in Continental Isotopic Records, Geophysical Monograph 78*, edited by P. K. Swart, K. C. Lohmann, J. McKenzie and S. Savin, p. 249-261.Washington: AGU.
- Marty, B. and N. Dauphas, The nitrogen record of crust-mantle interaction and mantle convection from Archean to present, *Earth Planet. Sci. Lett.*, 206: 397-410, 2003.
- Mattey, D. P. 1987. Carbon isotopes in the mantle. *Terra Cognita*. 7: 31-38.
- Mattey, D., D. Lowry and C. Macpherson. 1994. Oxygen isotope composition of mantle peridotite. *Earth Planet. Sci. Lett.* 128: 231-241.
- Mauersberger, K., 1987. Ozone isotope measurements in the stratosphere, Geophys. Res. Lett., 14:80-83.
- Milankovitch, M. 1938. Astronomische mittel zur erforschung der erdgeschlichtlichen klimate, Handbuch der Geophysik, 9: 593-698.
- Mojzsis, S. J., G. Arrhenius, K. D. McKeegan, T. M. Harrison, A. P. Nutman and C. R. L. Friend, Evidence for life on Earth before 3800 million years ago, *Nature*, 384: 55-59, 1996.
- Muehlenbachs, K. and G. Byerly. 1982. ¹⁸O enrichment of silicic magmas caused by crystal fractionation at the Galapagos Spreading Center, *Contrib. Mineral. Petrol.*, 79, 76-79.
- O'Leary, M. H. 1981. Carbon isotope fractionation in plants, *Phytochemistry*, 20, 553-567.
- O'Neil, J. R. 1986. Theoretical and experimental aspects of isotopic fractionation, in *Stable Isotopes in High Temperature Geologic Processes, Reviews in Mineralogy Volume 16*, J. W. Valley, H. P. Taylor and J. R. O'Neil, eds., Mineralogical Society of America, Washington, pp. 1-40, 1986.
- Ohmoto, H. and R. O. Rye, Isotopes of Sulfur and Carbon, in *Geochemistry of Hydrothermal Ore Deposits*, edited by H. Barnes, p. John Wiley and Sons, New York, 1979.
- Ohmoto, H., Stable isotope geochemistry of ore deposits, in *Stable Isotopes in High Temperature Geological Processes, Reviews in Mineralogy 16*, edited by J. W. Valley, H. P. Taylor and J. R. O'Neil, p. 491-560. Washington: Mineral. Soc. Am.
- Pagani, M., D. Lemarchand, A. J. Spivack and J. Gaillardet, A critical evaluation of the boron isotope-ph proxy: The accuracy of ancient ocean pH estimates, *Geochim. Cosmochim. Acta*, 69: 953-961, 2005.
- Palmer, M. R. 1991. Boron-isotope systematics of Halmahera arc (Indonesia) lavas: evidence for involvement of the subducted slab. *Geology*. 19: 215-217.
- Palmer, M., M. Spivack and J. M. Edmond. 1987. Temperature and pH controls over isotopic fractionation during adsorption of boron on marine clay. *Geochim. Cosmochim. Acta*. 51: 2319-2323.
- Park, R. and S. Epstein. 1960. Carbon isotope fractionation during photosynthesis, *Geochim. Cosmochim. Acta*, 21, 110-126.
- Quade, J., T. E. Cerling and J. R. Bowman. 1989. Development of Asian monsoon revealed by marked ecological shift during the latest Miocene in northern Pakistan, *Nature*, 342, 163-166.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

- Richet, P., Y. Bottinga and M. Javoy. 1977. A review of hydrogen, carbon, nitrogen, oxygen, sulphur, and chlorine stable isotope fractionation among gaseous molecules, *Ann. Rev. Earth Planet. Sci. Lett.*, 5, 65-110.
- Rosing, M., ¹3C-depleted carbon microparticles in >3700 Ma sea-floor sedimentary rocks from West Greenland, *Science*, 283: 674-676, 1999.
- Sakai, H., D. J. Des Marais, A. Ueda and J. G. Moore. 1984. Concentrations and isotope ratios of carbon, nitrogen, and sulfur in ocean-floor basalts, *Geochim. Cosmochim. Acta*, 48, 2433-2442, 1984.
- Sanyal, A., N. G. Hemming, G. N. Hanson and W. S. Broecker. 1995. Evidence for a higher pH in the glacial ocean from boron isotopes in foraminifera. *Nature*. 373: 234-236.
- Savin, S. M. and S. Epstein, The oxygen and hydrogen isotope geochemistry of clay minerals, *Geochim. Cosmochim. Acta*, 34, 25-42, 1970.
- Schidlowski, M., A 3800-million year isotopic record of life from carbon in sedimentary rocks, Nature, 333: 1988.
- Schoeninger, M. J. and M. J. DeNiro. 1984. Nitrogen and carbon isotopic composition of bone collagen from marine and terrestrial animals, *Geochim. Cosmochim. Acta*, 48, 625-639.
- Schwarcz, H. P., E. K. Agyei and C. C. McMullen. 1969. Boron isotopic fractionation during clay adsorption from sea-water. *Earth Planet. Sci. Lett.* 6: 1-5.
- Shackleton, N. J. and N. D. Opdyke. 1973. Oxygen isotope and paleomagnetic stratigraphy of an equatorial Pacific core V28-238: oxygen isotope temperatures and ice volumes on a 10⁵ and 10⁶ year time scale, *Quat. Res.*, *3*, 39-55.
- Sheppard, S. M. F. and C. Harris. 1985. Hydrogen and oxygen isotope geochemistry of Ascension Island lavas and granites: variation with fractional crystallization and interaction with seawater, *Contrib. Mineral. Petrol.*, 91, 74-81.
- Sheppard, S. M. F. and S. Epstein. 1970. D/H and ¹⁸O/¹⁶O ratios of minerals of possible mantle or lower crustal origin. *Earth Planet. Sci. Lett.* 9: 232-239.
- Sheppard, S. M. F., R. L. Nielsen and H. P. Taylor. 1969. Oxygen and hydrogen isotope ratios of clay minerals from porphyry copper deposits. *Econ. Geol.* 64: 755-777.
- Smith, H. J., A. J. Spivack, H. Staudigel and S. R. Hart, The boron isotopic composition of altered oceanic crust, *Chem. Geol.*, 126: 119-135, 1995.
- Spivack, A. and J. M. Edmond. 1987. Boron isotope exchange between seawater and the oceanic crust. *Geochim. Cosmochim. Acta*. 51: 1033-1044.
- Spivack, A. J., C.-F. You and H. J. Smith. 1993. Foraminiferal boron isotope ratios as a proxy for surface ocean pH over the past 21 Myr. *Nature*. 363: 149-151.
- Taylor, H. P. 1974. The application of oxygen and hydrogen studies to problems of hydrothermal alteration and ore deposition, *Econ. Geol.*, 69, 843-883.
- Taylor, H. P. 1980. The effects of assimilation of country rocks by magmas on ¹⁸O/¹⁶O and ⁸⁷Sr/⁸⁶Sr systematics in igneous rocks., *Earth Planet. Sci. Lett.*, 47, 243-254.
- Taylor, H. P. and S. M. F. Sheppard, 1986. Igenous rocks: I. Processes of isotopic fractionation and isotope systematics, in *Stable Isotopes in High Temperature Geological Processes*, edited by J. W. Valley, H. P. Taylor and J. R. O'Neil, Washington: Mineral. Soc. Am., p. 227-271.
- Teng, F.-Z., W. F. McDonough, R. L. Rudnick, C. Dalpe, P. B. Tomascak, B. W. Chappell and S. Gao, Lithium isotopic composition and concentration of the upper continental crust, Geochim. Comsochim. Acta, 68:4167-4178, 2004.
- Tomascak, P. B., Developments in the understanding and application of lithium isotopes in the Earth and planetary sciences, in C. M. Johnson, et al. (ed.), *Geochemistry of Non-Traditional Stable Isotopes*, *Reviews in Mineralogy Volume* 55, 153-195, 2004.
- Urey, H. 1947. The thermodynamic properties of isotopic substances, J. Chem. Soc. (London),1947: 562-581.
- Valley, J. W., H. P. Taylor and J. R. O'Neil (editors). 1986. *Stable Isotopes in High Temperature Applications, Reviews in Mineralogy Volume* 16. Washington: Mineral. Soc. Am.
- Vengosh, A., Y. Kolodny, A. Starinsky, A. Chivas and M. McCulloch. 1991. Coprecipitation and isotopic fractionation of boron in modern biogenic carbonates. *Geochim. Cosmochim. Acta*. 55: 2901-2910.

Geochemistry

CHAPTER 9: STABLE ISOTOPES

- Wang, Y., C. E. Cerling and B. J. McFadden. 1994. Fossil horse teeth and carbon isotopes: new evidence for Cenozoic dietary, habitat, and ecosystem changes in North America, *Paleogeogr., Paleoclim., Paleoceol.* in press.
- You, C.-F. 1994. Comment on "Boron content and isotopic composition of oceanic basalts: geochemical and cosmochemical implications" by M. Chaussidon and A. Jambon. *Earth Planet. Sci. Lett.* 128: 727-730.
- Zeebe, R., Stable boron isotope fractionation between dissolved B(OH)₃ and B(OH)₄, *Geochim. Cosmo-chim. Acta*, 69:2753-2766, 2005.

Problems

1. Using the procedure and data in Example 9.1, calculate the fractionation factor Δ_{CO-O_2} for the ¹⁷O/¹⁶O ratio at 300°C. What is the expected ratio of fractionation of ¹⁷O/¹⁶O to that of ¹⁸O/¹⁶O at this temperature?

2. What would you predict would be the ratio of the diffusion coefficients of H₂O and D₂O in air?

3. A granite contains coexisting feldspar (3% An content) and biotite with $\delta^{18}O_{SMOW}$ of 9.2‰ and 6.5‰ respectively. Using the information in Table 9.2, find the temperature of equilibration.

4. Spalerite and galena from a certain Mississippi Valley deposit were found to have $\delta^{34}S_{CDT}$ of +13.2‰ and +9.8‰ respectively.

a. Using the information in Table 9.4, find the temperature at which these minerals equilibrated.

b. Assuming they precipitated from an H_2S -bearing solution, what was the sulfur isotopic composition of the H_2S ?

5. Glaciers presently constitute about 2.1% of the water at the surface of the Earth and have a $\delta^{18}O_{SMOW}$ of \approx -30. The oceans contain essentially all remaining water. If the mass of glaciers were to increase by 50%, how would the isotopic composition of the ocean change (assuming the isotopic composition of ice remains constant)?

6. Consider the condensation of water vapor in the atmosphere. Assume that the fraction of vapor remaining can be expressed as a function of temperature (in kelvins):

$$f = \frac{T - 223}{50}$$

Also assume that the fractionation factor can be written as:

$$\ln \alpha = 0.0018 + \frac{12.8}{RT}$$

Assume that the water vapor has an initial $\delta^{18}O_{\text{SMOW}}$ of -9‰. Make a plot showing how the $\delta^{18}O$ of the remaining vapor would change as a function of *f*, fraction remaining.

7. Calculate the δ^{18} O of raindrops forming in a cloud after 80% of the original vapor has already condensed assuming (1.) the water initial evaporated from the ocean with δ^{18} O = 0, (2.) the liquid-vapor fractionation factor, $\alpha = 1.0092$.

8. A saline lake that has no outlet receives 95% of its water from river inflow and the remaining 5% from rainfall. The river water has δ^{18} O of -10% and the rain has δ^{18} O of -5%. The volume of the lake is steady-state (i.e., inputs equal outputs) and has a δ^{18} O of -3. What is the fractionation factor, α , for evaporation?

CHAPTER 9: STABLE ISOTOPES

9. Consider a sediment composed of 70 mole percent detrital quartz ($\delta^{18}O = +10\%$) and 30 mole percent calcite ($\delta^{18}O = +30\%$). If the quartz/calcite fractionation factor, α , is 1.002 at 500° C, determine the O isotopic composition of each mineral after metamorphic equilibrium at 500° C. Assume that the rock is a closed system during metamorphism, i.e., the $\delta^{18}O$ of the whole rock does not change during the process.

10. Assume that α is a function of the fraction of liquid crystallized such that $\alpha = 1 + 0.008(1-f)$. Make a plot showing how the δ^{18} O of the remaining liquid would change as a function of 1-f, fraction crystallized, assuming a value for Δ of 0.5%.

11. A basaltic magma with a δ^{18} O of +6.0 assimilates country rock that δ^{18} O of +20 as it undergoes fractional crystallization. Assuming a value of α of 1.002, make a plot of δ^{18} O vs. *f* for R = 5 and R = 10 going from *f* = 1 to *f* = 0.05.

CHAPTER 10: COSMOCHEMISTRY

CHAPTER 10: THE BIG PICTURE: COSMOCHEMISTRY

INTRODUCTION

n the previous eight chapters we acquired a full set of geochemical tools. In this and subsequent chapters, we will apply these tools to understanding the Earth. Certainly any full understanding of the Earth includes an understanding of its origin and it relationship to its neighboring celestial bodies. That is our focus in this chapter.

The question of the origin of the Earth is closely tied to that of the composition of the Earth, and certainly the latter is central to geochemistry. Indeed, one of the primary objectives of early geochemists was to determine the abundance of the elements in the Earth. It is natural to wonder what accounts for these abundances and to ask whether the elemental abundances in the Earth are the same as the abundances elsewhere in the solar system and in the universe. We might also ask why the Earth consists mainly of magnesium, silicon, oxygen, and iron? Why not titanium, fluorine and gold? Upon posing these questions, the realm of geochemistry melds smoothly into the realms of cosmochemistry and cosmology. Cosmochemistry has as its objective an understanding of the distribution and abundance of elements in the Solar System, and, to a lesser degree, the cosmos (the latter is relegated to secondary stature only because the data on objects outside the Solar System are much less complete).

The composition of the Earth is unique: there is undoubtedly no other celestial body that has exactly the same composition of the Earth. Nevertheless, the composition of the Earth is similar to that of the other terrestrial planets: Mercury, Venus, Mars, and the Moon. The Earth also shares a common geochemical heritage with the remainder of the Solar System, and all bodies in the Solar System probably have the same relative abundances of some elements, and the same isotopic compositions of most elements. What we know of the composition of the remainder of the universe suggests that it has a composition that is grossly similar to our Solar System: it is dominated by hydrogen and helium, with lesser amounts of carbon, oxygen, magnesium, silicon and iron, but there are local differences, particularly in the abundances of elements heavier than hydrogen and helium.

The unique composition of the Earth is product of three sets of processes. These include those processes responsible for the creation of elements, that is, nucleosynthetic processes; the creation of the Solar System, and finally the formation of the Earth itself. We will begin by considering nucleosynthesis. Meteorites are the principal record of formation of the Solar System and of the planetary bodies within it, so we devote considerable space to understanding these objects. Perhaps ever since he acquired the capacity to contemplate the abstract, man has wondered about how and when the Earth formed. The answer to the question can be provided by applying the tools of geochemistry to meteorites. We close the chapter by attempting to construct a history of Solar System and planetary formation from the meteorite record.

IN THE BEGINNING...NUCLEOSYNTHESIS

Astronomical Background

Nucleosynthesis is the process of creation of the elements. While we could simply take for granted the existence of the elements, such an approach is somehow intellectually unsatisfactory. The origin of the elements is both a geochemical and astronomical question, perhaps even more a cosmological one. Our understanding of nucleosynthesis comes from a combination of observations of the abundances of the elements (and their isotopes) in meteorites and from observations on stars and related objects. Thus to understand how the elements formed we need to understand a few astronomical observations and concepts. The universe began some 10 to 20 Ga ago with the Big Bang^{*}, since then the

^{*} The age of the universe is currently a hotly debated topic in astronomy and cosmology. Various lines of evidence appear to be contradictory and are not well explained by current theories. This makes one

CHAPTER 10: COSMOCHEMISTRY

universe has been expanding, cooling, and evolving. This hypothesis follows from two observations: the relationship between red-shift and distance and the cosmic background radiation, particularly the former.

Stars shine because of exothermic nuclear reactions occurring in their cores. The energy released by these processes creates a tendency for thermal expansion that, in general, exactly balances the tendency for gravitational collapse. Surface temperatures are very much cooler than temperatures in stellar cores. For example, the Sun, which is an average star in almost every respect, has a surface temperature of 5700 K and a core temperature thought to be 14,000,000 K.

Stars are classified based on their color (and spectral absorption lines), which is in turn related to their surface temperature. From hot to cold, the classification is: O, B, F, G, K, M, with subclasses designated by numbers, e.g., F5. (The mnemonic is 'O Be a Fine Girl, Kiss Me!'). The Sun is class G. Stars are also divided into Populations. Population I stars are second or later generation stars and have greater heavy element contents than Population II stars. Population I stars are generally located in the main disk of the galaxy, whereas the old first generation stars of Population II occur mainly in globular clusters that circle the main disk.

On a plot of luminosity versus wavelength of their principal emissions (i.e., color), called a Hertzsprung-Russell diagram (Figure 10.1), most stars (about 90%) fall along an array defining an in-

verse correlation, called the "Main Sequence" between these two properties. Since wavelength is inversely related to the fourth power of temperature, this correlations means simply that hot stars give off more energy (are more luminous) than cooler stars. Mass and radius are also simply related to temperature for these so-called main sequence stars: hot stars are big, small cool stars are small. Thus O and B stars are large, luminous and hot; K and M stars are small, faint, and cool. Stars on the main sequence produce energy by 'hydrogen burning', fusion of hydrogen to produce helium. Since the rate at which these reactions occur depends on temperature and density, hot, massive stars release more energy than smaller ones. As a result they exhaust the hydrogen in their cores much more rapidly. Thus there is an inverse relationship between the lifetime of a star, or at least the time it spends on the main sequence, and its mass. The most massive stars, up to 100 solar masses, have life expectancies of only about 10⁶ years, whereas small stars, as small as 0.01 solar masses, remain on the main sequence more than 10^{10} years.

The two most important exceptions to the main sequence stars, the red giants and the white dwarfs, represent stars that have burned all the H fuel in the cores and



Figure 10.1. The Hertzsprung-Russell diagram of the relationship between luminosity and surface temperature. Arrows show evolutionary path for a star the size of the Sun in pre- (a) and post- (b) main sequence phases.

wonder whether astronomy is not ripe for revolution, in which some of its most fundamental theories might be replaced by entirely new ones, or at least be radically altered.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

have moved on in the evolutionary sequence. H in the core is not replenished because the density difference prevents convection between the core and outer layers, which are still H-rich. The interior part of the core collapses under gravity. With enough collapse, the layer immediately above the He core will begin to "burn" H again, which again stabilizes the star. The core, however, continues to collapse until temperature and pressure are great enough for He burning to begin. At the same time, and for reasons not fully understood, the exterior expands and cools, resulting in a *red giant*, a star that is over-luminous relative to main sequence stars of the same color. When the Sun reaches this phase, in perhaps another 5 Ga, it will expand to the Earth's orbit. A star will remain in the red giant phase for something like 10^6-10^8 years. During this time, radiation pressure results in a greatly enhanced solar wind, of the order of 10^{-6} to 10^{-7} , or even 10^{-4} , solar masses per year (the Sun's solar wind is 10^{-14} solar masses per year, so that in its entire lifetime the Sun will blow off 1/10,000 of its mass through the solar wind).

The fate of stars after the red giant phase (when the He in the core is exhausted) depends on their mass. Nuclear reactions in small stars cease and they simply contract, their exteriors heating up as they do so, to become *white dwarfs*. The energy released is that produced by previous nuclear reactions and gravitational potential energy. This is the likely fate of the Sun. White Dwarfs are underluminous relative to stars of similar color on the main sequence. They can be thought of as little more than glowing ashes. Unless they blow off sufficient mass during the red giant phase, large stars die explosively, in supernovae. (Novae are entirely different events that occur in binary systems when mass from a main sequence star is pull by gravity onto a white dwarf companion.) Supernovae are incredibly energetic events. The energy released by a supernova can exceed that released by an entire galaxy (which, it will be recalled, consists of on the order of 10⁹ stars) for days or weeks!

THE POLYGENETIC HYPOTHESIS OF BURDIDGE, BURDIDGE, FOWLER AND HOYLE

Our understanding of nucleosynthesis comes from three sets of observations: (1) the abundance of isotopes and elements in the cosmos, (2) experiments on nuclear reactions that determine what reactions are possible (or probable) under given conditions, and (3) inferences about possible sites of nucleosynthesis and about the conditions that prevail in those sites. The abundances of the element in



Figure 10.2. Solar system abundance of the elements relative to silicon as a function of atomic number

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

primitive meteorites are by far our most important source of information of elemental abundances. Some additional information can be obtained from spectral observations of stars. The abundances of the elements in the Solar System are shown in Figure 10.2. Any successful theory of nuclear synthesis must explain this abundance pattern. Thus the chemical and isotopic composition of meteorites is a matter of keen interest, not only to geochemists, but to astronomers and physcists as well.

The cosmology outlined above provides two possibilities for formation of the elements: (1) they were formed in the Big Bang itself, or (2) they were subsequently produced. One key piece of evidence comes from looking back into the history of the cosmos. Astronomy is a bit like geology in that just as we learn about the evolution of the Earth by examining old rocks, we can learn about the evolution of the cosmos by looking at old stars. The old stars of Population II are considerably poorer in heavy elements than are young stars. In particular, Population II stars have an Fe/H ratio typically a factor of 100 lower than the Sun. This suggests that much of the heavy element inventory of the galaxy has been produced since these stars formed some 10 Ga ago. There are also significant variations in the Fe/H ratio between galaxies. In particular, dwarf spheroidal galaxies appear to be deficient in Fe, and sometimes in C, N, and O, relative to our own galaxy. Other galaxies show distinct radial compositional variations. For example, the O/H ratio in the interstellar gas of the disk of the spiral galaxy M81 falls by an order of magnitude with distance from the center. Finally, one sees a systematic decrease in the Fe/H ratio of white dwarfs (the remnants of small to medium size stars) with increasing age. On the other hand, old stars seem to have about the same He/H ratio as young stars. Indeed ⁴He seems to have an abundance of $25\pm 2\%$ everywhere in the universe.

Thus the observational evidence suggests that (1) H and He are everywhere uniform, implying their creation and fixing of the He/H ratio in the Big Bang, and (2) elements heavier than Li were created by subsequent processes. The production of these heavier elements seems to have occurred steadily through time, but the efficiency of the process varies between galaxies and even within galaxies.

Early attempts (~1930–1950) to understand nucleosynthesis focused on single mechanisms. Failure to find a single mechanism that could explain the observed abundance of nuclides, even under varying conditions, led to the present view that relies on a number of mechanisms operating in different environments and at different times for creation of the elements in their observed abundances. This view, which has been called the polygenetic hypothesis, was first proposed by Burbidge, Burbidge, Fowler and Hoyle (1957). The abundance of trace elements and their isotopic compositions in meteorites were perhaps the most critical observations in development of the theory. An objection to the polygenetic hypothesis was the apparent uniformity of the isotopic composition of the elements, but variations in the isotopic composition have now been demonstrated for a few elements in some meteorites. The isotopic compositions of other elements, such as oxygen and the rare gases, vary between classes of almost all meteorites. Furthermore, there are significant variations in isotopic composition of some elements, such as carbon, among stars. These observations provide strong support for this theory.

To briefly summarize it, the polygenetic hypothesis proposes four phases of nucleosynthesis. *Cosmological nucleosynthesis* occurred shortly after the universe began and is responsible for the cosmic inventory of H and He, and some of the Li. Even though helium is the main product of nucleosynthesis in the interiors of main sequence stars, not enough helium has been produced in this manner to significantly change its cosmic abundance. The lighter elements, up to and including Si and a fraction of the heavier elements, but excluding Li and Be, may be synthesized in the interiors of larger stars during the final stages of their evolution (*stellar nucleosynthesis*). The synthesis of the remaining elements occurs as large stars exhaust the nuclear fuel in their interiors and explode in nature's grandest spectacle, the supernova (*explosive nucleosynthesis*). Finally, Li and Be are continually produced in interstellar space by interaction of cosmic rays with matter (*galactic nucleosynthesis*).

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

Cosmological Nucleosynthesis

Immediately after the Big Bang, the universe was too hot for any matter to exist. But after a second or so, it had cooled to 10^{10} K so that a few protons and neutrons existed in an equilibrium dictated by the following reactions:

$${}^{1}H + e^{-} \rightleftharpoons n + \nu$$
 and $n + e^{+} \rightleftharpoons {}^{1}H + \nu^{\dagger}$

A somewhat worrisome question is why matter is exists at all. Symmetry would seem to require production of equal number of protons and anti-protons that would have annihilated themselves by:

$$p + \overline{p} \rightarrow 2\gamma^*$$

The current theory is that the hyperweak force was responsible for an imbalance favoring matter over anti-matter. The same theory predicts a half-life of the proton of 10^{32} yrs, a prediction not yet verified.

It took another 3 minutes to for the universe to cool to 10^9 K, which is cool enough for ²H to form:

$$H + {}^{1}n \rightarrow {}^{2}H + \gamma$$

At about the same time, the following reactions could also occur:

$${}^{2}H + {}^{1}n \rightarrow {}^{3}H + \gamma; {}^{2}H + {}^{1}H \rightarrow {}^{3}H + \gamma$$
$${}^{2}H + {}^{1}H \rightarrow {}^{3}He + \beta^{+} + \gamma; {}^{3}He + n \rightarrow {}^{4}He + \gamma$$
$${}^{3}He + {}^{4}He \rightarrow {}^{7}Be + \gamma; {}^{7}Be + e^{-} \rightarrow {}^{7}Li + \gamma$$

and

Formation of elements heavier than Li, however, was inhibited by the instability of nuclei of masses 5 and 8. Shortly afterward, the universe cooled below 10^9 K and nuclear reactions were no longer possible. Any free neutrons decayed away (neutrons outside the nucleus are unstable, having a half-life of 10 min). Thus the Big Bang created H, He and Li (⁷Li/H = 10^{-9}). Some 700,000 years later, the universe had cooled to about 3000 K, cool enough for electrons to be bound to nuclei, forming atoms.

Nucleosynthesis in Stellar Interiors

Hydrogen, Helium, and Carbon Burning Stars

For quite some time after the Big Bang, the universe was a more or less homogeneous, hot gas. "Less" turns out to be critical. Inevitably (according to fluid dynamics), inhomogeneities in the gas developed. These inhomogeneities enlarged in a sort of runaway process of gravitational attraction and collapse. Thus were formed protogalaxies, perhaps 0.5 Ga after the Big Bang. Instabilities within the protogalaxies collapsed into stars. Once this collapse proceeded to the point where density reaches 6 g/cm and temperature reaches 10 to 20 million K, nucleosynthesis began in the interior of stars by *hydrogen burning*, or the *pp process*, which involves reactions such as:

$${}^{1}H + {}^{1}H \rightarrow {}^{2}H + \beta^{+} + \nu;$$
 ${}^{2}H + {}^{1}H \rightarrow {}^{3}He + \gamma$ and ${}^{3}He + {}^{3}He \rightarrow {}^{4}He + 2{}^{1}H + \gamma$

There are also other reaction chains that produce ⁴He that involve Li, Be, and B, either as primary fuel or as intermediate reaction products. Later, when some carbon had already been produced by the first generation of stars and supernovae, second and subsequent generation stars with masses greater than about 1.1 solar masses produced He by another process as well, the *CNO cycle*:

$$^{12}C(p,\gamma) \ ^{13}N(\beta^+,\gamma) \ ^{13}C(p,\gamma) \ ^{14}N(p,\gamma) \ ^{15}O(\beta^+,\nu) \ ^{15}N(p,\alpha) \ ^{12}C^{\frac{1}{2}}$$

In this process carbon acts as a sort of nuclear catalyst: it is neither produced nor consumed. The net effect is consumption of 4 protons and two positrons to produce a neutrino, some energy and a ⁴He nucleus.

$${}^{^{12}}C(p,\gamma){}^{^{13}}N$$

$${}^{^{12}}C + p \rightarrow {}^{^{13}}N + \gamma$$

425

is equivalent to:

⁺ The v (nu) is the symbol for a neutrino.

^{*} γ (gamma) is used here to symbolized energy in the form of a gamma ray.

[‡] Here we are using a notation commonly used in nuclear physics. The notation:

CHAPTER 10: COSMOCHEMISTRY

The ¹⁴N(p, γ) ¹⁵O is the slowest in this reaction chain, so there tends to be a net production of ¹⁴N as a result. Also, though both ¹²C and ¹³C are consumed in this reaction, ¹²C is consumed more rapidly, so this reaction chain should increase the ¹³C/¹²C ratio.

The heat produced by these reactions counterbalances gravitational collapse and these reactions proceed in main sequence stars (Figure 10.1) until the hydrogen in the stellar core is consumed. How quickly this happens depends, as we noted earlier, on the mass of the star.

Once the H is exhausted in the stellar core, fusion ceases, and the balance between gravitational collapse and thermal expansion is broken. The interior of the star thus collapses, raising the star's temperature. The exterior expands and fusion begins in the shells surrounding the core, which now consists of He. This is the *red giant* phase. If the star is massive enough for temperatures to reach 10^8 K and density to reach 10^4 g/cc in the He core, *He burning* can occur:

$$He + {}^{4}He \rightarrow {}^{8}Be + \gamma$$
 and ${}^{8}Be + {}^{4}He \rightarrow {}^{12}C + \gamma$

The catch in this reaction is that the half-life of ⁸Be is only 10⁻¹⁶ sec, so 3 He nuclei must collide essentially simultaneously, hence densities must be very high. He burning also produce O, and lesser amounts of ²⁰Ne and ²⁴Mg, in the red giant phase, but Li, Be, and B are skipped: they are not synthesized in these phases of stellar evolution. These nuclei are unstable at the temperatures of stellar cores. Rather than being produced, they are consumed in stars.

There is a division of evolutionary paths once helium in the stellar core is consumed. Densities and temperatures necessary to initiate further nuclear reactions cannot be achieved by low-mass stars, such as the Sun, (because the gravitational force is not sufficient to overcome coulomb repulsion of electrons) so their evolution ends after the Red Giant phase, the star becoming a white dwarf. Massive stars, those greater than about 4 M_0^{T} , however, undergo further collapse and and the initiation of *carbon and oxygen burning* when temperatures reach 600 million K and densities $5 \times 10^5 \text{ g/cc}$. How-



Figure 10.3. Evolutionary path of the core of star of 25 solar masses (after Bethe and Brown, 1985). Note that the period spent in each phase depends on the mass of the star: massive stars evolve more rapidly.

ever, stars with intermediate masses, between 4 and 8 M_o , can be entirely disrupted by the initiation of carbon burning. For stars more massive 8 M_o , evolution now proceeds at an exponentially increasing pace (Figure 10.3) with reactions of the type:

$${}^{12}C + {}^{12}C \rightarrow {}^{20}Ne + {}^{4}He + \gamma$$

and ${}^{12}C + {}^{16}O \rightarrow {}^{24}Mg + {}^{4}He + \gamma$

Also, ¹⁴N created during the hydrogen burning phase of second generation stars can be converted to ²²Ne. A number of other less abundant nuclei, including Na, Al, P, S and K are also synthesized at this time, and in the subsequent process, *Ne burning*.

During the final stages of evolution of massive stars, a significant fraction of the energy released is carried off by neutrinos created by electron-positron annihilations in the core of the star. If the star is sufficiently oxygen-poor that its outer shells are reasonably transparent, the outer shell of the red giant may collapse during last few 10^4 years of evolution to form a *blue giant*.

 $^{{}^{\}rm T}$ The symbol O is the astronomical symbol for the Sun. M_{\circ} therefore indicates the mass of the Sun.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

The *e*-process

Eventually, a new core consisting mainly of ²⁸Si is produced. At temperatures near 10⁹ K and densities above 10^7 g/cc a process known as *silicon burning*, or the *e process*, (for equilibrium) begins, and lasts for a week or less, again depending on the mass of the star. This process includes reactions of the type:

 ${}^{28}\text{Si} + \gamma \rightleftharpoons {}^{24}\text{Ne} + {}^{4}\text{He}$ ${}^{28}\text{Si} + {}^{4}\text{He} \rightleftharpoons {}^{32}\text{S} + \gamma$ ${}^{32}\text{S} + {}^{4}\text{He} \rightleftharpoons {}^{36}\text{Ar} + \gamma$

While these reactions can proceed in either direction, there is some tendency for the build up of heavier nuclei with masses 32, 36, 40, 44, 48, 52 and 56. Partly as a result of the eprocess, these nuclei are unusually abundant in nature. In addition, because a variety of nuclei are produced during C and Si burning phases, other reactions are possible, synthesizing a number of minor nuclei.



Figure 10.4. Schematic diagram of stellar structure at the onset of the supernova stage. Nuclear burning processes are illustrated for each stage.

The star is now a cosmic onion of sorts, consisting of a series of shells of successively heavier nuclei and a core of Fe (Figure 10.4). Though temperature increases toward the interior of the star, the structure is stabilized with respective to convection and mixing because each shell is denser than the one overlying it.

Fe-group elements may also be synthesized by the e-process in Type I supernovae*. Type I supernovae occur when white dwarfs of intermediate mass (3-10 solar masses) stars in binary systems accrete material from their companion. When their cores reach the so-called Chandrasekhar limit, C burning is initiated and the star explodes.

The s-process

In second and later generation stars containing heavy elements, yet another nucleosynthetic process can operate. This is the slow neutron capture or *s-process*. It is so called because the rate of capture of neutrons is slow compared to the *r-process*, which we will discuss below. It operates mainly in the red giant phase (as evidenced by the existence of ⁹⁹Tc and enhanced abundances of several *s*-process nuclides in red giants) where neutrons are produced by reactions such as:

$${}^{13}C + {}^{4}He \rightarrow {}^{16}O + n$$
$${}^{22}Ne + {}^{4}He \rightarrow {}^{25}Mg + n$$
$${}^{17}O + {}^{4}He \rightarrow {}^{20}Ne + n$$

^{*}Astronomers recognize two kinds of supernovae: Type I and Type II. A Type I supernova occurs when a white dwarf in a binary system accretes mass from its sister star. Its mass reaches the point where carbon burning initiates and the star is explosively disrupted. The explosions of massive stars which we are considering are the Type II supernovae.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

(but even H burning produces neutrons, so that the s-process operates to some degree even in main sequence stars). These neutrons are captured by nuclei to produce successively heavier elements. The principle difference between the s and r process, discussed below, is the rate of capture relative to the decay of unstable isotopes. In the s process, a nucleus may only capture a neutron every thousand years or so. If the newly produced nucleus is not stable, it will decay before another neutron is captured. As a result, the s-process path closely follows the valley of stability on the chart of the nuclides and nuclear instabilities cannot be bridged.

Explosive Nucleosynthesis

The e-process stops at mass 56. In Chapter 8 we noted that ⁵⁶Fe had the highest binding energy per nucleon, i.e., it is the most stable nucleus. Thus fusion can release energy only up to mass 56; beyond this the reactions become endothermic, i.e., they consume energy. Once the stellar core has been largely converted to Fe, a critical phase is reached: the balance between thermal expansion and gravitational collapse is broken. The stage is now set for the catastrophic death of the star: a supernovaexplosion, the ultimate fate of stars with masses greater than about 8 solar masses. The energy released in the supernova is astounding. In its first 10 seconds, the 1987A supernova (Figure 10.5) released more energy than the entire visible universe, 100 times more energy than the Sun will release in its entire 10 billion year lifetime.

When the mass of the iron core reaches 1.4 solar masses (the Chandrasekhar mass), further gravitational collapse cannot be resisted even by coulomb repulsion. The supernova begins with the collapse of this stellar core, which would have a radius of several thousand km (similar to the Earth's radius) before collapse, to a radius of 100 km or so. The collapse occurs in a few tenths of a second.



Figure 10.5. Rings of glowing gas surrounding the site of the supernova explosion named Supernova 1987A photographed by the wide field planetary camera on the Hubble Space Telescope in 1994. The nature of the rings is uncertain, but they may be probably debris of the supernova illuminated by high-energy beams of radiation or particles originating from the supernova remnant in the center.

When matter in the center of the core is compressed beyond the density of nuclear matter $(3 \times 10^{14} \text{ g/cc})$, it rebounds, sending a massive shock wave back out. As the shock wave travels outward through the core, the temperature increase resulting from the compression produces a breakdown of nuclei by photodisintegration, e.g.:

$${}^{56}\text{Fe} + \gamma \rightarrow 13 \,{}^{4}\text{He} + 4 \,{}^{1}\text{n};$$

$${}^{4}\text{He} + \gamma \rightarrow 2 {}^{1}\text{H} + 2 {}^{1}\text{n}$$

Thus much of what took millions of years to produce is undone in an instant. However, photodisintegration produces a large number of free neutrons (and protons), which leads to another important nucleosynthetic process, the *rprocess*.

Another important effect is the creation of huge numbers of neutrinos by positron-electron annihilations, these particles having "condensed" as pairs from gamma rays. The energy carried away by neutrinos leaving the

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

supernova exceeds the kinetic energy of the explosion by a factor of several hundred, and exceeds the visible radiation by a factor of some 30,000. The neutrinos leave the core at nearly the speed of light. Though neutrinos interact with matter very weakly, the density of the core is such that their departure is delayed slightly. Nevertheless, they travel faster than the shock wave and are delayed less than electromagnetic radiation. Thus neutrinos from the 1987A supernova arrived a t Earth (some 160,000 years after the event) a few hours before the supernova became visible.

When the shock wave reaches the surface of the core, the outer part of the star is blown apart in an explosion of unimaginable violence. But amidst the destruction new nucleosynthetic processes occur. As the shock wave passes through outer layers, it may 'reignite' them, producing explosive Ne, O and C burning. These processes produce isotopes of S, Cl, Ar, Ca, Ti, and Cr, and some Fe.

The *r*-process

The neutrons produced by photodisintegration in the core are captured by those nuclei that manage to survive this hell. Because the abundance of neutrons is exceedingly high, nuclei capture them at a rapid rate; so rapid that even an unstable nucleus will capture a neutron before it has an opportunity to decay. The result is a build up of neutron-rich unstable nuclei. Eventually the nuclei capture enough neutrons that they are not stable even for a small fraction of a second. At that point, they β decay to new nuclides, which are more stable and capable of capturing more neutrons. This is the *rprocess* (rapid neutron capture), and is the principle mechanism for building up the heavier nuclei. It reaches a limit when nuclei beyond A \approx 90 are reached. These heavy nuclei fission into several lighter fragments. The r-process is thought to have a duration of 1 to 100 sec during the peak of the supernova explosion. Figure 10.6 illustrates this process.

The r-process produces broad peaks in elemental abundance around the neutron magic numbers (50, 82, 126, corresponding to the elements Sr & Zr, Ba, and Pb). This results because nuclei with magic numbers of neutrons are particularly stable, and have very low cross sections for capture of neutrons, and because nuclei just short of magic numbers have particularly high capture cross sections. Thus the



Figure 10.6. Diagram of the r-process path on a Z vs. N diagram. Dashed region is r-process path; solid line through stable isotopes shows the s-process path.

CHAPTER 10: COSMOCHEMISTRY



Figure 10.7. Z vs. N diagram showing production of isotopes by the r- s- and p-processes. Squares are stable nuclei; wavy lines are beta-decay path of neutron-rich isotopes produced by the r-process; solid line through stable isotopes shows the s-process path.

magic number nuclei are both created more rapidly and destroyed more slowly than other nuclei. When they decay, the sharp abundance peak at the magic number becomes smeared out.

The p-process

The r-process tends to form the heavier (neutron-rich) isotopes of a given element. Proton capture, or the *p*-process , also occurs in supernovae and is responsible for the lightest isotopes of a given element. The probability of proton capture is much less than that of neutron capture. You can easily imagine the reason for this: to be captured, the proton must have sufficient energy to overcome the coulomb repulsion and approach to within 10^{14} cm of the nucleus where the strong force dominates over the electromagnetic one. In contrast even low energy neutrons can be captured by nuclei since the neutron is uncharged and there is no coulomb repulsion. The production of nuclei by the p-process is much smaller than by neutron capture processes, and is significant only for those nuclides than cannot be produced in other ways. These tend to be the lightest isotopes of an element. Becuase of the improbability of proton capture, these light, p-process-only isotopes tend to be the least abundant.

Figure 10.7 illustrates how these three processes, the s-, r-, and p-process, create different nuclei. Notice the shielding effect. If an isotope with *z* protons and *n* neutrons has a stable isobar with n+x neutrons and *p*-*x* protons, this isotope is shielded from production by the r-process because β -decay will cease when that stable isobar is reached. The most abundant isotopes of an element tend to be those created by all processes; the least abundant are those created by only one, particularly by only the p-process. The exact abundance of an isotope depends on a number of factors, including its neutron-capture cross section^{*}, and the neutron capture cross section and stability of neighboring nuclei.

^{*} In a given flux of neutrons, some nuclides will be more likely to capture and bind a neutron that others, just as some atoms will be more likely to capture and bind an electron than others. The neutron capture cross section of a nuclide is a measure of the affinity of that nuclide for neutrons, i.e., a measure of the

CHAPTER 10: COSMOCHEMISTRY

Let's return to the exploding star. In the inner part of the stellar core, the reactions we just discussed do not take place. Instead, the core collapses to the point where all electrons are welded to protons to form a ball of neutrons: a neutron star. This inner core is hot: 100 billion Kelvin. And like a ballerina pulling in her arms, it conserves angular momentum by spinning faster as it collapses. The neutron star inside the expanding supernova shell of 1987A may be spinning at 2000 revolutions per second. Neutron stars emits radiation in beacon-like fashion: a *pulsar*. The collapse of cores of the most massive stars, however, may not stop at all. They collapse to a diameter of zero and their density becomes infinite. Such an object is called a singularity. Its gravitational attraction is so great even light cannot escape, creating a *black* hole.



Figure 10.8. Comparison of relative abundances in cosmic rays and the solar system.

Nucleosynthesis in Interstellar Space

Except for production of ⁷Li in the Big Bang, Li, Be, and B are not produced in any of the above situations. One clue to the creation of these elements is their abundance in galactic cosmic rays: they are over abundant by a factor of 10⁶, as is illustrated in Figure 10.8. They are believed to be formed by interactions of cosmic rays with interstellar gas and dust, primarily reactions of ¹H and ⁴He with carbon, nitrogen and oxygen nuclei. These reactions occur at high energies (higher than the Big Bang and stellar interiors), but at low temperatures where the Li, B and Be can survive.

SUMMARY

Figure 10.9 is a part of the Z vs. N plot showing the abundance of the isotopes of elements 68 through 73. It is a useful region of the chart of the nuclides for illustrating how the various nucleo-synthetic processes have combined to produce the observed abundances. First, we notice that even numbered elements tend to have more stable nuclei than odd numbered ones — a result of the greater stability of nuclides with even Z. We also notice that nuclides having a neutron-rich isobar (recall that isobars have the same value of A, but a different combination of N and Z) are underabundant, for example ¹⁷⁰Yb and ¹⁷⁶Lu. This underabundance results from these nuclides being 'shielded' from production by β^- decay of r-process neutron-rich nuclides. In these two examples, ¹⁷⁰Er and ¹⁷⁶Yb would be the ultimate product of neutron-rich unstable nuclides of mass number 170 and 176 produced during the r-process. Also notice that ¹⁶⁸Yb, ¹⁷⁴Hf and ¹⁸⁰Ta are very rare. These nuclides are shielded from the r-process and are also off the s-process path. They are produced only by the p-process. Finally, those nuclides that can be produced by both the s- and the r-process. ¹⁷⁶Yb cannot be produced by the s-process only while the latter can be produced by both the s- and the r-process. ¹⁷⁶Yb cannot be produced by the s-process.

probability of that nuclide capturing a neutron in a given neutron flux.

Geochemistry

7											
2										180	181
73 Ia										.01	99.9
72 Hf					174	175	176	177	178	179	180
					.17	70 d	5.2	18.5	27.1	13.8	35.2
1							175	176	177		
71 LU							97.4	2.6	6d		
70 Yb	168	169	170	171	172	173	174	175	176		
	.135	31 d	3.0	14.3	21.8	16.1	31.8	4 d	12.7		
69 Tm			169	170	171	172					
			100	129 d	1.9 y	63 h					
68 Er	166	167	168	169	170	171					
	33.4	22.9	27.1	9 d	14.9	8 h					
	98	99	100	101	102	103	104	105	106	107	108
$N \longrightarrow$											

CHAPTER 10: COSMOCHEMISTRY

Figure 10.9. View of Part of Chart of the Nuclides. Mass numbers of stable nuclides are shown in bold, their isotopic abundance is shown in italics as percent. Mass numbers of short-lived nuclides are shown in plain text with their half-lives also given.

because during the s-process, the flux of neutrons is sufficiently low that any ¹⁷⁵Yb produced decays to ¹⁷⁵Lu before it can capture a neutron and become a stable ¹⁷⁶Yb.

The heavy element yield of stellar and explosive nucleosynthesis will vary tremendously with the mass of the star. A star of 60 solar masses will convert some 40% of its mass to heavy elements. The bulk of this is apparently ejected into the interstellar medium. Stars too small to become supernovae will convert relatively small fractions of their mass to heavy elements, and only a very small fraction of this will be ejected. One the whole, stars in the mass range of 20-30 solar masses probably produce the bulk of the heavy elements in the galaxy. While such stars, which are already quite large compared to the mean stellar mass, convert a smaller fraction of their mass to heavy elements than truly massive stars, they are much more abundant than the very massive stars.

Novae may also make a significant contribution to the cosmic inventory of a few relatively light elements such as ¹⁹F and ⁷Li, as well as the rarer isotopes of C, N, and O. Novae occur when mass is accreted to a white dwarf from a companion red giant. If the material is mainly hydrogen and accretion is relatively slow, H burning may be ignited on the surface of the white dwarf, resulting in an explosion that ejects a relatively small fraction of the mass of the star.

METEORITES: ESSENTIAL CLUES TO THE BEGINNING

In subsequent sections we want to consider the formation of the Earth and its earliest history. The Earth is a dynamic body; its rock formations are continually being recycled into new ones. As a result, old rocks are rare. The oldest rocks are 4.0 Ga; some zircon grains as old as 4.2 Ga have been found in coarse-grained, metamorphosed sediments. The geological record ends there: there is no trace of the earliest history of the Earth in terrestrial rocks. So to unravel Earth's early history, we have to turn to other bodies in the Solar System. So far, we have samples only of the Moon and meteorites, and a

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

few rather marginal quality analyses of the surface of Venus and Mars⁺. The Moon provides some clues to the early history of planets; but meteorites provide the best clues as to the formation of planets and the Solar System. We now turn our attention to them.

Meteorites are divided into *Falls* and *Finds*. Falls are meteorites recovered after observation of a fireball whose trajectory can be associated the impact site. (A fireball is a just an intense falling star. Most falling stars are produced by meteoroids so small, a few grams or tens of grams perhaps, that they completely evaporate during their passage through the atmosphere and thus never reach the surface.) Finds are meteorites found but not observed falling. Because finds may have been on the surface of the Earth for very long times, they are often weathered and the compositional information they provide is less reliable than that of falls. An exception of sorts to this is the Antarctic meteorites. Meteorites have been found in surprising numbers in the last 20 years or so in areas of low snowfall in Antarctica where ice is eroded by evaporation and wind. Meteorites are concentrated in such areas by glaciers. Because of storage in the Antarctic deep freeze, they are little weathered.

Figure 10.10 illustrates the gross features of meteorite classification. Meteorites are classified according to (1) their composition, (2) their mineralogy, and (3) their texture. The classification may seem of little importance, but since the above factors are a reflection of the history of meteorites, un-

IRONS



Figure 10.10. Classification of meteorites, illustrating the various textural and chemical parameters that lead to the classification system.

⁺ As we shall see, a few meteorites probably come from Mars, providing additional information on the composition of that planet.
Geochemistry

CHAPTER 10: COSMOCHEMISTRY

derstanding the classification helps to understand the origin of meteorites and the early Solar System. These classification schemes have evolved over the years and multiple names exist for the same classes, which can lead to some confusion for the uninitiated.

The first order division is between Stones and Irons. You can pretty well guess what this means: stones are composed mainly of silicates while irons are mainly metal. An intermediate class is the Stony-Irons, a mixture of silicate and metal. Figure 10.11 illustrates the relative abundance of the various meteorite types among falls. Stones predominate among falls (irons are more common among finds because they are more likely to be recognized as meteorites, and because they are more likely to be preserved). Stony meteorites are rather friable and disintegrate rapidly in transit through the atmosphere and afterward. Even among the falls, irons are probably grossly overrepresented for these reasons. As of 1978, there were about 3000 meteorites in collections around the world, about 1000 of which were falls. Over 6000 have been added since then by the Antarctic collecting programs.



major types of meteorites Falls.

Chondrites: The Most Primitive Objects

Let's consider the stones in more detail. They are divided into *chondrites* and *achondrites* depending on whether they contain *chondrules*, at least in principle. Chondrules are small spherical or ellipsoidal particles that were once molten and can consitute up to 80 percent of the mass of chondrites (though the average is closer to perhaps 40%). Chondrites are generally of ultramafic composition. They consist of varying proportions of the following components: a high temperature component consisting of chondrules and other high-temperature inclusions (called refractory inclusions or calcium-aluminum inclusions), large aggregates of metal and sulfides, small aggregates of metal, sulfide and oxides, isolated coarse silicate grains, and a fine-grained, porous mixture of minerals and amorphous material called the matrix. These components formed while dispersed in the solar nebula and they subsequently aggregated to form the meteorite parent bodies (we will discuss the evidence that meteorites are derived from larger parent bodies subsequently). All chondrites have been variably



Figure 10.12. A chondritic meteorite from NASA's Antarctic meteorite collection.

metamorphosed in their parent bodies. Some are highly brecciated as a result of collisions and impacts on the surface of the parent bodies. The least metamorphosed have a distinctly porous and fluffy texture, as would be expected from the aggregation of nebular dust. An example of a chondrite is shown in Figure 10.12.

Chondrules

Chondrules are usually a few tenths of a mm to a few mm in diameter. Mean size varies between chondrite classes (see below), but is typically around 0.5 mm. In the least metamorphosed meteorites,

CHAPTER 10: COSMOCHEMISTRY

they consist of quench crystals (often, but not always monmineralic) and/or glass. Olivine and Capoor pyroxene (enstatite, hypersthene) generally constitute over half the volume of the chondrule, with troilite (FeS), kamacite (FeNi alloy), Ca-rich pyroxene (pigeonite, diopside), Mg-Al spinel, chromite, and feldspar being less abundant. Chondrules have remnant magnetism that was presumably acquired as they cooled through their curie point in the presence of a magnetic field, indicating the presence of such a field in the solar nebula. From the number of compound chondrules (two chondrules fused together) and those having indentations suggestive of collisions with other chondrules, the chondrule density was as high a few per m³ at times and places in the solar nebula. While "dents" are observed in chondrules, microcraters produced by high velocity impact are absent. Many chondrules are compositionally zoned, and most chondrules contain nuclei of relict crystals. Many are rimmed with fine-grained dark secondary coatings of volatile-rich material broadly similar in composition to the chondrite matrix.

Chondrules are highly variable in composition. The chondrules within a single meteorite may show more chemical variation that between all classes of chondrites. On the other hand, the mean composition of chondrules is the same in all classes of meteorites. Si, Mg, Al, Ca, Fe, and O are the most abundant elements in chondrules. Minor elements include Na, S, K, Ti, Cr, Mn, and Ni. They are typically depleted in siderophile, chalcophile, and volatile elements relative to the whole chondrite. These depletions, particularly of siderophile and calcophile elements appears to have preceeded formation of the chondrules, and hence are unrelated to the phenomenon that created the chondrules.

The presence of glass and their spheroidal shape indicates that chondrules represent melt droplets, as has been realized for at least 100 years. However, their origin remains a mystery. The main problem is that at the low pressures that must have prevailed in the solar nebula, liquids are not stable: solids should evaporate rather than melt. In addition, the solar nebula, with an O/H ratio of ~7 × 10⁻⁴ should have been much too reducing for iron-bearing silicates to exist, yet many chondrules nevertheless have significant Fe²⁺ dissolved in silicates such as olivine. Finally, the source of heat to melt the chondrules is a mystery. Chondrules seem to have been heated quite rapidly and apparently cooled through their crystallization range on time scales of an hour or less, and possibly minutes. Though cooling was rapid, it was considerably slower than the rate that would have resulted from radiative cooling in open space. This conclusion is based on compositional zonation of minerals and experimental reproduction of textures. It is strengthened by other experiments that show chondrules would have evaporated if they existed in the liquid state any longer than this. All these observations indicate they formed very quickly, and may never have reached equilibrium.

Over the past 100 years or so, many mechanisms for chondrule formation have been proposed. These include formation through volcanism on planets and planetisimals or asteroids, impact melting resulting from collisions of planets or planetisimals, condensation from hot nebular gas, and transitory heating of pre-existing nebular or interstellar dust. Most of these hypotheses appear to be inconsistent with the compositional and textural properties of chondrules and there is a building consensus that they formed by transitory heating of 'cool' (<650 K) nebular dust. The exact mechanism and environment of formation, however, are uncertain. There are several possibilities, including collisions of small (<1 m) bodies, frictional heating of dust traveling through gas during infall, lightning, energy released by magnetic flares, or reconnection of magnetic field lines, and radiational heating resulting from high velocity outflows during the T-Tauri phase (see below) of the protosun. There are difficulties with all these hypotheses, and it is possible that chondrules were produced by an entirely unrecognized phenomenon in the olar nebula. More thorough reviews of chondrules, their properties, and the problem of their origin may be found in Wood (1988, Kerridge and Matthews (1988), and Taylor (1992).

Refractory Inclusions

There is also much interest in the other high-temperature component, the so-called Refractory Inclusions (RI's) or Ca-Al Inclusions (CAI's), because their compositions are similar to that of the first

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

condensates from a high-temperature gas. CAI's range in size from microscopic to 5-10 cm, though most are about the same size as chondrules (typically 0.5 mm). They are enriched, commonly by a factor of 20 or so, in refractory elements such as Ca, Al, Ti, Re and the noble metals, Ba, Th, Zr, Hf, Nb, Ta, Y, and the rare earths. Fe, Mg, and Si are also present, though depleted relative to bulk chondrites. The most common minerals are spinel (MgAl₂O₄), perovskite (CaTiO₃), melilite (Ca₂(Mg,Fe,Al,Si)₂O₇), and hibonite (Ca(Al,Mg,Ti)₁₂O₁₉. They can contain microscopic nuggets of refractory siderophile elements (Re, Os, Re, Pt, Ir, W, and Mo) that are mantled with Ni-Fe metal and sulfides. They clearly formed under high temperature, high oxygen fugacity conditions. An alternative origin to the high-temperature condensate scenario is that they are the refractory residues of vaporized material.

It is quite possible that some CAI's are high temperature condensates (particularly the finest ones), while others are refractory residues. There is, however, a broad consensus that they are the products of transitory and local heating events rather equilibrium condensates a hot solar nebula. Some CAI's show relative depletions in the both the most volatile REE (Eu and Yb) and the most refractory (Er and Lu), suggesting they are the products of multiple cycles of evaporation and condensation. Isotopic anomalies in a number of elements have been identified in some CAI's. This suggests a third possible origin for at least some: they are interstellar grains. Based on radiogenic isotope evidence, CAI's appear to be distinctly older, by several million years or so, than other material in meteorites.

The Chondrite Matrix

The matrix of chondrites is dark, FeO and volatile-rich material that is very fine-grained (typcial grain size is about 1 μ m). It can be quite heterogeneous, even on a 10 μ m scale. It also varies between material classes with an order

between meteorite classes, with an order of magnitude variation in Mg/Si, Al/Si The primary constituents and Na/Si. appear to be Fe- and Ca-poor pyroxene and olivine and amorphous material, but Fe-metal and a wide variety of silicates, sulfides, carbonates, and other minerals are also be present. In the most volatilerich meteorites the olivine and pyroxene have been altered to serpentitine; in the carbonaceous chondrites, carbonaceous material is present in substantial quantitites. One the whole, the composition of the matrix is complementary to that of the chondrules: whereas the latter are depleted in Fe and volatiles, the former are enriched in them. Very significantly, the matrix includes grains of SiC and diamond of anomalous isotopic composition. This material may well be debris of a supernova and may be the oldest solid material in the Solar System. We will discuss these isotopic variations in greater detail in a subsequent section.



Figure 10.13. Abundances of the elements (in molar units) in the Sun's photosphere vs. their abundances in the carbonaceous chondrite Orgueil (CI1). Abundances for most elements agree within analytical error. The highly volatile elements H, C, N, O and the rare gases (not plotted) are depleted in chondrites relative to the Sun. Lithium is depleted in the Sun.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

Chondrite Classes and Their Compositons

Chondrites are richest in volatiles and have compositions closest to that of the Sun (as determined from spectral studies of its photosphere and the composition of solar wind), though relative to the Sun they are severely depleted in H, He and other highly volatile elements (Figure 10.13). Achondrites, on the other hand, generally have textures similar to terrestrial igneous rocks and compositions suggestive of igneous differentiation. *Chondrites are thus considered the most primitive objects.* Some meteorites lacking chondrules, the CI chondrites, are nevertheless classed as chondrites on compositional grounds. These are the most compositionally primitive objects of all.

Chondrites are subdivided into *C* (*Carbonaceous*), *E* (*Enstatite*), and *Ordinary* chondrites. The subclassification into ordinary and E chondrites is based on their iron and nickel content and the degree of oxidation of the iron. The *ordinary chondrites* consist of classes H (High Iron or Bronzite), L (Low Iron or Hypersthene), and LL. The name LL reflects low total iron (and other siderophile metals) and low metallic Fe (and other siderophiles). The E-, or enstatite, chondrites are the most highly reduced, virtually all the iron is present as metal rather than as Fe^{2+} in silicates. The E-chondrites can be further subdivided into EH (high iron) and EL (low iron) classes. Reduction of iron increases the Si/(Fe²⁺ + Mg) ratio in silicates and results in enstatite, rather than olivine, being the dominant mineral in these objects. Besides enstatite, metal and sulfides, enstatite chondrites contain a number of other exotic minerals, such as phosphides, carbides and a oxynitride of Si, that indicate they formed under highly reducing conditions. Figures 10.14 and 10.15 illustrate the compositional differences between these chondrite classes.

The diagonal lines on Figure 10.14 are lines of constant total iron content. There has been long debate as to the origin of the variations in iron content. Some have supposed it relates to processes in the parent body, such as loss of iron to form the cores of planetisimals (with iron meteorites represent-



Figure 10.14. Ratio of reduced and oxidized iron to Si in various chondrite groups. The diagonal lines represent constant total iron concentration. After Wasson (1974).

ing the cores), but the present consensus among meteorists is that the differences between these groups reflect variations in the oxygen fugacity prevailing at the time and place of condensation of the grains which later coalesced to form meteorite parent bodies. Thus the various chondrite groups probably formed in different regions of the psolar nebula. Because hydrogen is by far the most abundant element in the solar system, oxygen fugacity would be controlled by the amount of hydrogen present. In a region where the H concentration is high, it would bind much of the available O as H₂O, leaving the Fe in a metallic state. Where the H concentration is low, O would be available to oxidize Fe. The H concentration might have varied either radially or vertically through the disk of the solar nebula, providing different environments for chondrite formation.

As may be seen in Figure 10.14 and Table 10.1, the variations in oxidation state among ordinary chondrites are paralleled by other chemical variations. Thus, variations in oxygen fugacity alone cannot account for the compositional range seen in these objects.

CHAPTER 10: COSMOCHEMISTRY



Figure 10.15. Histogram of Al/Si, Mg/Si and Ca/Si ratios in various groups and petrologic types of chondrites. (After Wasson, 1974).

Geochemistry

Among the chondrites, and for that matter among all meteorites, carbonaceous chondrites are the most volatile rich and the most primitive. They may be further subdivided, but again, we run into some variation in classification. The older classification system divides them into subclasses I, II, and III, reflecting decreasing C and H₂O contents (hence CI are the most primitive). Wasson (1974) initiated a slightly different classification using a type meteorite for the name of the group and dividing the CIII's into two classes. Thus CI = CI (Ivuna), CII = CM(Mighei), CIII = CV (Vigarano) and CO (Ornans). CI are the most primitive objects known. As we noted earlier, CI's lack chondrules. Unfortunately, only one find exists in enough quantity for serious chemical analysis: Orgueil. Table 10.1 lists the general characteristics of the various chondrite groups.

Table 10.2 lists the current best estimate of the abundances of the elements in the Solar System (in units of atoms per 10⁶ S i atoms) and for Orgueil (in various units). While this is the best current summary, it is imperfect and can be expected to improve in the future.

	Principal	$E_0/(E_0/M_0)$	Motallic	Moan	Moan	Moan			Chon	drulae
	formana ann aoian	of all anto	E	M_{∞}/C	$\Lambda 1/C$:		\$180	\$170		ui uies
	terromagnesian	or sincate	ге	Mg/ 51	AI/ 51	Ca/51	0.0	00	Size Fi	requency
	silicate	(mole %)	(wt %)	(molar)	(mole %)	(mole %)) ‰	‰	(mm)	(%)
Carb	onaceous									
CI	serpentine	+	0	1.05	8.6	6.2	16.4	8.8		
CM	serpentine	+	0 - 1	1.05	9.7	6.8	12.2	4.0	0.2	≤15
CO	olivine	9 - 23	0 - 5	1.05	9.3	6.8	-1.1	-5.1	0.2	35-45
CV	olivine	6 - 14	0 - 8	1.07	11.6	8.4	0	-4.0	1.0	35-45
Ordi	inary									
Η	olivine	16 – 19	15 – 19	0.96	6.8	4.9	4.1	2.9	0.3	65-75
L	olivine	21 - 25	4 - 9	0.93	6.6	4.7	4.6	3.5	0.7	65-75
LL	olivine	25 - 32	0.3 – 3	0.94	6.5	4.7	4.9	3.9	0.9	65-75
Enst	atite									
E	enstatite	0.04 - 1.4	19 – 25	0.80	5.3	3.6	5.5	2.9	0.5	15-20

 Table 10.1.
 Characteristics of Chondrite Groups

⁺ Too fine grained to determine.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

TADLE	10.2.	Abundance	s of	тне	ELEMENTS
-------	-------	-----------	------	-----	----------

Ele	ement	Solar System Abundance*	ι σ (%)	Concentration in Orgueil (C1)	Eleme	ent	Solar System Abundance*	σ (%)	Concentration in Orgueil (C1) [†]
1	ц	2.79×1010	(/ 8 /	2 020	44 D		1 96	5.4	0.714
2	Ho	2.79×10^{-4} 2.72×10^{9}	_	$\frac{2.02}{0}$	44 N	h b	0.344	9. 4 8	0.714
2	I i	5.72×10^{10}	92	1 49	46 P	d d	1 39	6.6	0.154
4	Bo	730×10^{-1}	9.5	0.0249	40 Ι 47 Δ	u Va	0.486	6	0.000
т 5	B	7.50×10^{-10} 2.12 × 10 ¹	24	0.024)	48 (ъв Гd	1.61	65	0.197
6	C	1.01×10^{7}	<u> </u>	3 45%	49 Ir	n	0 184	6.4	0.000
7	N	3.13×10^{6}		3180	50 5	n	3.82	94	1.68
8	0	2.38×10^{7}	10	46.4%	51 S	h	0.309	18	0.133
9	F	8.43×10^2	15	58.2	52 T	'е	4 81	10	2 27
10	Ne	3.44×10^{6}	14	203 pL/g	53 I	C	0.9	21	0.433
11	Na	5.74×10^4	71	4900	54 X	e	47		86pL/g
12	Mø	1.07×10^{6}	3.8	9.53%	55 C	`s	0.372	6	0.186
13	Al	8.49×10^4	3.6	8690	56 B	a	4.49	4.5	2.340
14	Si	1.00×10^{6}	4.4	10.67%	57 L	a	0.446	2.0	0.236
15	P	1.04×10^{4}	10	1180	58 C	le	1.136	1.7	0.619
16	S	5.15×10^{5}	13	5.25%	59 P	'n	0.1669	2.4	0.09
17	C1	5.24×10^{3}	15	698	60 N	Jd	0.8279	1.3	0.463
18	Ar	1.01×10^{5}	6	7.51 pL/g	62 S	m	0.2582	1.3	0.144
19	Κ	3.77×10^{3}	7.7	566	63 E	'n	9.73×10^{-2}	1.6	0.0547
20	Ca	6.11×10^{4}	7.1	9020	64 C	Gd	0.330	1.4	0.199
21	Sc	3.42×10^{1}	8.6	5.86	65 T	b	6.03×10^{-2}	2.2	0.0353
22	Ti	2.40×10^{3}	5	436	66 D) _V	0.3942	1.4	0.246
23	V	2.93×10^{2}	5.1	56.7	67 H	Ío	8.89×10^{-2}	2.4	0.0552
24	Cr	$1.35 imes 10^4$	7.6	2660	68 E	lr	0.2508	1.3	0.162
25	Mn	9.55×10^{3}	9.6	1980	69 T	'n	3.78×10 ⁻²	2.3	0.022
26	Fe	9.00×10^5	2.7	18.51%	70 Y	Ъ	0.2479	1.6	0.166
27	Со	2.25×10^{3}	6.6	507	71 L	u	3.69×10^{-2}	1.3	0.0245
28	Ni	$4.93 imes 10^4$	5.7	1.10%	72 H	Ηf	0.154	1.9	0.108
29	Cu	5.22×10^{2}	11	119	73 T	a	2.07×10^{-2}	1.8	0.014
30	Zn	1.26×10^{3}	4.2	311	74 V	V	0.133	5.1	0.0923
31	Ga	37.8	6.9	10.1	75 R	le	5.17×10^{-2}	9.7	0.0371
32	Ge	1.19×10^{2}	9.6	32.2	76 C)s	0.675	6.3	0.483
33	As	6.56	12	1.85	77 II	r	0.661	6.1	0.474
34	Se	62.1	6.4	18.2	78 P	't	1.34	7.4	0.973
35	Br	11.8	19	3.56	79 A	u	0.187	15	0.145
36	Kr	45.0	18	8.7pL/g	80 H	Ig	0.34	12	0.258
37	Rb	7.09	7	2.3	81 T	1	0.184	9.4	0.143
38	Sr	23.8	8.1	7.80	82 P	b	3.15	7.8	2.43
39	Y	4.64	6.0	1.53	83 B	i	0.144	8.2	0.111
40	Zr	11.4	6.4	6.4	90 T	'n	3.29×10^{-2}	5.7	0.0289
41	Nb	6.98×10^{-1}	14	0.246	92 U	J	8.20×10^{-3}	8.4	0.0074
42	Мо	2.55	5.5	0.928					

 $\boldsymbol{\sigma}$ is the estimated uncertainty in the solar system abundances.

* Atoms relative to Si = 10^6 * in ppm unless otherwise indicated. Modified from Anders and Grevesse (1989).

CHAPTER 10: COSMOCHEMISTRY

Group	Petrologic Type						
	1	2	3	4	5	6	
Carbonaceous Chondrites							
CI 5							
CM		14					
CV		4	4				
СО			5	1			
Ordinary Chondrites							
H		6	23	53	32		
L			9	11	28	17	
LL		6	1	7	20		
Enstatite Chondrites							
E				3	2	6	

TABLE 10.3. DISTRIBUTION OF FALLS AMONG CHONDRITES GROUPS AND TYPES

from Wasson (1974).

Though the principal compositional variations among chondrites probably reflect variations in conditions in the solar nebula, all meteorites have undergone subsequent metamorphism on their parent bodies.⁺ Van Schmus and Wood (1967) divided all chondrites into petrographic types 1 through 6,

Class	1	2	3	4	5	6	
I. Homogeneity of olivine and pyroxene		Greater than 5% mean deviation		Less than 5% mean deviation	Uniform		
II. Structural state of low-Ca pyroxene		Predominantly monoclinic		Abundant monoclinic crystals	undant Orthorhombic noclinic stals		
III. Development of secondary feldspar		Absent		Predominantly crystalline agg	as micro- gregates	Clear, interstitial grains	
IV. Igneous glass		Clear and isotr glass: variable	opic primary abundance	Turbid if present	Absent		
V. Metallic minerals (Max. Ni content)		Taenite absent or minor (<20%)	Kamacite and taenite present (>20%)				
VI. Average Ni of sulfide minerals		>0.5%		< 0.59	%		
VII. Overall texture	No Chondrules	Very sharpl chondrules	y defined	Well-defined chondrules	Chondrules readily seen	Poorly defined chondrules	
VIII. Texture of matrix	All fine- grained, opague	Much opaque matrix	Opaque Transparent matrix Transparent microcrystal- line matrix Recrystal		Recrystalliz	zed matrix	
IX Bulk carbon content	3 – 5%	0.8 - 2.6%	0.2 – 1%	<0.2%			
X. Bulk water content	18-22%	2 - 16%	0.3 – 3%	- 3% <1.5%			

TA ble 10.4 .	VAN SCHMUS	and Wood Petro	ographic Cla	ASSIFICATION OF	CHONDRITES
----------------------	------------	----------------	--------------	-----------------	------------

⁺ As we shall see, there is evidence that meteorites must have been part of larger planetary bodies

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

based on increasing degree of metamorphism and decreasing volatile content. Types 1 and 2 represent low temperature hydrothermal alteration and types 4-6 represent increasing high-temperature metamorphism. Type 3 objects have probably undergone the least metamorphism. Table 10.3 shows the distribution of falls among the classes and petrologic types of chondrites. Table 10.4 summarizes the Van Schmus and Wood classification scheme.

The petrologic types are used together with the above groups to classify meteorites as to origin and metamorphic grade, e.g., CV-2. Types 1 and 2 are missing among the ordinary chondrites, types 1-3 are missing among enstatite chondrites, and class 1 is restricted to CI. CI-1 meteorites have probably not been heated above 300 K since their accumulation into larger bodies.

Differentiated Meteorites

Achondrites

While all the chondrites seem reasonably closely related, the achondrites are a more varied group. A principle distinction from the chondrites, besides the absence of chondrules, is the textural evidence for an igneous origin. Achondrites are generally poorer in volatile and moderately volatile (e.g., the alkalis) elements. The principal Ca-poor classes are the *aubrites* (= enstatite achondrites) and the *urelites* and *diogenites* (= hypersthene achondrites). The aubrites are highly reduced; they resemble enstatite chondrites except the metal and sulfide is missing. Urelites consist of olivine, pyroxene and a few metal grains plus a percent or so carbon, present as graphite and diamond. The principle mineral of diogenites is, as you might have guessed, hyperstheme. They resemble the silicate fraction of ordinary chondrites except for a deficiency in olivine. The Ca-rich achondrites include the *Eucrites* and *Howardites*. The eucrites resemble lunar and, to a lesser extent, terrestrial basalts, and are also called *basaltic achondrites*. The howardites are extremely brecciated (most achondrites are brecciated), and are an heterogeneous mixture of ecurite and diogenite material. The brecciated Howardites, the basaltic Eucrites, and the cummulate-textured diogenites may all have come from a single parent body. The angrites (of which there are two, the name of the class being derived from Angra dos Reis) consist mostly of Al-Ti augite (Ca-rich pyroxene) and have a composition indicating a complex igneous differentiation history. Chemical variation in the achondrites is illustrated in Figure 10.16.

A unique group of achondrites, the SNC meteorites have much younger formation ages (~ 1.25 Ga)

than virtually all other meteorites. This, and certain features of their bulk compositions and trapped noble gases, led to the interpretation that these meteorites come from Mars, having been ejected by an impact event on that planet. This interpretation, initially controversial, is now the consensus view.

Irons

Iron meteorites were originally classified based largely on phase and textural relationships. Compositionally, they all consist primarily of Fe-Ni alloys with lesser amounts of (mainly Fe-Ni) sulfides. Octahedral taenite, one of the Fe-Ni alloys, is the stable Fe-Ni metal phase at T> 900°C



Figure 10.16. Chemical variation among achondrites and terrestrial and lunar basalts. From Wasson (1974).

ranging up to 1000 km in diameter.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

(Figure 10.17). At lower temperature, kamacite, a Ni-poor Fe-Ni alloy, exsolves on the crystal faces of the octahedron. If the Ni content falls below 6%, all the metal converts to kamacite at lower temperature. Thus the phases and textures of iron meteorites are related to their composition and cooling history. Iron meteorites consisting only of kamacite are named hexahedrites. If Ni exceeds 6%, some taenite persists and the overall pattern is octahedral (= octahedrites). At low Ni contents, kamacite dominates and forms large crystals (coarse octahedrites). At higher Ni, kamacite and crystal size diminish (fine and medium octahedrites). *Ataxites* are Ni-rich (>14%) iron meteor-



Ataxites are Ni-rich (>14%) iron meteor- Figure 10.17. Phase diagram for iron-nickel alloy. ites consisting of a fine-grained intergrowth

of kamacite and taenite. The 20% or so of irons with silicate inclusions form a separate class.

The current classification by Wasson is based on Ga and Ge abundances. These are named I - IV, based on decreasing Ga and Ge. It turns out that many other chemical parameters can be related to Ga and Ge concentrations. Subgroups within these classes are named A, B, etc. Figure 10.18 illustrates chemical variation among the irons. Figure 10.10 illustrates the relationship between texture and composition, as well as these two classification systems.

The chemical variations within individual subclasses of irons are consistent with those produced by fractional crystallization of metal liquid. The clear implication then is that all irons from an individual subclass come from a single parent body. Perhaps some 60 parent bodies are represented by



Figure 10.18. Ge-Ga and Ge-Ni plots showing location of iron meteorite groups. Points are anomalous irons. After Wasson (1985).

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

the suite of analyzed irons. There is a general consensus that iron meteorites, with a few notable exceptions, represent the cores of planetary bodies whose diameters ranged from a few tens to a few hundred kilometers, called planetesimals. Cooling rates, estimated from textures and diffusion profiles, are typically in the range of a few 10's of degrees per million years, provide the principle evidence for the planetesimal hypothesis. A few classes of irons, most notably the IAB, may represent impact melts rather than segregated cores (Wasson et al., 1980).

Stony-Irons

The main classes of stony-irons are the pallasites and the mesosiderites. Pallasites consist of a network of Fe-Ni metal with nodules of olivine. They probably formed at the interface between molten metal and molten silicate bodies, with olivine sinking to the bottom of the silicate magma. Mesosiderites are similar to the achondrites, but contain a large metallic component.

METEORITE MINERALOGY

By and large, meteorites consist of the same minerals we find on Earth. The three most common are olivine, pyroxene and plagioclase. Olivine predominates in carbonaceous chondrites. In reduced chondrites (E's and ordinaries), the Fe reduction means not only the ferromagnesian silicates become more magnesian, they also become more silica-rich, for obvious reasons. Thus pyroxenes dominate the more reduced ordinary chondrites. In extreme cases, these are pure Mg end-members, which are rare on Earth. Hydrated layered silicates (which tend to be extremely fine and difficult to identify) occur and can dominate the matrix of the CI and CM's. Table 10.5 gives an incomplete list of minerals in meteorites. Some of these minerals are extremely rare in the Earth. Indeed in meteorites some are restricted to the CAI's. Another group of minerals you may not be familiar with is those occurring in the irons, principally taenite and kamacite.

Relationships Among Meteorites and Meteorite Parent-Bodies

The limited variability in composition within meteorite classes and the compositional gaps between different classes suggests all meteorites of a common class share a close genetic history. Relatively young cosmic ray exposure ages, to be discussed below, extensive thermal metamorphism (reach perhaps 1000° C), evidence of melting in achondrites, and slow cooling rates of iron meteorites men-

Mineral	Formula	Mineral	Formula
Augite	(Ca, Mg, Fe ²⁺ , Al) ₂ (Si, Al) ₂ O ₆	Oldhamite	CaS
Baddeleyite	ZrO_2	Olivine	$(Mg, Fe)_2 SiO_4$
Chlorapatite	$Ca_5(PO_4)_3Cl$	Orthopyroxene	(Mg, Fe)SiO ₃
Christobalite	SiO ₂	Pentlandite	(Fe,Ni) ₉ S ₈
Chromite	FeCr ₂ O ₄	Pervoskite	CaTiO ₃
Cohenite	Fe ₃ C	Plagioclase	NaAlSi ₃ O ₈ — CaAl ₂ Si ₂ O ₈
Corrundum	Al_2O_3	Quartz	SiO ₂
Daubreelite	FeCr ₂ S ₄	Rutile	TiO ₂
Farringtonite	$Mg_3(PO_4)_2$	Serpentine	$Mg_3Si_2O_5(OH)_4$
Graphite, Diamond	C	Sodalite	3NaAlSiO₄∙NaCl
Enstatite	MgSiO ₃	Schreibersite	(Fe,Ni) ₃ P
Hibonite	$CaO \cdot 6Al_2O_3$	Spinel	$MgAl_2O_4$
Illmenite	FeTiO ₃	Taenite	$Fe_{<0.8}Ni_{>0.2}$
Kamacite	Fe _{0.93-0.96} Ni _{0.07-0.04}	Tridymite	SiO ₂
Magnetite	Fe ₃ O ₄	Troilite	FeS
Melitite	Ca[Al,Mg][SiAl] ₃ O ₇	Whitlockite	$Ca_3(PO_4)_2$
Nepheline	NaAlSiO ₄	Zircon	$ZrSiO_4$

Table 10.5. Meteorite Minerals

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

tioned above all indicate that meteorites were once parts of larger *parent* bodies. All meteorites from a given class probably have been derived from the same parent body, though this need not always be the case. Estimates of the radii of parent bodies of various meteorite classes range from 10 to 1000 km. There are some compositional similarities between different classes suggesting a genetic relationship between classes, and possible derivation from the same parent body. Thus the stony-iron mesosiderites are closely related to the achondritic eucrites, howardites and diogenites. These, together with the a subset of pallasites and IIIAB irons may come from the same parent body. Other pallasites seem more closely related to ordinary chondrites and to IAB irons. The achondritic aubrites share with the enstatite chondrites the characteristic of being highly reduced.

Although meteorites represent a wide range of compositions, there is no particular reason to believe that those in collections are are representative sample of the compositions of Solar System bodies. As we shall see, most meteorites probably come from the asteroid belt, where they are continually produced by collisions. Meteorites in collections represent only those that have been produced relatively recently. Most likely, they come from only a small fraction of these asteroids, and hence represent only a part of their compositional range.

Time and the Isotopic composition of the Solar System

METEORITE AGES

Conventional methods

Meteorite ages are generally taken to be the age of Solar System. The off cited value for this age is 4.56 Ga. Before we discuss meteorite ages in detail, we need to consider the question of precisely what event is being dated by radiometric chronometers. In Chapter 8, we found that radioactive clocks record the last time the isotope ratio of the daughter element, e.g., ⁸⁷Sr/⁸⁶Sr, was homogenized. This is usually some thermal event. In the context of what we know of early solar system history, the event dated might be (1) the time solid particles were removed from a homogeneous solar nebula, (2) thermal metamorphism in meteorite parent bodies, or (3) crystallization (in the case of chondrules and achondrites), or (4) impact metamorphism of meteorites or their parent bodies. In some cases, the nature of the event being dated is unclear.

The oldest reliable high precision age is from CAI inclusions of *Allende*, a CV3 meteorite. These give a Pb isotope age of 4.568 ± 0.003 Ga. The matrix of Allende seems somewhat younger, although this is uncertain. Thus this age probably reflects the time of formation of the CAI's. Precise Pb-Pb ages of 4.552 Ga have been reported by several laboratories for the *St. Severin* LL chondrite. The same age (4.552 ± 0.003 Ga) has been reported for 2 L5 chondrites. U-Pb ages determined on phosphates in equilibrated (i.e., petrologic classes 4-6) ordinary chondrites range from 4.563 to 4.504 Ga. As these phosphates are thought to be secondary and to have formed during metamorphism, these ages apparently represent the age of metamorphism of these meteorites. Combined whole rock Rb-Sr ages for H, E, and LL chondrites are 4.498 ± 0.015 Ga. However, within the uncertainty of the value of the 87 Rb decay constant, this age could be 4.555 Ga (uncertainties normally reported on ages are based only on the scatter about the isochron and the uncertainty associated with the analysis, they do not include uncertainty associated with the decay constant). The age of *Allende* CAI's thus seems 5 Ma older than the best ages obtained on ordinary chondrites. No attempt has been made at high-precision dating of CI chondrites as they are too fine-grained to separate phases.

Pb isotope ages of the unusual achondrite *Angra dos Reis*, often classed by itself as an 'angrite' but related to the Ca-rich achondrites, give a very precise age of 4.5578±0.0004 Ma. Ibitira, a unique unbrecciated eucrite, has an age of 4.556±0.006 Ga. Perhaps surprisingly, these ages are the same as those of chondrites. This suggests that the parent body of these objects formed, melted, and crystallized within a very short time interval. Not all achondrites are quite so old. A few other high precision ages (those with quoted errors of less than 10 Ma) are available and they range from this value down to 4.529±0.005 Ga for *Nueve Laredo*. Thus the total range of the few high precision ages in achondrites is about 30 million years.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

K-Ar ages are often much younger. This probably reflects Ar outgassing as a result of collisions. These K-Ar ages therefore probably date impact metamorphic events rather than formation ages.

The present state of conventional meteorite chronology may be summarized by saying that it appears the meteorite parent bodies formed around 4.56±0.005 Ga, and there is some evidence that high-temperature inclusions (CAI's: calcium-aluminum inclusions) and chondrules in carbonaceous chondrites may have formed a few Ma earlier than other material. Resolving events on a finer time-scale than this has proved difficult using conventional techniques. There are, however, other techniques which help to resolve events in early solar system history, and we now turn to these.

Initial Ratios

The reference 'initial' 87 Sr / 86 Sr of the solar system is taken as 0.69897±3, based on the work of Papanastassiou and Wasserburg (1969) on basaltic achondrites (BABI: <u>basaltic achondrite best initial</u>). Basaltic achondrites were chosen since they have low Rb/Sr and hence the initial ratio (but not the age) is well constrained in an isochron. Subsequent high precision analyses of individual achondrites yield identical results (earlier reported low *Angra Dos Reis* have subsequently been shown to be in error). CAI's and Rb-poor chondrules from Allende have an even lower initial ratio: 0.69877±3 consistent with the idea that these formed slightly earlier than the meteorite parent bodies.

The initial ¹⁴³Nd/¹⁴⁴Nd ratio of the solar system is taken as 0.506609 ± 8 (normalized to ¹⁴³Nd/¹⁴⁴Nd = 0.72190) based on the work on chondrites of Jacobsen and Wasserburg (1980). Achondrites seem to have slightly higher initial ratios, suggesting they formed a bit later.

The initial isotopic composition of Pb is taken from the work of Tatsumoto et al. (1973) on troilite from the Canyon Diablo iron meteorite as ²⁰⁶Pb/²⁰⁴Pb: 9.307, ²⁰⁷Pb/²⁰⁴Pb: 10.294, ²⁰⁸Pb/²⁰⁴Pb: 29.476. These values agree with the best initial values determined from chondrites, including Allende chondrules. More recent work by Chen and Wasserburg (1983) confirms these results, i.e.: 9.3066, 10.293, and 29.475 respectively.

Extinct Radionuclides

There is evidence that certain short-lived nuclides once existed in meteorites. This evidence consists of anomalously high abundances of the daughter nuclides in certain meteorites and fractions of meteorites. A list of these extinct radionuclides is presented in Table 10.6. These provide evidence for a nucleosynthetic event shortly before

TADLE 10.6.	SHORT-LIVED	RAdionuclides	in	тне	Early
SOLAR SYSTEM	М				,

n n s	Radio- nuclide	Half-life Ma	Decay	Daughter	Abundance Ratio
f	¹⁴⁶ Sm	103	α	¹⁴² Nd	$^{146}\text{Sm}/^{144}\text{Sm} \sim 0.005$
n	²⁴⁴ Pu	82	α, SF	Xe	244 Pu/ 238 U ~ 0.005
-	^{129}I	16	β	¹²⁹ Xe	$^{129}\mathrm{I}/^{127}\mathrm{I} \sim 1 \times 10^{-4}$
-	¹⁰⁷ Pd	7	β	¹⁰⁷ Ag	107 Pd/ 108 Pd ~ 2 × 10 ⁻⁵
.	⁵³ Mn	3.7	β	⁵³ Cr	$^{53}Mn / ^{55}Mn \sim 4 \times 10^{-5}$
L	²⁶ A1	0.7	β	²⁶ Mg	$^{26}\text{Al}/^{27}\text{Al} \sim 5 \times 10^{-5}$

the solar system formed. To understand why, consider the example of ¹²⁹I. It decays to ¹²⁹Xe with a half-life of 16 Ma. Hence 16 Ma after they were created, only 50% of the original atoms of ¹²⁹I would remain. After 2 half-lives, or 32 Ma, only 25% would remain, after 4 half-lives, or 64 Ma, only 6.125% of the original ¹²⁹I would remain, etc. After 10 half lives, or 160 Ma, only 1/2¹⁰ (0.1%) of the original amount would remain. Anomalously high abundance of ¹²⁹Xe relative to other Xe isotopes in a meteorite indicates some ¹²⁹I was present when the meteorite, or its parent body, formed. From this we can conclude that an event that synthesized ¹²⁹I occurred not more than roughly 10⁸ years before the meteorite formed.

These short-lived 'fossil' radionuclides also provide a means of relative dating of meteorites and other bodies. Of the various systems, the ¹²⁹I –¹²⁹Xe decay is perhaps most useful. Figure 10.19 shows relative ages based on this decay system. These ages are calculated from ¹²⁹I/¹²⁷I ratios, which are in turn calculated from the ratio of excess ¹²⁹Xe to ¹²⁷I. Since the initial ratio of ¹²⁹I/¹²⁷I is not known, the ages are relative (the age of the *Bjurböle* meteorite, an L4 chondrite, is arbitarily taken as 0).

CHAPTER 10: COSMOCHEMISTRY



Figure 10.19. Summary of I-Xe ages of meteorites relative to Bjurböle, which is used as a standard because it gives highly reproducible results. Based on data in Swindle and Podosek (1988).

The total range is apparent ages is roughly 20 million years, though most fall within a narrower time span. The ages "date" closure of the systems to Xe and I mobility, but there is considerable debate as to what this represents. Some, perhaps the majority, argue it reflects accretion onto parent bodies or metamorphism. Others argue it somehow dates nebular proceeses, though the range in apparent ages is much greater than the life span of the solar nebula, which based on theoretical calculations and astronomical observations, is not longer than 1 million years. Perhaps both are involved.

An important point is that the range in ages is narrow and there is little systematic variation in age with meteorite class. Carbonaceous and ordinary chondrites do seem to be slightly older than enstatite chondrites, which might in turn be slightly older on average than than LL chondrites. Differentiated meteorites are perhaps slightly younger.

Another important chronometer is isotopic variations of Xe due to fission of ²⁴⁴Pu, half-life 80 Ma (²⁴⁴Pu fissions mainly to the heavier Xe isotopes, ¹³²Xe, ¹³⁴Xe, and ¹³⁶Xe, and isotopic variations due to Pu fission are readily distinguished from those due to decay of ¹²⁹I). When ²⁴⁴Pu decays, a great deal of energy is given off in the process. The newly formed atoms are propelled through the surrounding crystal lattice at high velocities, leaving behind a trail of destruction referred to as a *fission tracks*. Since damaged crystal lattices dissolves more quickly than intact lattices, applying acid to a crystal will etch out and enlarge the fission tracks until they are visible under a microscope. Other naturally occurring radioactive nuclei such as uranium also produce fission tracks, but to a predictable degree. Any excess over this expected number of fission tracks for a known uranium concentration is indicative of the former presence of ²⁴⁴Pu. Such evidence for fossil ²⁴⁴Pu has been found in chondrites, indicating that the chondrite must have formed during the period in which ²⁴⁴Pu still existed in the solar system.

The mere existence of radiogenic ¹²⁹Xe requires the time span between closure of the presolar nebula to galactic nucleosynthesis and formation of the solar system be no more than about 150 Ma. This time constraint is further reduced by the identification of radiogenic ²⁶Mg, produced by the decay of ²⁶Al. That ²⁶Al is the source of the ²⁶Mg is evidenced by the correlation between ²⁶Mg and the Al/Mg ratio (Figure 10.20). Given the half-life of ²⁶Al of 0.72 Ma, assuming a production ratio for ²⁶Al/²⁷Al around 10^{-3} to 10^{-4} in nucleosynthesis and an apparent ²⁶Al/²⁷Al ratios in CAI's of about 10^{-5} suggests a nucleosynthetic event, such as a red giant or supernova, less than 8 Ma before formation of these CAI's.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY



Figure 10.20. Al-Mg evolution diagram for Allende samples. Labeled lines are slopes that would result if the inclusions formed with the stated ${}^{26}\text{Al}/{}^{27}\text{Al}$ ratio. ${\delta}^{26}\text{Mg}$ is the percent deviation from the terrestrial ${}^{26}\text{Mg}/{}^{24}\text{Mg}$ ratio; ${}^{27}\text{Al}$ is the sole stable isotope of Al. After Lee et al. (1976).

variations of Ag from the decay of ¹⁰⁷Pd (half-life 6.5 Ma) in iron meteorites indicate core formation in meteorite parent bodies began, and

was largely complete, within about 15 Ma of the nucleosynthetic event.

Cosmic Ray Exposure Ages and the Origin of Meteorites

As we saw in Chapter 8, cosmic rays colliding with matter in meteorites and planetary bodies produce new nuclides through spallation. The cosmic rays only penetrate to a limited depth (of the order of a meter or less: there is no cutoff, the flux falls off exponentially), so that only small bodies or the surfaces of larger bodies are exposed to cosmic rays. The rate of production of nuclides by cosmic ray bombardment can be estimated from experimental physics if the cosmic ray flux is known. Thus, assuming a more or less constant cosmic ray flux, the length of time an object has been exposed to cosmic rays, the 'exposure age', can be calculated from the amounts of cosmogenic nuclides. Exposure ages are only accurate to within about a factor of 2, due to all the assumptions that are required in estimating the rate of production of nuclides. Exposure ages for chondrites are shown in Figure 10.21; expoimportant observation is that most CAI's show ²⁶Mg exesses that are consistent with a relatively uniform ²⁶Al/²⁷Al value of about 5×10^{-5} at the time they formed. Chondrules, on the other hand, rarely show ²⁶Mg excesses, implying the ²⁶Al had decayed by the time they formed, which in turn implies that formation of the CAI's preceeded formation of the chondrules by several million years. This is a surprisingly long period and one which theorists have difficulty explaining.

Since ⁴¹Ca also should have been produced (half-life 80,000 yrs), and anomalous abundances of its decay product, ⁴¹K, have not been found, it seems probable that at least 2 Ma must have elapsed between nucleosynthesis and formation of the CAI's. Isotopic variations of Cr have also been identified in *Allende* inclusions, an enstatite chondrite and iron meteorites. They arise from the decay of ⁵³Mn (half-life: 3.7 Ma). These and isotopic



Figure 10.21. Cosmic ray exposure ages of meteorites. Filled histogram is for meteorites with regolith histories (i.e., brecciated meteorites). After Crabb and Schultz (1981).

Geochemistry

W. M. White

CHAPTER 10: COSMOCHEMISTRY

sure ages for irons were shown in Figure 8.26 (page 363). As can be seen, the ages for chondrites are considerably less than formation ages. From this we may conclude that meteorites became small bodies accessible to cosmic rays only comparatively recently. Before that, they must have been stored in larger parent bodies where they would have been protected from cosmic ray bombardment. Apparently, meteorites are continually produced by collisions of these larger bodies. Irons tend to have longer exposure ages than stones. This simply reflects their greater strength and resistance to break up. The tendency for exposure ages of individual meteorite classes to cluster suggests many, or perhaps all, members of the class were derived from a single parent body.

Orbits for a few observed falls have been reconstructed, and these reconstructed orbits confirm the suspicion that many meteorites originate in the asteroid belt (Figure 10.22 shows an example of an asteroid). Reflectance spectra of some asteroids can

be matched to specific groups of meteorites. Based on spectroscopic studies, Ceres, the largest asteroid (radius of 479 km) as well as several other large asteroids, appears to be compositionally similar CI and CM chondrites. Vestra has a spectra that closely matches that of the eucrites (Figure 10.23) and thus almost certainly has a basaltic crust. Other asteroids, notably those of class "M", appear to be composed of Fe-Ni metal and are analogous to iron meteorites. The reflectance spectra of other asteroids do not match those of any of the meteorite groups, which emphasizes the point made earlier

that meteorites represent and incomplete sampling of Solar System material. It is quite interesting and important that the asteroid belt appears to be compositionally zoned, with "igneous" asteroids, analagous to the differentiated meteorites, predominating in the inner part and "primitive" meteorites, analogous to carbonaceous chondrites, dominating in the outer part of the belt.

In the past the had been much speculation about the possibility that carbonaceous chondrites, particularly the CI chondrites, might be derived from comets rather than asteroids. Comets come from the outer reaches of the Solar System, and thus it was thought their compositions might preserve that of the primordial solar nebula. However, the composition of Comet Halley, as determined by the 1986-1987 Giotto probe found that it certainly does not have a "primitive" composition in the sense of matching that of the Sun and CI chondrites. Clearly, comets are composed of chemically processed material, but what these processes might have been is



Figure 10.23. Comparison of the laboratory-determined reflectance spectrum of eucrite meteorites (basaltic achondrites) with that of the asteroid Vestra. The close match suggests that the crust of Vestra is basaltic, and similar to eucrites. After McSween (1987).

groups, which emphasizes the point made earlier

Figure 10.22. The asteroid Gaspra, photographed by the Galilleo spacecraft in 1991. Gaspra is an Stype asteroid, consisting principally of olivine, pyroxenes, and metal with dimensions of about 19 by 12 by 11 kilometers. It appears to be collisional fragment of a larger body.

CHAPTER 10: COSMOCHEMISTRY

uncertain (the Giotto measurements were very imprecise, so perhaps this is not the final work on this subject).

ISOTOPIC ANOMALIES IN METEORITES

Neon Alphabet Soup and Interstellar Grains in Meteorites

Since Thomson's discovery that elements could consist of more than one isotope in 1912, scientists have realized that the isotopic composition of the elements might vary in the They also realized that these universe. variations, if found, might provide clues as to how the elements came into being. As the only available extraterrestrial material, meteorites were of obvious interest in this respect. However, isotopic analyses of meteorites, by Harold Urey among others, failed to reveal any differences between meteorites and terrestrial materials. This apparent isotopic homogeneity was raised as an objection to the polygenetic hypothesis of Burbidge, Burbidge, Fowler, and Hoyle (1957), since isotopic variations in space and time were an obvious prediction of this model. Within a few years of its publication, however, John Reynolds, a physi-



Figure 10.24. Neon isotopic compositions in a stepheating experiment on Orgueil CI chondrite, which produced the first evidence of 'pre-solar' or exotic Ne. The points connected by the line show the changing Ne isotope ratios with increasing temperature. Shaded area was the original estimate of the composition of the pure Ne-E component. Also shown are the compositions of Ne-A ('solar'), Ne-B ('planetary'), and Ne-S ('spallogenic'). After Black and Pepin (1969).

cist at the University of Califonia, Berkeley, found isotopic variations in noble gases, particularly neon and xenon (Reynolds, 1960).

Noble gases are present in meteorites at concentrations that are often as low as 1 part in 10¹⁰. Though they are fairly readily isolated and analyzed at these concentrations, their isotopic compositions are nonetheless sensitive to change due to processes such as radiogenic decay (for He, Ar, and Xe), spallation and other cosmic ray-induced nuclear processes, and solar wind implantation. In addition, mass fractionation can significantly affect the isotopic compositions of the lighter noble gases (He and Ne). Through the late 1960's, all isotopic variations in meteoritic noble gases were thought to be related to these processes. For example, Ne isotopic variations could be described as mixtures of three components, "Neon A" or planetary (similar in composition to the Earth's atmosphere), "Neon B", or solar, and "Neon S", or spallogenic (cosmogenic) (Figure 10.23). The isotopic variations in Xe discovered by Reynolds were nonetheless significant because they were due to the decay of ¹²⁹I and ²⁴⁴Pu, which must have been only recently (on a cosmic time scale) synthesized.

In 1969, the picture became more complex when evidence of a ²²Ne-rich component, named "Neon E" was found in the high temperature (900-1100°C) release fractions of six carbonaceous chondrites (Black and Pepin, 1969). However, the carrier of Neon-E proved difficult to identify. Many scientists participated in an intensive search over nearly 2 decades for the carrier phase of these components. The search quickly focused on the matrix, particularly that of CM2 meteorites. But the fine-grained nature of the matrix, together with the abundance of sticky and refractory organic compounds, made work with the groundmass difficult. In the late 1980's, E. Anders and his colleagues at the University of Chicago (e.g., Tang and Anders, 1988) found that Neon-E is associated with fine-grained (<6 μ m) graphite and SiC (silicon carbide) of the matrix. Ne-E actually consists of two isotopically distinct components: Ne-E(L), which was found to reside in graphite, and Ne-E(H) which resides in SiC. The ²⁰Ne/²²Ne ratio of Ne-E(H) is less than 0.01, while that of Ne-E(H) is less than 0.2.

CHAPTER 10: COSMOCHEMISTRY

The other key noble gas in this context is xenon. Having 9 isotopes rather than 3 and with contributions from both ¹²⁹I decay and fission of Pu and U, isotopic variations in Xe are bound to be much more complex than those of Ne. On the other hand, its high mass minimizes mass fractionation effects, so "solar" (more properly solar wind) and "planetary" Xe are isotopically similar. The first evidence of isotopic variations in Xe came in the early 1960's, but these variations were thought to be fissogenic (at one time it was argued they were produced by fission of short-lived superheavy elements). However, subsequent investigation revealed two distinct components that could not be expalined by decay of unstable nuclides. The Xe-HL component, so named because it shows enrichments in both the heaviest and lightest Xe isotopes (Figure 10.25), and the Xe-S component, enriched in ¹²⁸Xe and ¹³⁰Xe.

Anders' University of Chicago group eventually identified the carrier of Xe-HL as microdiamonds and that of Xe-S as SiC. These diamonds are extraordinarily fine, averaging only 10 Å in diameter and containing typically only 10^3 or so atoms. Roughly one in every four atoms is at the surface. Because the surface free energy is so large as a result (see Chapter 5), the properties of this material differs significantly from normal diamond. For example, it readily forms a colloid in aqueous solutions of pH > 3. This considerably complicated the effort to isolate it. The other interesting observation is that diamond is unstable at low pressures, and the conditions of interstellar space is are certainly low pressure. The question then arises as to how the could be produced. The best guess at present is condensation from vapor. While diamond is unstable relative to graphite at low pressure, it's formation can be kinetically favored under certain conditions, including condensation from vapor with high hydrogen-carbon ratios.

Once these interstellar grains were isolated, it was possible to study their isotopic compositions in detail, particularly through the use of ion microprobes[‡]. Very large variations in the isotopic composition of carbon and nitrogen, as well as smaller variations in the composition of Si, Mg, Ca, Ti, Sr, and Ba have been found. The characteristics of these interstellar grain are summarized in Table 10.7.



Figure 10.25. The isotopic composition of Kr and Xe of the 'Xe-HL ' component in Allende matrix. Xe-HL is characteristically enriched in both the light and heavy isotopes, while the lighter noble gases show enrichment only in the heavy isotopes. After Anders (1988).

In addition, a fourth type of interstellar grain, Al_2O_3 , has also been isolated from the matrix of carbonaceous chondrites.

Discovery of isotopically anomalous instellar grains has inspired theorists to attempt to explain them, and there has been considerable progress on understanding stellar and explosive nuclosynthesis in the past few years as a result. However, even the very limited treatement of nucleosynthetic processes in stars earlier in the chapter is sufficient to allow us to identify the environment in which some of these grains were produced. Thus, if we examine a chart of the nuclides, we quickly see that the lightest Xe isotope, ¹²⁴Xe is a p-process-only nuclide, while the heaviest Xe isotope, ¹³⁶Xe, is a rprocess-only nuclide (in addition to production by fission, which, however, produces other Xe isotopes as well). The p- and rprocesses occur in supernovae, thus Xe-HL, as

[‡] Ion microprobes fire a narrow beam of ions (often O or Ar) at a surface. This produces ions from the surface that can be analyzed in an attached magnetic sector mass spectrometer. This is known as secondary ionization mass spectrometry or SIMS). Because the ion beam can be focused very finely (a few microns in diameter), very small areas can be analyzed.

CHAPTER 10: COSMOCHEMISTRY

	TADLE TO.7. ISOTOPIC CHARACTERISTICS OF INTERSTELLAR ORALINS								
Phase	Diamond	SiC	Graphite						
Isotopic Component	Xe-HL	Xe-S, Ne-E(H)	Ne-E(L)						
Enriched in Isotopes	¹²⁴ Xe, ¹³⁶ Xe	¹²⁸ Xe, ¹³⁰ Xe, ²² Ne	²² Ne						
Nuclear Process	p, r	s, ${}^{22}Na(\beta^+,\nu){}^{22}Ne$	$^{14}N+2\alpha \rightarrow ^{22}Ne$						
Grain Size, μ	0.001	0.03-10	0.8-7						
Abundance in C2 chondrites, ppr	n 400	7	<2						
$({}^{13}C/{}^{12}C)/({}^{13}C/{}^{12}C)_{\odot}$	0.96	0.03 - 50	0.012 - 50						
$({}^{15}N/{}^{14}N)/({}^{15}N/{}^{14}N)_{\odot}$	0.66	0.015 - 20	0.55 - 6.7						

Table 10.7. Isotopic Characteristics of Interstellar Grains

Modified from Anders and Zinner (1993).

well perhaps as the diamonds that carry it, must have been produced in supernovae. Xe-S is enriched in ¹²⁸Xe and ¹³⁰Xe, which are s-process-only isotopes. The s-process, of course, operates mainly in red giants, so we might guess the SiC was produced in red giants. Carbon and nitrogen in the SiC is, in most cases, enriched in ¹⁴N and ¹³C relative to normal solar system nitrogen and carbon. As we noted earlier in the chapter, there tends to be some net production of ¹⁴N, and consumption of ¹²C in the CNO cycle, which operates in main sequence stars, but also in the H-burning shell of red giants. As it turns out, our guess of red giants as sources of this SiC would be a good one. Theoretical studies show a close match between the observed isotopic patterns and those produced in the red giant phase (also called asymptotic giant branch) of medium-sized stars ($1-3M_{\odot}$). These studies show that such stars could also produce the ¹⁰⁷Pd and ²⁶Al that was present when the meteorites formed (e.g., Wasserburg et al., 1995).

The Ne-E(L) in graphite is interesting because it appears to consist of pure, or nearly so, ²²Ne. Its origin posed something of a mystery: any process that synthesizes ²²Ne should also synthesize other Ne isotopes as well as isotopes of other noble gases. The answer lies in theoretical calculations of nucleosynthesis, which indicate that under conditions for synthesis of ²⁰Ne-poor Ne, most of mass 22 would be ²²Na, which has a half-life of 2.6 years. Na could readily separate from Ne and other noble gases by condensation into grains if ambient temperatures are below several hundred degrees C, as would certainly be the case in the interstellar medium. Both novae and the He-burning shells of presupernovae stars are possible sources of this material.

Other Exotic Components in Meteorites

Besides isotopic anomalies apparently produced by decay of short-lived radionuclides such as ²⁶Al, ¹²⁹I, and ²⁴⁴Pu, and those associated with the interstellar grains in the matrix, other isotopic anomalies have been identified in CAI's. Some of these have been observed only in a few CAI's from *Allende*: Ca, Ba, Sm, Nd, and Sr. They have been termed FUN anomalies: fractionation and unknown nuclear. Other variations are widespread, but appear to be due to fractionation, e.g., Si and Mg. A systematic search by G. Lugmair and colleagues at the University of California at San Diego has revealed that isotopic variations of the iron peak elements (elements clustered around the cosmic abundance peak at Fe), Ca through Zn, are ubiquitous in many CAI's. Variations are up to per mil size relative to terrestrial isotope ratios, and are characterized by overabundance of the most neutron-rich nuclides (⁴⁸Ca, ⁵⁰Ti, ⁵⁴Cr, ⁶⁴Ni, and ⁶⁶Zn), often accompanied by underabundance of the most neutron-served with the ion microprobe. In general, these isotopic variations are most readily explained if the inclusions contain an admixture of neutron-rich e-process material from a supernova.

Oxygen Isotope Variations

Another element commonly showing isotopic variations is O. Until 1973, O isotope variations in meteorites were thought to be simply the result of fractionation, as they are on Earth. But when R. Clayton of the Univ. of Chicago went to the trouble of measuring ¹⁷O (0.037% of O) as well as ¹⁸O and ¹⁶O, he found that these variations were not consistent with simple mass-dependent fractionation.

CHAPTER 10: COSMOCHEMISTRY



Figure 10.26. Variation of O isotope ratios in meteorites and terrestrial and lunar samples. Most of the data from Allende inclusions is off the plot to the left. From the work of R. N. Clayton. This is illustrated in Figure 10.26. On a plot of ${}^{17}\text{O}/{}^{16}\text{O}$ vs. ${}^{18}\text{O}/{}^{16}\text{O}$, variations created by fractionation should plot along a line with slope of 1/2. Terrestrial and lunar samples do indeed define such a line, but other meteorites or their components fall along a line with slope = 1. One interpretation is that this reflects mixing between a more or less pure ${}^{16}\text{O}$ component, such as might be created by helium



Figure 10.27. Mass independent fractionation during the production of ozone from molecular oxygen. Squares are ozone, dots are oxygen. After Thiemens and Heidenreich (1983).

burning, and a component of 'normal' isotopic composition. However, Thiemens and Heidenreich (1983) conducted experiments in which ozone produced by a high frequency electric discharge showed "mass independent fractionation", i.e., where the ozone was equally enriched in ¹⁷O and ¹⁸O (Figure 10.27) relative to ¹⁶O. The experiment demonstrates that a slope of 1 on the $\delta^{17}O$ — δ^{18} O diagram could be produced by chemical processes. Thiemens suggested this kind of fractionation results from a kinetic fractionation mechanism, which arises because non-symmetric (e.g., ¹⁶O¹⁷O or ¹⁸O¹⁶O) molecules have more available energy levels than symmetric (e.g., ¹⁶O¹⁶O) molecules (as we saw in Chapter 9, symmetry enters into the calculation of the partition function). However, it is unclear how this mechanism could produce the observed fractionations, and debate still rages as the whether the O isotopic variations are ultimately of chemical or nuclear origin.

Whether the O isotope variations in meteorites are nuclear or fractionation effects, their discovery is one of the most important in the field of

Geochemistry

CHAPTER 10: COSMOCHEMISTRY



Figure 10.28. O isotope variation among minerals of various meteorite classes (Clayton et al., 1976).

cosmochemistry. What we will call 'mass independent' effects, be they nucleosynthetic or chemical fractionation, almost certainly occurred before meteorites accreted. As Figure 10.28 shows, while variations be*tween* classes are mostly mass-independent, variations within groups of meteorites fall along mass-dependent fractionation lines. This strongly suggests that, for the most part, different groups could not have come from the same parent body and that the different groups probably formed in different parts of the presolar nebula. There are a few exceptions: IIE irons fall on a mass-dependent fractionation line (MDFL) with Hchondrites, IVA irons plot on a MDFL with L and LL chondrites, basaltic and hy-

persthene achondrites plot on a MDFL with IAB irons and some stony-irons, and the Moon and the Earth plot on a MDFL together with enstatite chondrites and achondrites. This suggests a genetic relationship between these objects, perhaps derivation from a single a single parent body.

Oxygen isotope compositions of the minerals of the FUN inclusions of *Allende* are erratic and do not fall on any line. These are certainly nuclear effects.

Mass-independent effects have not been found among isotopes of other elements such as Si, S and Mg (except as noted above), though large mass-dependent fractionations are ubiquitous. Thus oxygen appears unique. The reason may be related to the partition of oxygen between the gaseous and solid phases of the presolar nebula. Over a large temperature range, only about 17% of oxygen will condense, the remainder being in the gas as H_2O and CO. Isotopic variations could arise by reaction between gas and solid of different isotopic composition. Hydrogen and carbon have only two stable isotopes so mass-dependent and mass-independent effects cannot be distinguished.

FORMATION OF THE SOLAR SYSTEM

The idea that the Solar System formed from a cloud of gas and dust is an old one and can be traced back at least to the German philosopher Immanuel Kant (1724-1804), if not to French philosopher and mathematician René Descarte (1590-1650). Like many other aspects of science, however, most of the progress on this problem has been made in the last 50 years. This progress was accelerated by two events in 1969: man's exploration of the Moon and the fall of the Allende meteorite.

There are a number of observations that allow us to place constraints on the way in which the solar system formed. These constraints include astronomical observations of star formation and the physics and chemistry of the solar system itself. Further, any successful model of solar system formation clearly must produce the distribution of mass, angular momentum, and composition, including isotopic composition, that we now observe. The chondritic meteorites, whose various components formed in and accreted from the solar nebula, are also an important constraint on models of solar system formation: a successful model of solar system formation should explain these features.

It must be stated at the outset that no model of solar system has successfully predicted all the relevant observations. This essentially means that our understanding of the process is still incomplete. In this section, we will discuss the various constraints, and outline in a general way the processes involved in solar system formation, emphasizing geochemical aspects. The geochemical goal of models of solar system formation is to predict the composition of the planets. Understanding the composition of the Earth is perhaps the most important aspect of this subject from our perspective.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

Astronomical Observations and Theoretical Models of Star Formation

Astronomical observations have established that stars form when fragments of large molecular clouds collapse. Such clouds may have dimensions in excess of 10⁶ AU[‡] and masses greater than $10^{6} M_{\odot}$ (solar masses). Gravity will tend to make such clouds collapse upon themselves. This tendency may be resisted by thermal motion that tends to expand the cloud, rotational motion that stabilizes it against collapse, and internal pressures generated by magnetic fields. A careful balance between the forces tending to collapse the cloud and forces tending to expand or collapse it can result in the cloud being stable indefinitely.

Density inhomogeneities in the cloud may develop that are gravitationally unstable. The Taurus-Auriga cloud complex is a good example of a region in which low mass stars similar to the Sun are currently forming. The cloud is about 6×10^5 AU across, has a mass of



Figure 10.29. The Great Nebula in Orion, shown in a Hubble Space Telescope photograph, is cloud of gas and dust within which stars are forming. Credit: NASA.

roughly $10^4 M_{\odot}$, a density of 10^2-10° atoms/cm³, and a temperature around 10 K. Embedded within the cloud are clumps of gas and dust with densities 2 orders of magnitude higher than the surrounding cloud. Within these clumps are luminous protostars. About 100 stars with mass in the range of 0.2-3 M_{\odot} have been formed in this cloud in the past few million years.

In addition to random density perturbations that can spontaneously collapse, two 'events' have been identified as possible triggers of cloud collapse and star formation. The first is shock waves produced by supernovae explosions. As the shock wave propagates through an interstellar cloud, it produces instability that leads to collapse of parts of the cloud. Several such examples are known. A luminous gas cloud in Canis Major is a region of active star formation. The young stars occur in an arc whose radius suggests a shock wave produced by a supernova about a million years ago. The Great Nebula in Orion (Figure 10.29) also has an approximately circular shape, suggestive of a supernova shock wave. In this case, several large supernovae would be necessary to produce the observed motion. Since stars generally form in groups and the longevity of stars depends on their mass, several large stars forming simultaneously could ultimately lead to a series of approximately 'simultaneous' (on an astronomical—geological time scale) supernovae.

The second mechanism of cloud collapse and star formation is the density wave associated with the spiral arms of the galaxy. The spiral arms of galaxies are bright because they contain many young, massive stars (since the lifetime of a massive star will be short, all such stars are astronomically young). It is thought that the arms are actually a standing density, or gravity, wave. The clustering of many stars in the arm pulls in other stars and dust clouds. Eventually, the galactic orbits of these stars and clouds will take them out of the arms, but during the time they are in the arms, they contribute to this self-perpetuating wave of gravitational attraction. We can think of the galactic

^{\pm} AU stands for Astronomical Unit, which is the Earth-Sun distance or 1.49×10^8 km.

CHAPTER 10: COSMOCHEMISTRY

arms as being similar to a traffic jam on the galactic orbital freeway. Though traffic continues to flow through the region of the traffic jam, there is nevertheless a sort of self-perpetuating high concentration of stars in the traffic jam itself. The Sun is thought to pass through an arm about once every 100 Ma on its journey around the galactic center. As clouds are pulled into the arms, they are compressed by collisions with other matter in the arms, leading to collapse of the clouds and initiation of star formation.

These two causes of dust and gas cloud collapse are not mutually exclusive. The time spent by galactic matter in the spiral arms is long enough so that very massive stars will spend their entire lives within the arms. A high concentration of massive stars in the arms means that supernovae will be more common within the arms than elsewhere. These supernovae can contribute to further star formation within the arms.

Model calculations show that once a cloud fragment or clump becomes unstable, it will collapse into a thin disk, the solar nebula, with much of the material collecting as a compact object at the center (the protostar), but much of the angular momentum being contained in the nebular disk. Such protoplanetary disks have recently been photographed by the Hubble Space Telescope (e.g., Figure 10.30). The time scale for collapse of an individual fragment to form a protostar and associated nebular disk depends on the mass of the fragment. For a star of about 1 solar mass, the time scale should be roughly 10⁵ to 10⁶ years (Cameron, 1995). For very massive stars, the time scale may be two orders of magnitude shorter; for very small stars, it may be two orders of magnitude longer. In most cases, collapse leads to multiple or binary star systems.

Cameron (1995) has divided the evolution of the solar nebula into 4 stages. The first stage is the one we have just described: collapse of a molecular cloud to form a disk-like nebula. At the end of stage 1 a small core of gas of gas and dust that will eventually become the Sun is located at the center of the nebula. One of the most controversial and problematic aspects of models of stellar nebulae is the temperatures attained in the nebula. While early nebular theories (e.g., Cameron, 1962) predicted the inner solar system reached high temperatures, subsequent ones predicted much lower temperatures, 600 K at the orbit of Earth. More recent three-dimensional models of the solar nebula pre-

dict temperatures up to 1500 K in the inner solar system (e.g., Boss, 1990), but even these more recent models are greatly simplified and may significantly under- or over-estimated nebular temperatures. Nebular mass is an important factor in the maximum temperatures reached and is not well constrained. A nebula with a mass of 0.04 M_{\odot} would reach a maximum temperature of ~1400 K at 2-3 AU (the asteroid belt), while one of 0.01 M_{\odot} would only reach a temperature of ~500 to 800 K (Boss, 1995). Temperatures decrease with distance, as illustrated in Figure 10.31. There is also a strong vertical thermal gradient, so that at 0.1 AU above the nebular midplane temperatures would fall to ~500 K.

High temperatures appear to be reached only in the early stages of nebular evolution: calculations show the nebula would cool on a time-scale of 10⁵ years (Boss, 1990). The heating is due to release of gravitational energy. The temperature obtained depends on the relative rates of heating and heat loss. The rate of heating depends on the infall rate, i.e., the rate at which gravitational potential energy is converted to thermal energy. The rate of heat dissipation depends to a considerable degree on the density and opacity of the disk. The opacity will depend particularly on the fraction of mass present as condensed matter, i.e., as dust, because dust will strongly absorb radiated energy. As temperatures in the nebula in-



Figure 10.30. Dusty Protoplanetary disk (dark region) surrounding a newly formed star (bright central object) in the Orion Nebula photographed by the Hubble Space Telescope. Credit: Mark McCaughrean (Max-Planck-Institute for Astronomy), C. Robert O'Dell (Rice University), and NASA.

CHAPTER 10: COSMOCHEMISTRY

crease, more and more of the dust will evaporate (pressures are probably quite low in the nebula, in the vicinity of 10^{-4} atm and less). As opacity decreases, more energy will be radiated away, stabilizing, or lowering, temperature. A critical temperatures occur around 1420 K where iron grains evaporate. Above this temperature, there is little dust remaining, so the nebular would be quite transparent. As a result, temperatures probably never exceed 1500 K. However, much of the material that experiences this strong heating within the inner nebula at this stage may ultimately accrete to the Sun, so the relevance of this early heating for the chemical composition of planets and meteorites has been questioned.

Another matter of controversy is the mass of the nebula. Estimates for the mass of the nebula surrounding a star the size of the Sun, range from 0.01 M_{\odot} (the mass of material with solar abundances necessary to yield the mass and composition of the planets as observed—their mass is about 0.0015 solar masses) to about 1 M_{\odot} , in excess of the mass of the Sun. Most recent models seem to favor relatively low mass nebulas. For example, Cameron (1995) uses a value that is several times the minimal value.

In Cameron's stage 2 of nebular evolution, the mass of nebula is roughly steady state, with the mass lost by accretion to the growing Sun being balanced by infall from the cloud. The Sun accretes most of its mass during this stage. One of the most difficult problems in understanding the origin of the Solar System is how the Sun ended up with most of the mass while the planets ended up with most of the angular momentum. This problem has puzzled physicists for more than a century. Indeed, it has been so vexing that some physicists have argued that the planets were formed from a captured nebula rather than one that formed with the Sun (to do so, however, requires that the chemical evidence be totally ignored). Some progress may have been made on this problem in the last few years, however. If the Sun is not located at the center of mass of the system, and if the nebula is sufficiently dense that its gravity is significant, the density waves can develop within the nebula. These density waves are much like the spiral arms of galaxies. What is important is that it appears they can transport both mass inward and angular momentum outward. These theoretical models may have a counterpart in observational astronomy. The density wave calculations indicate that mass can be added to stars in clumps. Some accreting stars can experience sudden increases in luminosity, which is

thought to be due to sudden increases in the rate at which mass is accreted to the star. Such stars are known as FU Orionis stars and their outbursts may result from density waves delivering clumps of nebular material to their surfaces.

Because of both radioactivity within the nebula (²⁶Al is probably the most important radionuclide in the nebula at this point) and radiation from the growing Sun, significant ionization may occur within the nebula. If so, magnetic fields will play an important role in nebular processes. These can also transfer angular momentum outward and contribute to radial mixing of the nebula. Another characteristic of FU Orionis stars is that they eject significant amounts of gas at high velocity normal to the plane of the nebula. These "bipolar outflows" also appear to be due to magnetic coupling between the star and the nebula. The rate of outflow is quite high, as much as 10% of the rate of accretion of material from the nebula to the Sun.



Figure 10.31. Variation of temperature with distance in a 3D computational model of the solar nebula by Alan Boss (1990).

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

Solar accretion rates are rapid during stage 2, the FU Orionis stage, and its duration is short. Cameron estimates accretion rates of 10^{-5} M_o, and as high as 10^{-4} M_o, during periods of rapid accretion, and a duration of this stage of 50,000 years. Much of the material that was in the inner Solar System (and that experienced strong heating in stage 1) is accreted to the Sun in this stage.

Cameron's third stage of nebular evolution is the T Tauri stage. T Tauri stars are young stars. They have cool surfaces (4000 K), but have luminosities several times that of the Sun. They are often surrounded dusty disks, that characterized by very enhanced solar winds directed radially outward; mass losses of $10^{-8} M_{\odot}/yr$ are typical (the present solar wind amounts to a loss of $10^{-14} M_{\odot}/yr$). Like the FU Orionis stars, there are bipolar jets, but they are smaller. The strong solar winds inhibit accretion of additional material to the Sun. At the beginning of this stage accretion rates have decreased to about $10^{-7} M_{\odot}/yr$ and they drop by another order of magnitude by the end of this stage (Cameron, 1995) The T Tauri solar winds were sufficiently strong that the wind pressure on particles with radii of a few cm and smaller would exceed the gravitational force of the Sun. As a result, the winds begin to sweep any gas and dust that has not accumulated into large particles out of the solar nebula, beginning with the region closest to the Sun and working slowly outward.

It is during this stage, which lasts several million years, that planetismals and planets form. The rate at which dust, which is largely in the submicron size to begin with, accumulates into large bodies depends strongly on a number of parameters that are poorly constrained. These include turbulence, sticking efficiency, and relative velocities. Computer simulations of the process by Cameron (1995) show that grains grow most rapidly in size when they are some distance above (or below) the nebular plane. They then settle to the plane of the nebula. The simulations also suggest that once this coagulation process begins, it proceeds very rapidly. Within time of 10³ to 10⁵ years (fastest in the region of Earth, slowest in the beyond Neptune), most of the mass of the solid material in the nebula is in meter to kilometer size bodies. According to calculations by Wetherill and Stuart (1993), planetismals only slightly larger than this (10 km radius) would accumulated into "planetary embryos", bodies the size of the Moon or Mercury, within 10⁵ years.

By stage 3, the nebula had lost most of its original heat. Cameron (1995) proposed the recombination reactions in the solar wind kept the inner solar system sufficiently warm, however, that water ice does not condense. Other processes may have kept the inner solar system warm as well. Jupiter, and to a lesser degree Saturn, may accumulate much of there mass during this stage of nebular evolution (Cameron, 1995). Indeed, the gas-rich nature of these planets appears to require that they form early, before the T Tauri winds dissipate the nebular gas. Though Jupiter is gas-rich, it is nevertheless depleted in gas relative to the Sun, so some dissipation of nebular gas apparently occurred before the assembly of the giant planets was complete. The location of the "snow line", i.e., the distance as which temperature falls below the point where water ice condenses (170 K) may have been located roughly at the orbit of Jupiter during stage 3. Sticking efficiency of grains is greatly enhanced if the grains are coated with frost (of H_2O , CO_2 , NH_4 , etc.). Thus the position of the "snow line" at the orbit of Jupiter may have accelerated the process of planet formation in this region. The early formation of a rocky planet of roughly 10 earth masses would have then been able to capture enough gas to become Jupiter.

The early formation of Jupiter probably has the effect of preventing the formation of a planet within the asteroid belt. This could occur either because Jupiter itself robs the belt of material, or because Jupiter perturbs the orbits of planetary embryos in the asteroid region so that they are ejected from the Solar System. Evolution of planetary embryos into planets in the inner solar system is not affected by Jupiter, but does take longer because of mutual perturbation to high velocity orbits Wetherill and Stuart (1993).

The Sun has completely accreted by the end of Cameron's stage 3. During stage 4, there is little additional accretion of material to the nebula, which is slowly dissipated by photoevaporation beyond the orbit of Saturn and from inside out by the T Tauri solar winds. As the solar winds drive out the nebula, the inner edge of the remaining nebula, called to bow, becomes a location where formation of planetismals is particularly likely as particles are driven to the bow from both directions. The es-

CHAPTER 10: COSMOCHEMISTRY

timated time for this stage is 3 to 30 million years. By the end of this period, the nebula has dissipated and planet formation is well underway and perhaps complete.

FORMATION OF CHONDRITIC METEORITES AND PROCESSES IN THE SOLAR NEBULA

After long debate, it is now recognized that compositional variation between classes of chondrites primarily (though not necessarily exclusively) reflects differences in conditions and composition of the nebula in the region where the various parent bodies formed rather than metamorphic processes occurring within parent bodies. Thus the composition and petrography of chondrites provide an important constraint on conditions and processes in the solar nebula.

It is now generally agreed that planetary bodies did not condense from a hot solar nebula, rather they accreted from a cool one. Nevertheless, much of the chemical variation in chondrites appears to be related to volatility. The inner part of the solar nebula may have reached high enough temperatures for much of the dust to sublime. Although this episode appears to have preceded formation of planetismals and planets by a considerable period, some of the material that later accreted to chondrite parent bodies may have experienced high temperatures in the early nebula. The nature and composition of CAI's and chondrules within meteorites are clear evidence that there were transient local heating events within the nebula. For these reasons, it is useful to consider the sequence in which the elements would condense from a hot nebular gas.

The Condensation Sequence

The theoretical condensation sequence has been calculated from thermodynamic data, largely by Larimer, Grossman, and Lewis, all of whom have worked with E. Anders at the University of Chicago (e.g., Larimer, 1967; Grossman, 1972). The condensation temperature of an element reflects its vapor pressure, its tendency to react with other elements to form compounds in the gas or solid solutions or alloys in the solid, and its abundance in the gas. Let's consider two examples of condensation sequence calculations. First, consider iron, which is a particularly simple case since it condenses as Fe metal:

$$\operatorname{Fe}_{(g)} \to \operatorname{Fe}_{(s)}$$
 10.1

Assuming ideality, the partial pressure of iron is simply its mole fraction in the gas times the total pressure (P_T). Since hydrogen is by far the dominant element in the gas, the mole fraction can be approximated as 2 times the solar Fe/H ratio (the 2 arises from hydrogen's presence as H₂). Thus the partial pressure of Fe is written as:

$$p_{\text{Fe}} = \frac{[\text{Fe}]_{\odot}}{1/2[\text{H}]_{\odot}} \times P_{\text{T}}$$
 10.2

where $[Fe]_O$ and $[H]_O$ are the solar abundances of Fe and H and P_T is total pressure. Once condensation begins, we can express the equilibrium constant for this reaction as the ratio of the partial pressure of Fe in the gas to the concentration in the solid:

$$K = \frac{p_{Fe}}{[Fe]_s}$$
 10.3

where $[Fe]_s$ is the concentration in the solid and $[Fe]_g$ is the concentration in the gas. Condensation begins when the partial pressure of Fe exceeds the vapor pressure of solid Fe. Since:

$$G = -RT \ln K$$
(3.98)

the equilibrium constant can also be written as:

$$\ln K = -\frac{H_V}{RT} + \frac{S_V}{R}$$
 10.4

 ΔH_v and ΔS_v are the enthalpy and entropy of vaporization. Once condensation of an element begins, its partial pressure drops by $(1 - \alpha)$ where α is the fraction condensed. Hence, equation 10.2 becomes:

$$p_{\text{Fe}} = \frac{(1 - \alpha) [\text{Fe}]_{\text{O}}}{1/2[\text{H}]_{\text{O}}} \times P_{\text{T}}$$
 10.5

November 25, 1997

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

Combining 10.3, 10.4, and 10.5, the equation describing condensation is:

$$\ln (1 - \alpha) = -\frac{H_V}{RT} + \frac{S_V}{R} - \ln \frac{[Fe]_O}{1/2[H]_O} - \ln P_T + \ln [Fe]_s$$
 10.6

Now consider elements such as Mg and Si, which form compounds in the gas and solid, which complicates the calculation considerably. The reaction for the condensation of forsterite is:

$$Mg_{(g)} + SiO_{(g)} + 3H_2O \rightarrow Mg_2SiO_{4(s)} + 3H_2$$
 10.7

The equilibrium constant for this reaction is:

$$K = \frac{a_{Mg_2SiO_4} p_{H_2}^3}{p_{Mg}^2 p_{SiO} p_{H_2O}^3 P_T^3}$$
 10.8

The first step is to compute partial pressures of the gaseous species. For example, for SiO and H_2O above, we may write reactions:

$$2H + O \rightleftharpoons H_2O$$
 and $Si + O \rightleftharpoons SiO$

and calculate equilibrium constants for them from the free energies of these species, e.g.:

$$K = \frac{p_{SiO}}{p_{Si} p_{O}}$$
 10.10

Values of K can be computed from thermodynamic data. For each element, an additional constraint is imposed by the total abundance of that element in the gas. Thus for example:

$$[SiS] + [Si] + [SiO] + [SiO_2] = [Si]_{\odot} - [Si]_{S}$$
 10.11

where $[Si]_s$ is the total of silicon in condensed phases. Combining equilibrium constant equations such as 10.10 with mass balance equations such as 10.11 leads to a series of simultaneous equations. These can be solved by successive approximation using a computer. Values of equilibrium constants such as 10.8 can then be computed from Δ H and Δ S using equation 10.4.

Further complications arise when solid solutions form. For example, the forsterite term in 10.7 is 1 if forsterite is treated as a pure phase. If the solid solution with fayalite is considered, then the value of the activity must also be calculated. Values for activity coefficients are difficult to obtain, and ideal solid solution is generally assumed. Thus the activity of forsterite in 10.8 would be replaced by its mole fraction. Solid solution results in the condensation of an element at a higher temperature than if a pure component were the condensed phase.

Some elements, such as Au and Cu, will condense primarily either by reaction with already condensed Fe metal grains, or by condensing with Fe metal. At lower temperatures (around 670 K for $P_T = 10^4$ atm), the Fe metal will react with H₂S gas to form FeS. Moreover, from the onset of condensation a small but increasing amount of the Fe will react with H₂O gas to form FeO that dissolves in the silicates. A marked increase in the Fe content of silicates occurs around 400-500 K.

Figure 10.32 shows 2 theoretical sequences calculated by Larimer (1967) and subsequently refined by Grossman (1972). In the 'fast cooling' sequence, matter does not react with nebular gas after it has condensed. At any given time and temperature, the solid phases are a mixture of material condensed over a range of temperature. In the 'slow cooling' sequence, condensed material continually reacts and re-equilibrates with the gas as temperature drops, so that at any time, only an assemblage in equilibrium with the gas at that temperature is present. Figure 10.33 illustrates the minerals expected at a given temperature somewhat more clearly. The condensation sequence depends critically on total pressure and H pressure; the sequence shown is for relatively low total and H pressure. At relatively high pressure, metallic Fe liquid is the first phase to condense.

In a nutshell, the sequence goes like this. The first minerals to condense would be Re and the most refractory of the platinoid metals (Os, Ir, Ru), which would condense as metallic phases. Since these are extremely rare elements, they would likely form very small grains. Interestingly, small nuggets of such metal, called "Fremdlinge" (the German word for "strangers"), are found as inclusions in CAI's. There has been some speculation that they may have served as nuclei for the CAI's. Following this would be condensation of oxides and silicates of Ca, Al and Ti. They should be rich in refrac-

CHAPTER 10: COSMOCHEMISTRY

tory trace elements such as U, Th, Zr, Ba and the REE. This match closely the composition of the CAI's, suggesting the possibility that CAI's are high temperature condensates. The CAI's have about 20 times the 'chondritic' abundances of the refractory elements, suggesting they reflect 1/20 = 5% condensation.

In detail, the CAI's are a varied lot. Some may indeed represent early condensate, but most are probably residual: i.e., the remaining solid after 95% of the dust has evaporated. There is evidence that some CAI's were once liquid, though this is difficult to reconcile with pressures in the nebula, which should have been too low for such liquids to be stable. Some CAI's show evidence of several cycles of evaporation and condensation.

Next in the condensation sequence should come metallic Fe-Ni and compounds richer in the moderately refractory elements such as Mg and Si: olivines and pyroxenes. If the cooling takes place under equilibrium conditions the high temperature assemblage should react to form anorthite as well, and at lower temperature when Na condenses, plagioclase. These phases are the ones that predominate in chondrules, with the important caveat that chondrules are metal poor. Since these phases condense at temperatures similar to Fe-Ni metal, some process must have separated metal from silicates before formation of the chondrules.

The Fe should also largely react out to form more Fe-rich olivine and pyroxene. At lower temperature, S condenses and reacts with Fe to form sulfides. At even lower temperature, the Fe reacts with O to form magnetite and the silicates react with water in the gas to form hydrated silicates. Sulfates, carbonates and organic compounds will also form around these temperatures.

If equilibrium conditions prevail, only the last named compounds would exist when condensation was complete, but all might exist if disequilibrium prevails.

The CI chondrites seem to be very similar in composition and petrography to equilibrium condensates (down to 300 K or so), or, more accurately, accretions of equilibrium condensates. The other carbonaceous chondrites approximate to varying degrees aggregates of disequilibrium condensates. In particular, the CV and CO's chondrites contain both the highest temperature condensates (CAI's) and lowest temperature material (hydrated silicates) and much of everything expected to condense in



Figure 10.32. Condensation sequence of a gas with solar composition. In slow cooling diffusional equilibrium is assumed. Fast cooling leads to disequilibrium and condensation of pure elements and compounds.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY



Figure 10.33. Simplified mineralogical condensation sequence.

silicates and FeS are classed as *moderately volatile*, and those condensing below this temperature are classed as *highly volatile* (Larimer, 1988). Those elements generally found in metal phases in meteorites are classed as *siderophile*. This classification is illustrated in Figure 10.34. There is a some similarity to Goldschmidt's classification that we discussed in Chapter 7. The refractory elements are mostly *lithophile*, the highly volatile elements include all the *atmophile* elements, and the moderately volatiles include many *chalcophile* elements. The similarity results in part because Goldschmidt used the distribution in meteorites as part of the basis for his classification.

Anders noted that the chemistry of the carbonaceous chondrites can be approximated as two component mixtures: a volatile-rich matrix and volatile-poor chondrules and CAI's. CI chondrites consist only of matrix, while CV chondrites contain ~30% chondrules and CAI's, and are ~40% depleted in volatiles as a result. This mixing hypothesis can also be applied to ordinary chondrites, though less well. Here the cosmochemical grouping comes more into play. Figure 10.35 illustrates how several of the chondrite groups can be explained as mixtures of various identifiable components. The refractory elements are associated with chondrules and CAI's. The siderophile elements are, of course, associated with Fe metal. The association of the moderately volatile elements is somewhat more complex: those with chalcophile tendencies are associated with FeS, those with lithophile tendencies (e.g., the alkalis) are associated with silicates in the matrix. The highly volatile elements are associated exclusively with the matrix. Thus much of the compositional variation in chondrites appears to

between. CV and CO are depleted in the more volatile elements and compounds (such as water, but also the moderately volatiles such as Ga and Ge, alkalis).

The Cosmochemical Classification of the Elements

The chemical variations in meteorites seem to be dominated by two processes: (1) condensation and evaporation and (2) oxidation and reduction. This leads to a natural classification of the elements into 4 groups based on how they are affected by these processes. The division relative to volatility is based on the condensation sequence discussed above. There are two benchmarks: the condensation of the Mgsilicates and Fe-Ni metal at 1300-1400 K (depending on pressure) and the condensation of FeS at 670 K. Those elements with condensation temperatures higher than or equal to the former are classed as *refractory*, those condensing between the Mg-

CHAPTER 10: COSMOCHEMISTRY



Figure 10.34. The cosmochemical periodic table (after Larimer, 1988).

result from mixing of different components that formed under different conditions in the solar nebula. However, the concentrations of the most volatile elements and compounds (e.g., water) vary between petrographic types, whereas no variation with petrographic type occurs for the moderately volatile elements. Variation in the moderately volatiles may be nebular in origin, whereas some of the variation in the highly volatiles may reflect parent body metamorphism.

Constraints on Conditions in the Solar Nebula

In summary, most of the compositional variation between classes of chondritic meteorites falls into 4 categories: variation in the amount of refractory material (as evidence by, for example, Al/Si ratios), variation in the content of volatile material such as water, variation in the amount of sidero-



Figure 10.35. Ordinary chondrites as mixtures of nebular components. The CI chondrites contain 'solar' abundances of the condensable elements, all contained in the matrix. From Larimer (1988).

phile material, and variation in oxygen fugacity. The nature of meteorites place on the formation of the solar system, which are summarized below. These constraints are not all mutually consistent, and cannot all be explained in the context of our present understanding of solar system formation.

- The evidence for pervasive ¹²⁹I suggests nucleosynthesis no more than 10⁸ years before the solar nebula collapsed. This is rather a long time on the scale of galactic processes, so it would be perhaps more surprising if ¹²⁹I were absent.
- The evidence of ²⁶Al in CAI's indicates that newly synthesized material was injected into the molecular cloud that ultimately formed the Solar System not more than about 10⁷ years before the CAI's formed. This is supported by evidence of the existence of ¹⁰⁷Pd and ⁵³Mn, as well as the recent discovery of evidence of ⁴¹Ca (half-life: 10⁵ yr). The most likely site of synthesis of this material is red giant stars.

CHAPTER 10: COSMOCHEMISTRY

- The matrix of carbonaceous chondrites contains a collection of interstellar dust. Most of this material was apparently produced well before the solar system formed. This dust includes diamonds that are either the ash of a supernova explosion, or were implanted with Xe produced in a supernova. That this material has retained its distinct noble gas isotopic signature indicates it was never heated above ~700 K.
- The pervasive presence of chondrules requires that temperatures in the nebula locally reached 1700 K or more. However, chondrules clearly cooled rapidly (time scales of an hour or less), so that these high temperatures were transient. At least locally, much of the nebular dust would have been processed through such events, since chondrules can constitute up to 80% of the mass of some meteorite classes.
- CAI's also require high temperatures and suggest that, at least locally, 95% of the nebular dust was evaporated. Studies have shown that a few of the CAI's also cooled fairly rapidly (0.5-50 K/hr), though others may have remained at high temperature for longer periods. They may have formed in the inner solar system where temperatures were high. Regardless of whether CAI's are condensates or refractory residues, some process must have separated them from the gas before cooling, otherwise lower temperature components should have condensed as well.
- Most CAI's show ²⁶Mg anomalies, indicating they formed while ²⁶Al still existed. With rare exceptions, chondrules do not show ²⁶Mg anomalies, indicating they formed after ²⁶Al had decayed. From this we may conclude that the CAI's formed several million years before the chondrules.
- Planetismals that were the parent bodies of achondrites and irons underwent sufficient heating to melt and differentiate. The oldest achondrites are do not appear to be significantly younger than many chondrites. Thus planetismals formed, melted, and differentiated within a few million years of the formation of chondrites.
- Oxygen isotope ratios vary between essentially all meteorites and planetary bodies, implying that the entire nebula was never entirely homogenized, and therefore was never entirely gaseous for a significant length of time. Despite the oxygen isotope heterogeneity, and isotope anomalies in rare interstellar grains, the isotopic composition of most elements investigated in homogeneous throughout all classes of meteorites. This dichotomy has yet to be satisfactorily explained.
- Variations in the chemical composition of chondrites clearly indicates chemical inhomogeneities within the solar nebula.
- Oxygen fugacity clearly varied within the solar nebula. Since the gas is dominated by H and O constitutes a significant fraction of condensed matter, this may reflect variation in the ratio of gas to dust. This could have been caused by a number of processes, such as turbulent eddies or differential settling of grains to the nebular mid-plane. Variation in siderophile content of meteorites could also preferential transport of metal grains due to differences in densities or magnetic properties.

SUMMARY: RECIPE FOR CHONDRITIC METEORITES

We have seen that chondrites are a mix of a variety of materials formed under different conditions, probably in different places in the solar nebula. Figure 10.36 is a cartoon illustrating the processes involved in formation of chondrites and their components. Most CAI's seem to be refractory residual solids heated to ~1700 K. This heating may have occurred in the inner part of the solar nebula and was probably transient and local. Chondrules, on the other hand, experienced melting as a result of rapid heating and subsequent quick cooling. The matrix represents the nebular dust, some, or all of which never experienced severe heating.



Figure 10.36. Cartoon illustrating the processes involved in the formation of chondrites and their components. Modified from McSween (1987).

FORMATION OF THE PLANETS

Table 10.8 lists the some data regarding the planets. Many planetary scientists do not consider Pluto a planet: it is better than an order of magnitude smaller than Mercury and smaller than some of the major satellites of Jupiter, Saturn, and Neptune, and has a very anomalous orbit. On the other hand, the Moon is so large, nearly a quarter of the size of Mercury, that it is sometimes counted as a planet. The planets can be divided into three groups based on their size, density and composition. These are:

• The terrestrial planets:

Mercury Venus Earth-Moon

- Mars
- The giant planets:
 - Jupiter
 - Saturn
- The outer icy planets:
 - Uranus
 - Neptune
 - (Pluto-Chiron)

The terrestrial planets consist primarily of silicates and Fe-Ni metal, the giant planets consist primarily of H and He, and the icy planets consist of outer gaseous shells with mantles of ices of H_2O ,

CHAPTER 10: COSMOCHEMISTRY

 CH_4 , and NH_3 and silicate-metal cores. In a gross way, this compositional pattern is consistent with radial decrease in nebular temperature (e.g., Figure 10.31): the terrestrial planets lost, or never had, their compliment of the highly volatile elements (e.g., H, He) and are also depleted in moderately volatile elements (e.g., K, Pb). From what can be judged from reflectance spectra, the asteroids also fit this pattern: the inner asteroids (sunward of 2.7 AU) are predominantly igneous and compositionally similar to the achondrites, which are highly depleted in volatile and moderately volatile elements. The outer asteroids (beyond 3.4 AU) appear to be similar to carbonaceous chondrites.

Comparing the volatile inventory of Mars with that of Earth, Mars at first appears depleted in volatile elements. It has a much smaller atmosphere than the Earth (surface pressures are 0.006 atm). Like Venus, the Martian atmosphere is dominated by CO_2 , with N_2 as the second most abundant component. Judging from the presence of erosional features on the Martian surface, it appears that liquid water once existed on the Martian surface. To attain the necessary temperatures, Mars would have had to have had CO_2 pressures at its surface of 5 to 10 atm. This early atmosphere has been lost, a consequence of the smaller gravitational field of Mars. Thus the depletion of highly volatile elements on Mars may be a secondary feature.

There is evidence that Mars in richer in moderately volatile elements than the Earth. Both analyses of Martian soil and the composition of SNC meteorites suggest a K/U ratio about twice that of the Earth. Sr-Nd isotope systematics (Figure 10.37) of SNC meteorites define an array shifted to higher ⁸⁷Sr/⁸⁶Sr ratios compared with that of the Earth, implying a higher Rb/Sr ratio for Mars. Pb isotope ratios of these meteorites indicate a ²³⁸U/²⁰⁴Pb ratio of 5 for Mars, compared with the terrestrial value of ~8. Thus Martian moderately volatile/refractory element ratios appear to be systematically higher than terrestrial ratios.

The comparison of Venus with the Earth is particularly interesting. Though the two planets are of similar size, the Venetian atmosphere is almost 100 times more massive than that of the Earth.

	Mass	radial distance	radius	density	1 atm densi	ty Principal
	(kg)	(AU)	(km)	(g/cc)	(g/cc)	atmospheric components
Sun	1.99×10 ³⁰		6.96×10 ⁵	1.4		
Mercury	3.35×10 ²³	0.39	2.44×10^{3}	5.42	5.3	_
Venus	4.87×10^{24}	0.72	6.05×10^3	5.24	3.95	CO ₂ , N ₂ , Ar
Earth	5.98×10^{24}	1.0	6.38×10^{3}	5.52	4.03	N ₂ , O ₂ , Ar
Moon	7.35×10^{22}		1.74×10^{3}	3.3	3.4	_
Mars	6.42×10^{23}	1.6	3.39×10^{3}	3.93	3.7	CO ₂ , N ₂ , Ar
asteroids	4×10^{21}	2.8	≤10 ³			
Jupiter	1.90 ×10 ²⁷	5.2	6.99×10 ⁴	1.31		H, He
Io	8.63×10^{22}		1.82×10^{3}	3.42		
Europa	4.71×10^{22}		1.55×10^{3}	3.03		_
Ganymede	1.51×10^{23}		2.63×10^{3}	1.98		_
Callisto	1.06×10^{23}		2.40×10^3	1.83		_
Saturn	5.69×10^{26}	9.6	5.95×10^{4}	0.69		H, He
Titan	1.38×10^{23}		2.58×10^{3}	1.88		N ₂ , CH ₄
Uranus	8.73×10^{25}	19.1	2.54×10^{4}	1.30		H, He, CH ₄
Neptune	1.03×10^{26}	30.8	2.13×10^{4}	1.76		H, He, CH ₄
Triton	2.14×10 ²² (?)	1.35×10^{3}	2.08		_
Pluto	2.88×10^{22}	39.4	1.15×10^{3}	~2.0		—
Chiron	1.74×10^{21}		5.93×10^{2}	~2.0		—
Anhydrous Ch	ondrites				3.4-3.9	

TABLE 10.8. DATA REGARDING THE SUN, THE PLANETS, AND THEIR MAJOR SATELLITES

Geochemistry

CHAPTER 10: COSMOCHEMISTRY



Figure 10.37. Comparison of terrestrial and martian Sr-Nd isotope systematics. The martian array is based on the assumption that SNC meteorites come from Mars. Mars appears to have higher ⁸⁷Sr/⁸⁶Sr and therefore higher Rb/Sr. After Taylor (1992).

that of Earth is about 300. Since 40 Ar is produced by radioactive decay, it is related to the planetary K/Ar ratio. The terrestrial 40 Ar/ 36 Ar is high because the Earth captured relatively little primordial Ar; most of the Ar in the terrestrial atmosphere has been produced by decay of 40 K. Apparently, Venus captured much more primordial Ar and has a much lower K/Ar ratio than the Earth. Thus on Venus, the amount of 40 Ar produced by decay of 40 Ar.

Although the Earth is depleted in noble gases relative to Venus, it appears to be richer in H_2O (Prinn and Fegley, 1989). Furthermore, the Earth and Venus appear to have similar K/U ratios, implying similar depletions in the moderately volatile elements (Taylor, 1991).

Thus in detail, we find the compositional differences between planets cannot simply be explained by radially decreasing temperature. The noble gas-rich nature of Venus is just one example. Another example is Mercury. Considering its density, Mercury appears to be much richer in iron than the other terrestrial planets, which is difficult to explain simply a temperature effect (since Fe condenses temperatures similar to Mg silicates). In the outer Solar System, we find that although Uranus and Neptune are further from the Sun, they are poorer in H and He than Saturn and Jupiter. Jupiter and Saturn appear to have H/C and N/H ratios that are about a factor of 2 lower than the solar values, and thus only modestly depleted in H. C/H ratios in Uranus and Neptune, however, are 25 times the solar value. Interestingly, while Uranus and Neptune have solar H/He ratios, the atmospheres of Jupiter and Saturn are depleted in He relative to H. However, this probably reflects internal fractionation of these elements within these planets. He becomes saturated in the H₂-dominated fluid that in outer part of Jupiter, exsolves and sink in the metallic H core, where it is more soluble.

Thus factors other that radial decrease in nebular temperature appear to also have been involved in creating the variety of chemical compositions we see in the Solar System. Some of these factors appear to be at least partly understood. For example, the compositional differences between the giant and icy planets may have been a matter of timing. The cores Jupiter and Saturn apparently formed earlier than those of Uranus and Neptune and thus were able to capture nearly their complete share of nebular gas. Uranus and Neptune apparently formed after much of the gas had dissipated.

The details of the accretion process may also have played an important role in governing the final composition of the planets. As we shall discuss below, there is something close to a consensus among planetary scientists and cosmochemists that the Moon formed as a result of a giant impact of a very large body on the Earth. Such an impact is consistent with current models of planetary accretion, which predict that the final stages of their assembly involve the collision of large planetary em-

Whereas the Venus's atmosphere is dominated by CO₂, that of the Earth is dominated by N₂ and O₂. However, the differences in both atmospheric mass and the abundance of CO₂ may reflect the difference in geological processes, and biological processes, on the two planets. Both planets appear to have similar relative abundances of carbon and nitrogen (Prinn and Fegley, 1989). In the case of the Earth, however, most of the surficial carbon is locked up in carbonates and organic carbon in rocks. This is in part due to biological activity, as is the presence of O_2 in the terrestrial atmosphere. There are, however, other differences that are more fundamental. The Venetian atmosphere is richer in noble gases than is the terrestrial one. Furthermore, the ⁴⁰Ar/³⁶Ar ratio of the Venetian atmosphere is about 1.15, whereas

CHAPTER 10: COSMOCHEMISTRY

bryos (e.g., Wetherill, 1990). The energy released in the hypothesized giant impact is enormous could have been sufficient to melt much of the Earth, and much of its volatile elements may have been lost as a result. In contrast, the accretionary histories of Venus and Mars were apparently such that they retained much of their volatile inventory (although the slow backward rotation of Venus suggests it also suffered a late impact). In this respect, we note that the Earth with a satellite that is a significant fraction of its own size is unique. Venus has no moon and the moons of Mars are merely captured asteroids.

A large impact may also explain the peculiar composition of Mercury. Mercury appears to consist of a large Fe-Ni core and a relatively thin silicate mantle. Benz et al. (1988) proposed that Mercury formed with a much larger silicate mantle, but that much of this mantle was lost as a result of a collision with a body roughly 20% the size of the present Mercury.

In summary, the temperature gradient in the solar nebula, the timing of planet formation, and accretionary histories may all play a role in determining planetary composition. This is not to say any real understanding of planet formation has been achieved. How much of the volatile depletion of the terrestrial planets should be attributed to their accretionary history and how much to high temperatures in the inner solar nebula remains unclear. And while there is a consensus that the gas-rich nature of Jupiter and Saturn is due to their formation before the nebular gas was lost, we can only guess as to why they formed early. The inner Solar System heat source is also a mystery. The heat generated by collapse of a molecular cloud should have dissipated early, before T-Tauri winds swept out the nebular gas. If so, the moderately volatile elements should have recondensed. Perhaps planetismal formation got underway while the inner Solar System was still hot. If so, volatiles would not have been accumulated by them. Any recondensed moderately volatiles elements may have remained in small particles that were swept out by later T Tauri winds. The available evidence suggests, however, that planetismals did not form until later. ²⁶Mg anomalies suggest an interval of several million years between the formation of CAI's, the oldest known "event" and chondrule formation, which probably preceded planetismal formation. This should have been enough time for the inner Solar System to cool to the point where moderately volatile elements would condense. At the least, the absence of ²⁶Mg anomalies in achondrites indicates ²⁶Al was effectively extinct by the time the parent bodies of these meteorites differentiated, implying it occurred millions of years after CAI formation. This observation also eliminates ²⁶Al as a heat source for melting and differentiation of achondrite parent bodies.

Perhaps there was another heat source. Radiation and solar wind from the Sun during its T Tauri stage have been suggested. But while this may be sufficient to keep the inner Solar System warm enough to prevent ice from condensing, it seems considerably less likely it would be sufficient to keep elements such as K from condensing. Rubin (1995) argues that collisional heating was sufficient to produce the metamorphism of meteorite parent bodies. This requires that these bodies be large. If so, they one must ask how the carbonaceous chondrite parent bodies, which apparently include large asteroids such as Ceres, avoided such heating.

Clearly, much remains to be learned about the early history of the Solar System. Progress has, however, been particularly rapid in the last 10 or 20 years. This progress in part been technologically driven, as astronomers work with increasingly powerful telescopes capable of looking at the full electromagnetic spectra, theorists obtain ever more powerful computers, and geochemists develop more powerful analytical tools. Yet the fall of the Allende meteorite in 1969, with its abundant and large CAI's, and man's exploration of the Moon beginning in the same year were perhaps the sparks for this revolution. Hopefully, exploration of other planetary bodies in the future, as well as continued technological advances, will help move us toward the solution to the difficult problems that remain.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

The Moon: Its Chemistry, History and Relation to Earth

The Moon is the one planetary sized body about whose early history we have gained some understanding. This history may have been partly shared by Earth, and is therefore worth considering. A consensus about the origin of the Moon may be emerging 25 years after the Apollo landings. Most investigators now believe the Moon originated as the result of the impact of a large body, similar in siz, perhaps to Mars on the Earth. This event probably occurred near the end of the main accretionary period.

THE GIANT IMPACT Hypothesis

The impact hypothesis was initially suggested to explain the angular momentum, which is much greater than that of other planets, and other "anomalous" features of the Earth-Moon system, including the relatively large mass of the Moon and the high inclination of its orbit (Hartmann and Davis, 1975). It also is able to explain many of the unusual compositional features of the Moon.

COMPO						
	Bulk	Bulk	Silicate	Silicate		
	Earth	Moon	Earth	Moon		
SiO ₂	30.38	43.4	45.0	44.4		
TiO ₂	0.14	0.3	0.201	0.31		
Al_2O_3	3.00	6.0	4.45	6.14		
FeO	5.43	10.7	8.05	10.9		
MgO	25.52	32	37.8	32.7		
CaO	2.40	4.5	3.55	4.6		
Na ₂ O	0.24	0.09	0.36	0.09		
K ₂ O	0.02	0.01	0.029	0.01		
Fe	28.43	2.166	87.5			
Ni	1.75	0.134	5.4			
S	1.62		5			
core%	32.5%	2.3%				
mantle	67.5%	97.7%				

Terrestrial composition from McDonough and Sun (1995), assuming the light element in the core is S. Lunar composition from Taylor (1992).

Table 10.9 compares the composition of the Earth and the Moon. There are several notable differences. First, the Moon is depleted in moderately volatile elements (Na, K) compared to the Earth. Second, although the silicate part of the Moon (mantle + crust) is richer in iron than the silicate Earth, the bulk lunar composition is poorer in iron that the terrestrial one. This reflects the small size of the lunar core, only about 2% of the mass of the planet. The lunar depletion in Fe compared to the Earth extends to all siderophile and chalcophile elements as well. The Moon is also much more depleted in the highly volatile elements as well. In the words of Newsom and Taylor (1989), it is "bone dry". Conversely, the Moon is enriched the refractory lithophile elements, having nearly twice the Al and Ca as the Earth. Despite these compositional differences, the Moon and Earth have the same oxygen isotope composition, which strongly suggests they formed from the same part of the solar nebula.

Fairly detailed models of the impact have been produced (e.g., Newsom and Taylor, 1989). The model that best explains the features of the Earth-Moon system can be best summarized as follows. During the latest stages of accretion, the Earth was struck at low angle, and relatively low velocity (5 km/s) by a body slightly larger than Mars. The impactor presumably shared the depletion in highly and moderately volatile elements that characterizes the terrestrial planets. Metallic cores had already formed in both the Earth and the impactor. The collision disrupts the impactor and partially disrupts the Earth's mantle, and much of the resulting debris goes into orbit around the Earth. Most of the disrupted core of the impactor quickly (a matter of hours) accretes to the Earth and the silicate material in orbit more slowly coalesces to form Moon. About 85% of the material forming the Moon is derived from the impactor, the remainder from the Earth's mantle. A small core then segregates from the largely molten Moon. Most of the core of the impactor quickly sinks through the disrupted mantle of the Earth to coalesce with the existing terrestrial core. The depletion of siderophile elements in the Moon is thus a result of its formation from the silicated portions of the impactor and Earth, in which the siderophile were already depleted by core formation. The volatile element depletion of the Moon is a consequence of the evaporation of these elements during impact.

TADLE 10.9.	Сомра	rison of	тне	
Compositio	ns of th	ε Ελστμ	AND	MOON

CHAPTER 10: COSMOCHEMISTRY

	Highland Anorthosite	Highland Anorth. Gabbro	Mare Qz Tholeiite	Low-K Mare Basalt	Fra Mauro KREEP	
SiO ₂	44.3	44.5	46.1	40.5	48.0	
TiO_2	0.06	0.39	3.35	10.5	2.1	
$Al_2\bar{O}_3$	35.1	26.0	9.95	10.4	17.6	
FeO	0.67	5.77	20.7	18.5	10.9	
MnO		_	0.28	0.28		
MgO	0.80	8.05	8.1	7.0	8.7	
CaO	18.7	14.9	10.9	11.6	10.7	
Na ₂ O	0.8	0.25	0.26	0.41	0.7	
K ₂ O		—	0.07	0.10	0.54	
P_2O_5		_	0.08	0.11		
Cr_2O_3	0.02	0.06	0.46	0.25	0.18	
Total	100.5	99.9	100.3	99.7	99.4	

Table 10.10. Representative Compositions of Lunar Rocks

Geology and History of the Moon

The Moon can be divided into three geologic provinces: the highlands, mountainous regions appar-



Figure 10.38. Rare earth patterns of representative lunar rocks. From Taylor (1975).

ently consisting largely of anorthosite, the uplands, areas of mild relief covered by a blanket of ejecta from the large impacts, and the Mare, the large craters which have been filled with basalt. Much of the surface of the Moon is covered with fine debris of impacts, called the regolith, consisting of rock and mineral fragments, glass, and some meteorite particles. For the most part, it seems to be locally derived; thus the regolith in the Mare differs from that of the highlands, though large impacts would have showered debris over large regions. Basalt from the Mare encompasses a variety of magma types, including both incompatible-element rich and incompatible-element poor types and both quartz-normative and olivine normative tholeiites. Highland rocks include anorthosite (nearly monomineralic calcic plagioclase), anorthosite gabbro (plagioclase and pyroxene with lesser amounts of olivine), dunite, and The highlands are ex-K-rich basalts. tremely brecciated; most of these rock types have been found only as clasts in breccias. Table 10.10 shows some representative compositions of lunar rocks, and Figure 10.38 shows rare earth patterns of the same rock types.

Most of the lunar Mare are thought to have been created between 4.2 and 3.8 Ga. Preexisting large impact craters were de-
CHAPTER 10: COSMOCHEMISTRY

stroyed in this time period. Subsequently, the Mare were flooded by basalt to a depth of 5 to 10 km. These were partial melts generated at 100 or so km depth. The flooding occurred over an extensive time: 3.9 to 3.1 Ga. Mare flooding was the last major lunar geologic event. Subsequent to that time, the only activity has been continual bombardment by meteorites and asteroids, which continued to produce minor disruption of the surface and build up of the regolith, and rare volcanism.

Figure 10.39 illustrates the highlights of lunar history. The oldest lunar rocks are nearly as old as meteorites. These are clasts of anorthositic highland rocks found in breccias. This suggests lunar differentiation began about 4.5 billion years ago. But most highland rocks have ages between 3.9 and 4.0 Ga. For the most part, these ages are interpreted as (and sometimes can be shown to be) the time of impact metamorphism. Apparently the Moon suffered very heavy bombardment by meteorites and asteroids as late as 3.9 Ga. The clustering of ages around this time gives rise the the idea of a 'late heavy bombardment' of the Moon. Alternatively, the age pattern could have been produced by decreasing impact rate between 4.5 to 3.9 Ga. According to this hypothesis, impact rates were so high in the first 500 Ma, that the probability of a rock surviving without having its age "reset" by impact metamorphism would be very low. Thus only a fraction would survive to record older ages. Around 3.9 billion years, the impact rate dropped to the point where the probability was such that most rocks would escape further impact metamorphism. Whether there was a 'late heavy bombardment' is still disputed.

One of the more important conclusions derived from study of the Moon is that it underwent very extensive melting just after its formation, perhaps forming a magma ocean 100 km deep. It is also possible that it was entirely molten, the energy supplied by the giant impact would have been sufficient to accomplish this. The anorthosite

Time B.P (Ga) 0 0.2 0.4 Low Cratering Rate; 0.6 **Regolith Formation** 0.8 1.0 3.0 Mare Flooding 3.2 Extrusion of 3.4 Mare Basalts 3.6 3.8 Mare Formation Heavy Bombardment 4.0 Igneous Activity and Extensive Metamorphism of Highlands 4.2 Initial Differentiation 4.4 of Crust Formation of Moon 4.6 Figure 10.39. Highlights of lunar chronology

of the highlands is thought to have originated by plagioclase flotation in the magma ocean, i.e., anorthosite icebergs. The lunar crust seems to have been largely in place within 100 to 200 Ma after the Moon's formation. Fractional crystallization of this magma ocean and floatation of plagioclase accounts for the general Eu depletion (Figure 10.38) observed in basalts derived from the lunar mantle. A particularly incompatible-element enriched basalt, called KREEP (K, REE, P: potassium, rare earth elements and phosphorus enriched) is thought to reflect extreme differentiation of the magma ocean.

There is almost complete agreement that extensive melting on the Moon occurred very early. Given this, and the much greater amounts of energy released by accretion of the Earth, it is hard to see how the Earth would not also have experienced extensive melting. The energy supplied by the giant impact would have been sufficient to melt the entire mantle of the Earth, though this depends on exactly how this energy is dissipated.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

References and Suggestions for Further Reading:

- Anders, E. and E. Zinner. 1993. Interstellar grains in primitive meteorites: diamond, silicon carbide, and graphite. *Meteoritics*. 28: 490-514.
- Anders, E. and N. Grevesse. 1989. Abundances of the elements: Meteoritic and solar, *Geochim. Cosmo-chim. Acta*, 53: 197-214.
- Anders, E., 1988. Circumstellar material in meteorites: noble gases, carbon and nitrogen, in *Meteorites and the Early Solar System*, ed. by J. F. Kerridge and M. S. Matthews. Tuscon: Univ. of Arizona Press
- Benz, W., W. L. Slattery and A. G. W. Cameron. 1988. Collisional stripping of Mercury's mantle. *Icarus*. 74: 516-528.
- Boss, A. P. 1990. 3D solar nebular models: implications for Earth origin, in *Origin of the Earth*, ed. by H. E. Newsom and J. H. Jones. Oxford: Oxford Univ. Press.
- Burbidge, E. M., G. R. Burbidge, W. A. Fowler, and F. Hoyle. 1957. Synthesis of the elements in stars, *Rev. Mod. Phys.*, 29: 547-650.
- Cameron, A. G. W. 1962. The formation of the sun and planets. *Icarus*, 1: 13-69.
- Cameron, A. G. W. 1995. The first ten million years in the solar nebula. *Meteoritics*. 30: 133-161.
- Chen, J. H. and G. J. Wasserburg. 1983. The least radiogenic Pb in iron meteorites. *Fourteenth Lunar and Planetary Science Conference, Abstracts*, Part I, Lunar & Planet Sci. Inst., Houston, pp. 103-104.
- Clayton, R. N., N. Onuma, and T. K. Mayeda. 1976. A classification of meteorites based on oxygen iso-topes, *Earth. Planet. Sci. Lett.*, 30: 10-18.
- Cowley, C. R. 1995. Cosmochemistry. Cambridge: Cambridge University Press.
- Crabb, J., and L. Schultz. 1981. Cosmic-ray exposure ages of the ordinary chondrites and their significance for parent body stratigraphy, *Geochim. Cosmochim. Acta*, 45: 2151-2160.
- Grossman, L., 1972. Condensation in the primitive solar nebula, *Geochim. Cosmochim. Acta*, 36: 597-619.
- Hartmann, W. K. and D. R. Davis. 1975. Satellite-sized planetismals and lunar origin. *Icarus*. 24: 504-515.
- Jacobsen, S. and G. J. Wasserburg. 1980. Sm-Nd isotopic evolution of chondrites, *Earth Planet. Sci. Lett.*, 50, 139-155.
- Kerridge, J. F., and M. S. Matthews (editors). 1988. *Meteorites and the Early Solar System*, Univ. of Arizona Press, Tucson, 1269pp.
- Larimer, J. W., 1967. Chemical fractionations in meteorites I: condensation of the elements, Geochim. Cosmochim. Acta, 31, 1215-1238.
- Larimer, J. W., The cosmochemical classification of the elements, in *Meteorites and the Early Solar System*, edited by J. F. Kerridge and M. S. Matthews, p. 375-389, Univ. Arizona Press, Tuscon, 1988.
- Lee, T., D. A. Papanastassiou and G. J. Wasserburg, 1976. Demonstration of ²⁶Mg excess in Allende and evidence for ²⁶Al, *Geophys. Res. Lett.*, 3: 41-44.
- McDonough, W. F. and S.-S. Sun. 1995. The composition of the Earth. Chem. Geol. 120: 223-253.
- McSween, H. Y., Jr. 1987. *Meteorites and their Planet Bodies*, New York: Cambridge University Press.
- Newsom, H. E. and J. H. Jones (editors). 1990. Origin of the Earth, Oxford: Oxford Univ. Press.
- Newsom, H. E. and S. R. Taylor. 1989. Geochemical implications of the formation of the Moon by a single giant impact. *Nature*. 338: 29-34.
- Papanastassiou, D. A., and G. J. Wasserburg. 1969. Initial strontium isotopic abundances and the resolution of small time differences in the formation of planetary objects. *Earth Planet. Sci. Lett.*, 5: 361-376.
- Podosek, F. A. 1970. Dating of meteorites by high temperature release of iodine correlated ¹²⁹Xe, *Geochim. Cosmochim. Acta*, 34: 341-365.
- Prinn, R. G. and B. Fegley. 1989. Solar nebula chemistry: origin of planetary, satellite and cometary volatiles. in *Origin and Evolution of Planetary and Satellite Atmospheres*, ed. S. K. Atreya, J. B. Pollack and M. S. Mathews. Tuscon: Univ. of Arizona Press.

Geochemistry

CHAPTER 10: COSMOCHEMISTRY

- Rubin, A. E. 1995. Petrological evidence for collisional heating of chondritic asteroids. *Icarus*. 113: 156-167.
- Talyor, S. R. 1991. Accretion in the inner nebula: the relationship between terrestrial planetary compositions and meteorites. *Meteoritics*. 26: 267-277.
- Tang, M. and E. Anders. 1988. Isotopic anomalies of Ne, Xe, and C in meteorites. II. Interstellar diamond and SiC: Carriers of exotic noble gases, *Geochim. Cosmochim. Acta*, 52, 1235-1244, 1988.
- Tatsumoto, M., R. J. Knight, and C. J. Allègre. 1973. Time differences in the formation of meteorites ad determined from the ratio of lead-207 to lead-206, *Science*, 180: 1279-1283.
- Taylor, S. R. 1975. Lunar Science: A Post-Apollo View. Pergamon Press, New York, 372 pp.
- Taylor, S. R. 1992. Solar System Evolution: A New Perspective. Cambridge: Cambridge University Press.
- Thiemens, M. H., and J. E. Heidenreich. 1983. The mass independent fractionation of oxygen A novel isotopic effect and its cosmochemical implications, *Science*, 219: 1073-1075.
- Van Schmus, W. R., and J. A Wood. 1967. A chemical petrologic classification for the chondritic meteorites. *Geochim. Cosmochim. Acta*, 31: 747-765.
- Wasserburg, G. J., M. Busso, R. Gallino and C. M. Raiteri. 1994. Asymptotic giant branch stars as a source of short-lived radioactive nuclei in the solar nebula. *Astrophys. J.* 424: 412-428.
- Wasson, J. T., 1974. Meteorites, Springer-Verlag, Berlin, 316 pp.
- Wasson, J. T., 1985. *Meteorites: Their Record of Early Solar System History*, New York: W. H. Freeman.
- Wetherill, G. W. 1990. Formation of the Earth, Ann. Rev. Earth Planet. Sci., 18: 205-256.
- Wetherill, G. W. and G. R. Stewart. 1993. Formation of planetary embryos: effects of fragmentation, low relative velocity, and independent variation of eccentricity and inclination. *Icarus*. 106: 190-209.
- Wood, J. A. 1988. Chondritic meteorites and the solar nebula, Ann. Rev. Earth Planet. Sci., 16, 53-72.

Problems

1. On the part of the chart of the nuclides below, identify the mode of origin (S, R, or P process) of the stable isotopes of W, Re, Os, and Ir by writing S, R or P in the box for each (remember some nuclides can be created by more than one process). Identify those isotopes you feel should be most abundant and those least abundant. On the chart below, mass numbers are given for only the stable isotopes. As a start, assume the S-process path starts at ¹⁸¹Ta. Assume the unstable isotopes will decay before capturing a neutron during the S-process.



Geochemistry

CHAPTER 10: COSMOCHEMISTRY

2. One calcium-aluminum inclusion in the Allende meteorite has δ^{26} Mg values which imply a 26 Al/ 27 Al ratio of 0.46×10^{-4} at the time of it's formation. A second inclusion apparently formed with a 26 Al/ 27 Al ratio of 1.1×10^{-4} . The half-life of 26 Al is 7.2 $\times 10^{5}$ years. Assuming both these inclusions formed from the same cloud of dust and gas and that the 26 Al/ 27 Al ratio in this cloud was uniform, what is the time interval between formation of the two inclusions?

3. Assuming that the oxygen in C3 chondrites are a mixture of oxygen having an oxygen isotope composition lying on the terrestrial fractionation line in Figure 10.23 and pure ¹⁶O, how much ¹⁶O would have to be added to oxygen lying on the terrestrial fractionation line to reproduce their oxygen isotopic composition?

4. Using the partition coefficients in Table 6.5, estimate the fraction of plagioclase that would have to fractionally crystallize from a lunar magma ocean to produce the Eu anomaliy of KREEP shown in Figure 10.32. Hint: concern yourself only with the Eu/Sm ratio.

5. Make a plot of the log of the fraction of Os condensed from a gas of "solar" composition as a function of temperature (e.g., a plot similar to Figure 10.27). Assume a total pressure of 10^4 atm, ΔH_V^o of 738 kJ/mol, ΔS_V^o of 139 J/mol, and the solar system abundances in Table 10.2. Assume the solid is pure Os metal. (*HINT: about 50% will be condensed at 1737 K.*)

CHAPTER 11: THE MANILE AND CORE

CHAPTER 11: GEOCHEMISTRY OF THE SOLID EARTH I: THE MANTLE AND CORE

INTRODUCTION

hough sizable pieces of mantle occasionally crop out on the Earth's surface as alpine massifs or ophiolites, and small pieces, xenoliths, are occasionally brought to the surface by volcanic eruptions, much of what we know about the mantle has been deduced indirectly. Indirect methods of study include determination of geophysical properties such as heat flow, density and seismic velocity. Another indirect method of study is examination of volcanic rocks produced by partial melting of the mantle. Finally, the composition of chondrites is an important constraint on the composition of the mantle.

The mantle was once viewed as being homogeneous, but we now realize that the chemistry of the mantle is heterogeneous on all scales. On a large scale, the mantle appears to consist of a number of reservoirs that have complex histories. The best evidence for this large-scale heterogeneity comes from trace element and isotope ratio studies of volcanic rocks, but there is also evidence that the major element composition of the mantle varies. While trace elements may vary by an order of magnitude or more, the major element variations are much more subtle, just as they are in volcanic rocks and in the crust. Isotope studies have proven tremendously valuable in understanding the mantle for several reasons. First, unlike trace element and major element concentrations, isotope ratios do not change during the magma generation process (except by mixing of the magma with other components such as assimilated crust). Second, radiogenic isotope ratios provide *time-integrated* information about the parent/daughter ratios, and therefore allow inferences about the history of the mantle.

Composition of the Earth's Mantle

Table 11.1 lists the relative masses of the various shells of the Earth. From a geochemical perspective, these are the Earth's major chemical reservoirs. The mantle is by far the largest single reservoir, and constitutes over 99% of the mass of the silicate Earth, the mass of the crust being less than 1%. The crust has apparently been created from the mantle through partial melting over geologic time. For the major elements, this process of crust formation does not change the composition of the mantle on the whole because the volume of the crust is so small (but extraction of partial melts undoubtedly has affected the major element composition of the mantle *locally*). With the exception of highly incompatible elements, which can be highly concentrated in the crust, and siderophile and possibly chalcophile elements, which are concentrated in the core, major lithophile elements are present in the mantle in the same proportions as in the Earth as a whole. Since we assume that chondritic meteorites are a good representation of Solar System composition, we naturally want to compare the composition of the mantle to that of chondrites. Thus the composition of meteorites is an important constraint on the composition of the mantle and the Earth as a whole.

Cosmochemical Constraints on the Earth's Composition

In Chapter 10 we considered the composition of that part of the Solar System accessible to sampling: meteorites, the solar surface (deduced from optical spectra and solar wind particles), and the Moon. The meteorite data provide first order constraints on the formation and composition of the Earth. Two important cosmochemical assumptions guide our thinking about the composition of the Earth: (1) the entire solar system formed more or less simultaneously from one gaseous, dusty nebula, (2) the composition from which the solar system formed was similar to that of CI chondrites except for the most volatile elements (e.g., H, N, noble gases). The last statement should not be construed to mean that all bodies that formed from this nebula are of chondritic composition.

CHAPTER 11: THE MANTLE AND CORE

TADLE TT.T. VOLUMES AND MASSES OF THE EARTH'S SHELLS						
	Thickness	Volume	Mean Density	Mass	Mass	
	(km)	$10^{27} \mathrm{cm}^3$	g/cc	10 ²⁴ kg	Percent	
Atmosphere				0.000005	0.00009	
Hydrosphere	3.80	0.00137	1.03	0.00141	0.024	
Crust	17	0.008	2.8	0.024	0.4	
Mantle	2883	0.899	4.5	4.016	67.2	
Core	3471	0.175	11.0	1.936	32.4	
Whole Earth	6371	1.083	5.52	5.976	100.00	

.

Thus far there is essentially no evidence to contradict the first assumption (provided we interpret "simultaneous" in a geologic sense; i.e., this process may have taken up to 100 million years); the second assumption certainly holds to a first approximation, beyond that there is some uncertainty. In defense of it, however, we should emphasize the isotopic homogeneity of the solar system. The isotopic anomalies we have discussed previously are very much the expectation to the rule and serve to emphasize the general homogeneity of the solar system (except, of course, for oxygen). The implication is that the presolar nebula was well mixed. We can reasonably conclude that it was initially chemically homogeneous as a result.

In summary, we can draw the following conclusions relevant to the formation and composition of the Earth from cosmochemistry:

1. The material from which the solar system formed was of CI chondritic composition (plus gases).

2. Despite (1), the composition of bodies formed from this nebula is variable. This is apparent from the composition of chondrites as well as from density variations of the planets.

3. Formation of planetary bodies begin shortly (a few million years at most) after newly synthesized elements were added to the solar nebula (as evidenced by the oldest ages from the Moon and differentiated meteorites).

4. Iron meteorites clearly reflect segregations of metallic liquid from material of approximately chondritic composition. Certainly the existence of stony-irons, which are irons with silicate inclusions, demonstrates that segregated silicate and iron metal fractions existed in the parent bodies. Since the ages of these meteorites are

nearly indistinguishable from those of chondrites, it suggests core formation in planets got under way very early. In the most likely case, cores had already segregated in the planetismals and planetary embryos from which the planets formed.

5. The Moon seems to have had a magma ocean. The Earth may well have had one as well.

Upper Mantle Composition

Most geochemists and geophysicists believe the mantle is ultramafic in composition, peridotitic to be more specific. Peridotite is a rock dominated by olivine (peridot is the gem name for olivine). The nomenclature of ultramafic rocks is summarized in Figure 11.1. If the olivine exceeds 90% of the rock, it is termed a dunite. If the rock consists al- Figure 11.1. Ultramafic rock nomenclature.



'. M White

CHAPTER 11: THE MANTLE AND CORE

most exclusively of olivine and orthopyroxene, it is called a harzburgite. A websterite is a rock consisting almost exclusively of olivine and clinopyroxene. A rock consisting of over 90% pyroxene is a pyroxenite (clinopyroxenite or orthopyroxenite as the case may be; a pyroxenite is an ultramafic rock, but not a peridotite). Rocks with substantial amounts of both pyroxenes as well as olivine are *lherzolites.* This can be prefaced by the name of the Al-bearing phase, e.g., spinel lherzolite, whose nature depends on pressure. In the upper mantle, olivine is thought to be the dominant mineral, followed by orthopyclinopyroxene, and an Al-bearing roxene, phase, which can be plagioclase, spinel, or garnet, depending on the pressure. The upper mantle is thought to be dominantly *lherzolitic.* An alternative term is *pyrolite* (pyroxene-olivine rock), coined by A. E. Ringwood. Pyrolite, however, is associated with a specific composition (e.g., Table 11.2), though this composition has been revised numerous times.

Any estimate of upper mantle composition must satisfy a number of constraints. First, it should match the composition of direct mantle samples such as xenoliths (Figure 11.2). Second, it must yield basalt upon melting. Third, it must have seismic velocities (Figure 11.3) and densities (Figure 11.4) matching those observed. Finally, it should have an approximately chondritic composition (as we shall see, the upper mantle does not appear to be exactly chondritic). A lherzolitic or pyrolitic composition appears to fit these constraints. However, lherzolitic compositions are not unique in meeting these some of these constraints, and some have argued that the mantle, or at least the upper mantle, is dominantly eclogitic. Eclogite is a rock consisting almost exclusively of garnet and omphacitic pyroxene (i.e., Na-, Al-rich). Figure 11.3. Mantle seismic velocities. This is the mineralogy of basalt at high

Clinopyroxene Number of Samples 2 - 5 6 - 20 21 - 40 >40

Geochemistry





pressure. It was once argued by some that the upper mantle composition was closer to eclogite than peridotite, because it seemed to fit seismic velocity distribution better. But the idea was not accepted, primarily because it failed to match the chondritic composition sufficiently well, and because

Geochemistry

CHAPTER 11: THE MANTLE AND CORE



Figure 11.4. Phase assemblages and 1 atm density of the mantle.

eclogites are considerably rarer as xenoliths than peridotites. Also seismic anisotropy^{*} observed in some parts of the upper mantle suggests the mantle is too rich in olivine to be eclogite. However, in the last decade Don Anderson (e.g., Anderson, 1989b) has argued for an upper mantle consisting of olivine bearing eclogite, for which he coined the term *piclogite*[†]. The eclogite, which he argues was produced by crystal accumulation in a primordial magma ocean, is the source of MORB in his model. This idea has not been particularly well received by geochemists, perhaps even less so by petrologists.

Most studies rely on the composition of mantle xenoliths to estimate mantle compositions (the composition of peridotite massifs, which also represent mantle samples, but these have inevitably been affected by secondary processes). However, while formation of the crust and other processes have not affected the major element composition of the bulk mantle, the compositions of individual xenoliths often have been affected by igneous processes. Despite this, xenoliths provide evidence that the upper mantle differs in composition from chondrites. Whether the estimated upper mantle composition is also the composition of the entire mantle depends on whether the mantle is chemically layered. This is an important and unresolved question, and one to which we shall return.

Table 11.2 compares several estimates for the major and minor element composition of the mantle made over the last decade with the composition of CI chondrites after subtracting the most volatile elements in the latter (column 1). Clearly chondrites are much richer in siderophiles (e.g., Fe, Ni) than all these estimated mantle compositions. The chondritic composition matches the mantle composition much better after a sufficient amount of the siderophile elements has been removed to form

^{*} Seismic anisotropy refers to the case where seismic waves travel faster in one direction than an other. This can ultimately be related to crystalographic properties. In olivine, seismic velocities are faster along one crystallographic axis than the others. Since garnet has a cubic (isotropic) structure, seismic velocity is equal in all crystallographic directions, as is the velocity of light. Seismic anisotropy can only occur only when there is some tendency for crystallographic axes of individual mineral grains to be aligned parallel. This is thought to arise in the mantle from flow associated with convection.

⁺ Piclogite comes from combining *picrite*, which is an olivine-rich basalt, and eclogite. Piclogite would be the high pressure equivalent of picrite.

Geochemistry

CHAPTER 11: THE MANILE AND CORE

the Earth's core. However, even after removing the siderophile and highly volatile elements, there are significant differences between the apparent composition of the mantle and chondrites.

First, the mantle is depleted in the alkali elements (e.g., K and Na in Table 11.2). There are several lines of evidence for this depletion of alkali elements. The depletion in moderately volatile elements is also apparent by comparing Sr isotope ratios of the mantle and chondrites, as Gast demonstrated in 1960. ⁸⁷Sr/⁸⁶Sr ratios in the mantle imply Rb/Sr ratios more than an order of magnitude lower than chondritic. Some of the Rb depletion of the mantle may be explained by extraction of the Rb into the crust. Indeed more than half the Earth's Rb may be in the crust. However, the terrestrial Rb/Sr ratio appears to be nearly an order of magnitude lower than chondritic (0.03 vs. 0.25) even when crustal Rb is considered. Independent of Sr isotope considerations, a number of other studies have demonstrated depletion of K, Rb, and Cs in the Earth.

	CI	CI Chondritic	Hart & Zindler	Ringwood ³	McDonough & Sun ⁴	Allegre et al. ⁵
	Chondrites	Mantle & Core ¹	LOSIMAG ²	Pyrolite	Pyrolite	PRIMA
SiO ₂	22.77	49.52	45.96	44.76	45.0	46.12
Al ₂ O	1.64	3.56	4.06	4.46	4.45	4.09
FeO	24.49	7.14	7.54	8.43	8.05	7.49
MgO	16.41	35.68	37.78	37.23	37.8	37.77
CaO	1.30	2.82	3.21	3.60	3.55	3.23
Na_2C	0.67	0.292	0.332	0.61	0.36	0.36
K ₂ O	0.067	0.028	0.032	0.029	0.029	0.034
Cr_2O_3	0.39	0.412	0.468	0.43	0.384	0.38
MnO	0.256	0.115	0.130	0.14	0.135	0.149
TiO ₂	0.073	0.159	0.181	0.21	0.20	0.18
NiO	1.39	0.244	0.277	0.241	0.25	0.25
CoO	0.064	0.012	0.013	0.013	0.013	0.07
P_2O_5	0.274	0.018	0.019	0.015	0.021	
Sum	69.79	100.0	100.0			
HVE	30.21	30.21	30.21			
MVE	7	1.258	1.258			
SVE^8			5.33			
Fe	74.44	85.62				79.39
Ni	4.53	5.16				4.87
Со	0.2081	0.237				0.253
S						2.3
0		20.81	8.99			4.10
Core	Sum	100.0	100.0			
Total	99.9	99.7	99.4			

Table 11.2. Comparison of Estimated Silicate Earth Compositions

¹removing HVE, MVE and siderophiles and some oxygen from mantle to form core; renormalizing silicate mantle and core separately. Hart and Zindler (1986)

²Low-Si-Mg Earth with siderophiles and some oxygen removed from mantle to form core; renormalizing silicate mantle and core separately. Hart and Zindler (1986)

³Estimated Primitive Upper Mantle Composition of Ringwood (1991)

⁴Esitmated Primitive Mantle Composition from McDonough and Sun (1995)

⁵Primitive Mantle ("PRIMA") of Ållegre et al. (1995)

 $^{6}_{-}$ HVE = Highly Volatile Elements, includes H₂O, S, C, organics, etc.

⁷ MVE = Moderately Volatile Elements, includes Na, K, P, Cr, Mn

⁸ SVE = Slightly Volatile Elements, Si and Mg

Geochemistry

CHAPTER 11: THE MANILE AND CORE

This depletion is thought to encompass all the moderately volatile elements (see Figure 9.29) as all appear to be depleted in the mantle. Many of the moderately volatile elements are siderophile or chalcophile, so their depletion in the mantle may also reflect extraction into the core.

The Earth's depletion in moderately volatile elements is not entirely surprising, given that it, along with the other terrestrial planets, is obviously depleted in the atmophile elements. Since the depletion in the highly volatile elements is a feature shared by all the terrestrial planets, it is probably due to high temperatures in the inner solar nebula during the initial stages of accretion. Thus these elements probably did not condense and were not accreted to the planetismals that ultimately formed the terrestrial planets. An enhanced solar wind during the T-Tauri phase of the Sun might also have caused loss of volatiles that were accreted. Finally, high temperatures achieved

during formation of the Earth (due to release of gravitational energy), particularly as a result of the giant impact, probably contributed to volatile loss.

The depletion of the Earth in volatile elements is well established, and more recent work has focused on comparing the relative concentrations of remaining major elements, particularly Mg, Si, Ca, and Al in the mantle with those of CI chondrites.

Table 11.2 reflects a general agreement that the upper mantle is depleted in silicon relative to a 'chondritic' upper mantle. This depletion can be demonstrated in several ways, but we will consider only the approach of Hart and Zindler (1986). Figure 11.5a shows the Mg/Si ratio plotted against the Nd/Si ratio for mantle xenoliths and various classes of chondritic meteorites. The xenoliths and chondrites form 2 distinct arrays. Hart and Zindler argued that the "meteorite array" reflects fractionation during processes occurring in the solar nebula or during planet formation (e.g., volatilization), whereas the "terrestrial array" reflects processes occurring in the Earth's mantle such as partial melting. For example, the dashed line shows that extraction of a partial melt would drive composition off the meteorite array in the general trajectory of the terrestrial array. Since the bulk mantle composition should not have been affected by crust formation, Hart and Zindler, following earlier work by Jagoutz et al. (1979), reasoned that it should lie at the intersection of the meteorite and terrestrial arrays. A similar sort of analysis can be done with Al/Si (Figure 11.5b) and Ca/Si



Figure 11.5. a.) Variation of Mg/Si as a function of Nd/Si in terrestrial mantle xenoliths and meteorites. b.) Variation of Mg/Si as a function of Al/Si in terrestrial mantle xenoliths and meteorites. The data suggest the Earth is depleted in Mg and Si relative to chondrites. Dashed line is the trajectory of residual mantle compositions following melt extraction with the % melt extracted shown by the tic marks. After Hart and Zindler (1986).

CHAPTER 11: THE MANILE AND CORE

ratios. Using this approach, Hart and Zindler concluded that the upper mantle, is depleted in Si. If the mantle is not chemically layered, then the entire Earth must be depleted in Si and Mg relative to chondrites. Ringwood (1989) argued that the depletion in Si is a general feature of the inner solar system, and reflects the more volatile nature of SiO compared to the gaseous phases of Al, Mg, Ca, and Fe.

Though it is less obvious, all the estimated mantle compositions in Table 11.2 also have high abundances of the most refractory elements, Ca, Al, and Ti, relative to Mg. For example, the chondritic Al/Mg ratio is 0.123 whereas estimates of upper mantle Al/Mg ratio ranges from 0.133 (Allegre et al., 1985) to 0.147 (Ringwood, 1991). Estimates of the upper mantle Ca/Mg ratio range from 0.101 (Hart and Zindler, 1986; Allegre et al., 1995) to 0.115 (Ringwood, 1991), and are all higher than the chondritic value of 0.094. Assuming these estimates are correct, the upper mantle appears to be enriched in the highly refractory elements by 7 to 50%. If the mantle is not chemically zoned, then the Earth as a whole is enriched in the highly refractory elements.

Palme and Nickel (1985) argued that the Ca/Al ratio of the mantle is also non-chondritic (0.887 for the upper mantle vs. 0.766 for chondrites). Since Ca and Al are both among the most refractory elements whose relative abundances is uniform in all classes of chondrites (despite variations in abundances of more volatile elements), it would be surprising indeed if the Ca/Al ratio of the Earth as a whole were not chondritic. Palme and Nickel argued that the difference between their estimated upper mantle Ca/Al ratio and the chondritic one reflects early mantle differentiation, perhaps during crystallization of a magma ocean, that resulted in a high Ca/Al ratio in the upper mantle and a low Ca/Al ratio in the lower mantle. Hart and Zindler (1986) disputed the high Ca/Al estimated by Palme and Nickel. The dispute points out the difficulties in using xenoliths, all of which have been affected by various processes, to estimate upper mantle compositions. Most recent estimates of the silicate Earth composition have Ca/Al values close to the chondritic one (0.76 to 0.78).

Phase Transitions in the Mantle

The mineral assemblage of the mantle changes as pressure in-Understanding creases. the phase relationships in the mantle is important for two reasons. First, seismic velocities and densities are the most important constraint on the bulk composition of the mantle, and these depend on the mineral assemblage. Second, the pressure at which phase transitions occur is composition dependent, so that the depths at which phase transitions occur, as reflected by changes in seismic velocity, constrain the composition of the mantle.

In the upper 200 km or so, the only important phase changes are the nature of the aluminous phase. This changes from plagio-clase to spinel (MgAl₂O₄) at about



Figure 11.6. Upper mantle phase diagram.

Geochemistry

CHAPTER 11: THE MANTLE AND CORE



Figure 11.7. Mineral assemblages in the upper 1000 km of the mantle. From Ringwood (1991).

30 km depth, and from spinel to garnet at 60 to 90 km depth, depending on composition and temperature.

The garnet peridotite assemblage remains stable to depths of about 300 km. At this depth, appreciable amounts of pyroxenes begin to dissolve in garnet, forming a solid solution with the general composition $M_2(MSi,Al_2)Si_3O_{12}$ where M is Mg, Fe, or Ca. This garnet, called *majorite*, differs from those found at lower pressure in that up to a quarter of the silicon atoms are in octahedral coordination (i.e., surrounded by 6 oxygens rather than 4). The octahedral coordination is favored because the anions such as oxygen are more compressible than are cations such as silicon. When compressed, more oxygens can be packed around each silicon atom. This phase change is a gradual one, with complete conversion of pyroxenes to majorite at about 460 km depth (Figure 11.7). The phase change results in a roughly 10% increase in density of the "pyroxene" component.

THE TRANSITION ZONE

Between 400 and 670 km depth seismic velocities increase more rapidly than elsewhere (Figure 11.3) and

this depth interval is often called the transition region. At about 400 km, or 14 GPa, olivine undergoes a structural change from the low pressure, or α form, to the β form. In contrast to the pyroxene-to majorite phase change, this phase boundary is relatively sharp, with a transition interval of 9-17 km. The α to β transition results in an 8% increase in density.

At about 500 km depth or so, olivine undergoes a further structural change to the γ -form. The structure is similar to that of MgAl₂O₄ spinel, and this phase is sometimes, somewhat confusingly, referred to simply as spinel, though the name *ringwoodite*, has been applied to γ -olivine found in shocked meteorites. The change from β to γ -olivine is thought to be more gradual than the α - β transition, occurring over a depth interval of 30 km, and involves only a 2% increase in density. In both the β and γ phases silicon remains in tetrahedral coordination.

Within the transition zone, some of the Mg and Ca in majorite begin to exsolve to form $CaSiO_3$ in the perovskite structure and MgSiO₃ in the ilmenite structure. The proportion of $CaSiO_3$ perovskite increases with depth until majorite disappears at about 720 km. MgSiO₃ ilmenite persists only to 660 km.

A very sharp and large increase in seismic velocity occurs at around 660^{\ddagger} km depth, and is called the 660 seismic discontinuity[§]. This depth marks the beginning of the lower mantle. Over the years there has been considerable debate as to whether this discontinuity is a chemical boundary or a phase change. There is now complete agreement that it is primarily due to a phase change, though debate continues whether it might also be a chemical boundary (a question discussed below). At this depth γ -olivine disproportionates to form (Mg,Fe)SiO₃ in the perovskite structure and (Mg,Fe)O magnesiowüstite, with the Fe going preferentially in the magnesiowüstite. This phase change results in a density increase of about 11%. The (Mg,Fe)SiO₃ perovskite has the chemical stoichiometry of pyroxene, but the silicons are in octahedral coordination. The structure, illustrated in Figure 11.8, is similar to that of the "high-temperature" cuprate superconductors discovered in the 1980's.

[‡] The exact depth of the transition has been debated and maybe anywhere between 650 and 690 km. In fact, the depth at which the transition occurs could vary due to lateral variations in temperature within the mantle, or because of mantle convection.

[§] This is also variously known as the 650 discontinuity and the 670 discontinuity.

CHAPTER 11: THE MANILE AND CORE

High pressure experiments carried out with the diamond anvil show that the transition is quite sharp, occurring within a pressure interval of 0.15 GPa at 1600° C. The transition has a negative Clapeyron slope (P (GPa) = 27.6- 0.0025 T (°C); e.g., Chopelas, et al., 1994), so that it will occur at somewhat shallower depth in hot regions, such areas of mantle upwelling, and at greater depth in cooler regions, such as subducted lithosphere, though since the Clapeyron slope is shallow, the effect is small. The affect of these differences is to oppose motion across the boundary, and therefore to stablize any chemical differences that might exist between the upper and lower man-To understand this, consider a sinking subducted tle. lithospheric slap. Such slabs sink through the upper mantle because they are denser than surrounding mantle. Let's arbitrarily suppose that it is 200° cooler than surrounding mantle. As a result it will undergo the γ -olivine-perovskite phase change at about 0.5 GPa greater pressure (15 km or so deeper) than surrounding mantle. As the slab penetrates to below 660 km, its density contrast between the slab and surrounding mantle will be reduced, because γ -olivine is the principal phase in the slab



Figure 11.8. The structure of $MgSiO_3$ perovskite. The structure consists of corner-sharing SiO_6 octahedra with Mg^{2+} in dodecahedral sites.

whereas Mg-perovskite is the major phase in the surrounding mantle. The reduced density contrast results in less gravitation energy to cause further sinking of the slab. For similar reasons, the phase transition will also tend to oppose the motion of rising hot mantle, mantle plumes for example. Indeed, there is seismological evidence that the discontinuity is deflected downward under subduction zones. Though the negative Clapeyron slope opposes motion across the boundary, it is probably not sufficient to prevent it entirely unless the 660 discontinuity is also a chemical boundary.

As we noted above, there is no question that phase transitions are primarily responsible for the increase in seismic velocity observed in the transition zone (Figure 11.3). Anderson (1989) argues, however, that change in seismic velocity at 400 km is smaller than what should occur in an olivine-rich (i.e., lherzolitic) mantle, and therefore that the upper mantle must be piclogitic rather than lherzolitic. Ringwood (1991) argued that, because of the negative Clapeyron slope, many slabs are not dense enough to sink through the 660 discontinuity and therefore pile up above it. In his view, the there is a layer of 50 km thickness or so of accumulated oceanic crust above the 670 discontinuity. Resolution of these questions requires better data on the elastic properties of phases within the transition zone.

Composition of the Lower Mantle

The lower mantle, the region between the 660 km seismic discontinuity and the core-mantle boundary at 2900 km, is substantially less accessible to study than is the upper mantle. It composition must be inferred indirectly. The only direct constraint on its composition is its seismic properties. The assumption that the Earth is approximately chondritic forms another constraint. It is generally agreed that the lower mantle is grossly similar in compositional to the upper mantle, i.e., composed dominantly of SiO₂, MgO, and FeO with lesser amounts of CaO, Al₂O₃, TiO₂, etc. However, there is substantial disagreement as to its precise composition, in particular whether it is richer in FeO and SiO₂ than the upper mantle.

Perovskite and magnesiowüstite appear to remain the principal phases throughout the lower mantle. Perovskite constitutes about 80% of the lower mantle, and thus is the most abundant mineral in the Earth. The magnesiowüstite/perovskite partition coefficient for Fe decreases with increasing pressure, so that the proportion of Fe in perovskite should increase somewhat with depth up to about

Geochemistry

CHAPTER 11: THE MANILE AND CORE

30 GPa (~900 km), and this is constant at greater pressure. At this depth, the mole fractions of Mg (Mg/(Mg+Fe)) in MgSiO₃-perovskite and magnesiowüstite ((Mg,Fe)O) for a pyrolite composition are about 0.93 and 0.80 respectively.

There has been some debate about how Al_2O_3 is accomodated in lower mantle minerals. Some have argued that a separate aluminous phase forms. The most recent experiments however, suggest that while a separate aluminous phase forms at relatively low pressure (20-30 GPa, correpsonding to the region of 600-800 km), at pressures above about 35 GPa, Mg-perovskite apparently accepts all available Al_2O_3 in its structure, so that there is no separate aluminous phase in the lower mantle(e.g., Kesson et al., 1994) below about 800 km. However, the solid solution between Mg-perovskite and Ca-perovskite appears to be quite limited, so that Ca perovskite is probably present as a separate phase, and probably accepts all the Na₂O.

There is considerable uncertainty about the composition of the lower mantle, and there are two principal viewpoints on the question. In the first, the lower mantle is compositionally similar to the upper mantle, i.e., an approximately pyrolitic composition, with no major chemical boundaries. In this case, the lower mantle would consist of 80% perovskite, with about 10% magnesiowüstite (Figure 11.7). Proponents of the second view claim that an essentially pure perovskite lower mantle better fits seismic observations. A second argument favoring a pure perovskite lower mantle is that it is richer in SiO₂ and therefore closer to chondritic than upper mantle compositions. At present, the density and elastic properties of lower mantle minerals are not sufficiently well known to distinguish between these alternatives, as the predicted densities of the pyrolite and chondritic models of the lower mantle differ by less than 0.06 g/cc. Uncertainties about the temperature of the lower mantle, and the coefficient of thermal expansion of lower mantle materials, compound the problem. Though a 'chondritic' lower mantle does seem to match the seismic observations slightly better, a pyrolite composition appears to match the seismic observations within about 1%, which is less than the combined uncertainty of the seismic observations and the values for seismic properties of lower mantle materials. D. Anderson and R. Jeanloz, among others, also argue that the lower mantle is richer in iron than the upper mantle. Jeanloz has suggested a molar Mg/(Mg+Fe) ratio of about 0.8 for the lower mantle, compared to 0.9 for the upper mantle.

The importance of this problem to both the present working of the Earth and its past evolution should be emphasized. If the mantle is chemically stratified or layered, that is if the composition below the 660 discontinuity differs from that above, it implies that convection within the mantle is layered; moreover, it implies there is little chemical exchange between the upper and lower mantles. In this case, it seems unlikely that subducted lithospheric slabs penetrate to depths greater than 660 km. Similarly, if the mantle is layered, mantle plumes, which are responsible for most interplate volcanism, probably arise from the 660 km discontinuity rather than the core-mantle boundary. A layered mantle also implies that the continental crust was derived only from the upper mantle, which has important implications for mass balance models of the Earth, which we consider below. On the other hand, a homogenous mantle allows, but does not require, whole mantle convection. Finally, it is possible, and perhaps likely, that if the mantle is layered, this layering originated during crystallization of an early magma ocean, a question to which we will return.

D" and the Core-Mantle Boundary

For the most part, seismic velocities increase smoothly throughout the lower mantle and appear to be consistent with self-compression of chemically homogeneous material. However, the lower two hundred kilometers or so of the mantle appear to be seismically peculiar. Within this layer, known as D'' (d-double-prime), both P- and S-wave velocity gradients are smaller than elsewhere. Furthermore, the thickness of D'' varies significantly. Though the anomalous gradients in this region have been recognized for 50 years, the extremely heterogeneous nature of D'' has only recently been discovered. The shallow seismic gradients can be fairly easily explained if this region is a thermal boundary layer between the core and the mantle, as it would be hotter. The heterogeneity cannot be so easily explained, and is the subject of debate. Again, there are two views. In the first view, the

CHAPTER 11: THE MANILE AND CORE

region is a chemical boundary layer and is produced by reaction between silicates of the mantle and the FeNi core. In the other view, the heterogeneity arises from the presence of subducted oceanic lithosphere that has sunk to the base of the mantle. In this latter view, the D´´ layer is the source of mantle plumes, which we discuss in a subsequent section.

THE "PRIMITIVE MANTLE"

An important concept in mantle geochemistry is that of the "primitive mantle" composition. Primitive mantle is defined as the composition of the mantle after separation of the core, but before separation of the crust. It is thus equivalent to the "Bulk Silicate Earth" composition. As we will see, it seems likely that the core formed very early, perhaps before accretion was complete, whereas the crust seems to have formed much more slowly, so equating primitive mantle with bulk silicate earth seems reasonable.

Table 11.3. Primitive Mantle Concentrations

Li	1.6	Ge	1.1	Nd	1.225
Be	0.066	As	0.13	Sm	0.381
В	0.5	Se	0.05	Eu	0.145
С	250	Br	0.075	Gd	0.526
F	26	Rb	0.60	Tb	0.093
Na	2545	Sr	19.9	Dy	0.651
Mg	219407	Y	4.05	Ho	0.146
Al	22985	Zr	10.5	Er	0.428
Si	214766	Nb	0.651	Tm	0.058
Р	95	Мо	0.065	Yb	0.439
S	350	Ru	0.0042	Lu	0.065
Cl	330	Rh	0.001	Ηf	0.283
Κ	240	Pd	0.005	Та	0.037
Ca	23858	Ag	0.008	W	0.021
Sc	15.5	Cd	0.04	Re	0.00028
Τi	1153	In	0.013	Os	0.0034
V	82	Sn	0.175	Ir	0.0033
Cr	2935	Sb	0.005	Pt	0.0068
Mn	1080	Те	0.013	Au	0.00075
Fe	65500	Ι	0.011	Hg	0.01
Со	105	Cs	0.021	T1	0.007
Ni	1890	Ва	6.189	Pb	0.18
Cu	30	La	0.624	Вi	0.0025
Zn	56	Ce	1.637	Th	0.0813
Ga	3.9	Pr	0.238	U	0.0203

All concentrations in ppm.



Figure 11.9. Apparent Th/U ratio of magma sources as a function of time. ²⁰⁸Pb*/²⁰⁶Pb* is the ratio of *radiogenic* ²⁰⁸Pb to *radiogenic* ²⁰⁶Pb. The plot shows data for komatiites (komatiites are ultramafic lavas produced by relatively high degrees of melting) and ophiolites (Bay of Islands and Oman), which are oceanic crustal sections. After Allègre et al. (1986).

In estimating primitive mantle composition, we start with the assumption of an approximately chondritic Earth. The primitive mantle composition will differ from chondritic, however, as a result of two processes: loss of volatiles from the inner Solar System or during formation of the Earth, and loss of siderophile elements to the Earth's core.

Though the compositions of the various classes of chondrites differ, they do have constant relative abundances of the highly refractory lithophile elements, i.e., those elements concentrated in CAI's (see Chapter 10): Al, Ca, Sc, Ti, Sr, Y, Zr, Nb, Ba, REE, Hf, Ta, Th, and U. Thus for example, the Sm/Nd ratio is nearly constant in different meteorites despite variations in the absolute abundances of Sm and Nd. This being the case, i t is reasonable to assume that these elements are present in the same relative abundances in the silicate Earth as in chondrites. This group includes the parents and daughters of the Sm-Nd and Lu-Hf radioactive decay schemes. It is this assumption that the Sm/Nd ratio of the Earth is

Geochemistry

CHAPTER 11: THE MANILE AND CORE

chondritic that makes this decay scheme so useful (among other things). Note also that some important elements in isotope geochemistry are absent from this group: Re, Os, Rb and Pb.

Though both Th and U are refractory elements, there is evidence that the Th/U ratio of the Earth might be somewhat higher than the chondritic ratio. The Anders and Grevesse tabulation (Table 10.2) yields a chondritic Th/U ratio of about 3.8. But Pb isotope systematics of meteorites suggest a higher ratio, closer to 4.0. Allègre et al. (1986) have argued the ratio is as high as 4.2. Figure 11.9 illustrates one piece of evidence they use to support this hypothesis. As we demonstrate in a subsequent section, a time-integrated Th/U ratio may be calculated from $^{208}Pb^{*/206}Pb^{*}$ ratios (the asterisk is used to designate radiogenic ^{208}Pb and radiogenic ^{206}Pb). When Allègre et al. (1986) did this for a series of mafic and ultramafic rocks of various ages, they found evidence for a decrease of the Th/U ratio of about 4.2 at 4.55 Ga. The importance of this derives from recent work on Th isotope ratios, which indicates that present upper mantle Th/U ratios are about 2.3, which is much lower than time-averaged ratio derived from $^{208}Pb^{*/206}Pb^{*}$ ratios, which is about 3.6-3.8. If the primitive mantle Th/U is as low as 3.8 and the present mantle ratio is 2.3, a time integrated ratio approaching 3.8 suggests the low present Th/U is a relatively young feature. However, if the primitive mantle Th/U is 4.2, as Allègre et al. argue, the a more gradual depletion of Th/U is possible.

The abundance of other elements in the silicate Earth can also be estimated through various indirect approaches. For example, the abundance of Rb has been estimated by comparing Rb-Sr and Sm-Nd isotope systematics, and also from the Rb/Ba ratio of mantle-derived basalts, which is constant (Hofmann and White, 1983). The Pb concentration in the Earth can be estimated by noting that most major reservoirs in the Earth have Pb isotope ratios consistent with a μ (²³⁸U/²⁰⁴Pb) of about 8. K concentration can be estimated by noting that the K/U ratio of most materials is about 10000. Table 11.3 lists estimated concentrations of the elements in the Primitive Mantle (= bulk silicate Earth). The values were calculated as follows. First a 'chondritic mantle' was calculated, following Hart and Zindler (1986), by subtracting an appropriate amount of the volatile and moderately volatile elements and enough of the major siderophiles to form the core. The Si was then adjusted downward to



Figure 11.10. Abundances of the elements in the Primitive Mantle compared to CI chondrites.

Geochemistry

CHAPTER 11: THE MANILE AND CORE

accord with observations on mantle xenoliths (Hart and Zindler, 1986). The observed concentrations of the major refractory elements, Al, Ca, and Ti, in this upper mantle (Hart and Zindler, 1986) appear to be about 15% enriched relative to this calculated 'chondritic mantle'. On this basis, it was assumed that refractory lithophile trace elements (e.g., REE) were also 15% overabundant. The remaining element concentrations were taken from McDonough (1987) or Hofmann (1988) with some minor adjustments. Figure 11.10 compares these estimated primitive mantle concentrations with CI chondrites. These primitive mantle concentrations should not, of course, but taken as gospel and will undoubtedly be revised in the future.

MAGMA OCEAN AND MANTLE DIFFERENTIATION

Several observations suggest the Earth underwent extensive melting when it formed. First, the energetics of the formation of the Earth and its core suggest extensive melting occurred. The gravitational energy released by accretion of the Earth from nebular dust is sufficient to melt the entire planet. Whether this occurs or not, however, depends on how efficiently heat is lost during accretion, and how quickly accretion occurs. If the Moon formed as a result of a giant impact, the impact would release enough energy to cause very extensive melting. Formation of the core would also release a large amount of energy. In addition, there is evidence for a magma ocean on the Moon. Given that the gravitational energy released by accreting the Earth is even greater, it is difficult to see how formation of a terrestrial magma ocean could not have formed. Second, the solidus and liquidus of peridotite converge at high pressure (Figure 11.11). Some petrologists have argued that this is unlikely to be coincidental and they conclude that the lherzolitic upper mantle is a minimum melt or eutectic composition of a "chondritic mantle" composition (i.e., Table 11.2 column 2).

According to some, crystallization of a magma ocean would produce a chemically stratified mantle. The magma ocean would have crystallized both from the top down, because heat can only be lost from the surface of the Earth. It would have simultaneously crystallized from the bottom up (Figure 11.12), because the P-T slope of the solidus is greater than that of an adiabatic gradient (vigorous convention within the magma ocean would have kept temperatures nearly adiabatic). Presumably, a protocrust would have quickly formed as the surface of the magma ocean crystallized. This protocrust, like that of the Moon, might have been rich in plagioclase because plagioclase is less dense than a melt of lherzolitic composition (provided it is relatively dry) and hence floats to the surface. Plagioclase, however, would have crystallized only within the upper 30 km of so, since it is not stable at greater pressure. If melting was extensive enough to extend into the Mg-perovskite stability field (shallower than at present because of higher temperature), Mg-perovskite crystallizing in the lower part of the magma ocean would have sunk out of the magma due to its higher density. Mg-



Figure 11.11. Solidus and Liquidus of peridotite at high pressure. Closed symbols are experiments in which there was no evidence of melting, red half-filled circles are experiments that partially melted, and open circles are experiments that totally melted. From Scarfe and Takahashi (1986).

perovskite has a slightly higher Si/Mg ratio than chondrites or lherzolite, so the residual liquid would have been depleted in Si by Mgperovskite crystallization. This liquid would also have been strongly enriched in Na and Ca and modestly enriched in Al, which are excluded from perovskite, and depleted in Ti, which is enriched in Mg-perovskite. Except in the uppermost part, olivine crystallizing above the depth of the present transition zone would be less dense than the melt and would have floated upward, enriching the upper mantle in olivine. Majorite garnet, which crystallizes later than perovskite and olivine, would have also been denser than the liquid and would have settled, depleting the upper mantle in Al and other elements,

CHAPTER 11: THE MANTLE AND CORE



Figure 11.12. Solidus, liquidus and thermal gradient in a terrestrial magma ocean. Crystallization occurs in the gray regions where temperature is between the solidus and liquidus. This occurs at the very top where temperatures are not adiabatic and at the bottom, where the solidus temperature exceeds the adiabatic temperature.

Thus the magma ocean would crystallize without producing a stratified mantle.

Even if crystallization of a magma ocean had produced a stratified mantle, this stratification may not persist, since subsequent convection could have destroyed it. Little, if any of the presently observed mantle heterogeneity can be related to a primordial magma ocean. However, there is some evidence that a stratified mantle did form and persisted for a billion years or so.

The Core and Core Formation

We know little about the composition of the core or how it formed. Various lines of evidence, primarily geophysical, indicate the core consists of Fe-Ni

Geochemistry

such as the heavy rare earths, that partition into garnet. Thus in the view of some, crystallization of a magma ocean could produce a chemically stratified mantle in which the upper mantle has a lower than chondritic Si/Mg ratio and a higher than chondritic Ca/Al ratio.

There are, however, several questions about this scenario. Kato et al. (1988) determined Mg-perovskite-liquid partition coefficients for a number of elements. Hf and Sc appear to strongly partition into perovskite, whereas rare earths, such as Sm, are not. Fractional crystallization of perovskite would change the Hf/Sm and Sm/Sc ratios more rapidly than the Si/Mg ratio (Figure 11.13). However, both the Hf/Sm and Sm/Sc ratios appear to be approximately chondritic in the upper mantle. Fractional crystallization of perovskite would also increase the Lu/Hf ratio, whereas Hf isotope data suggest the early mantle had an approximately chondritic Lu/Hf ratio. In addition, Tonks and Melosh (1990) argued that convection in the magma ocean would be so vigorous that nucleating crystals would be swept out of the crystallization zones before they had time to grow to sufficient size to settle out.



Figure 11.13. Effect of fractional crystallization of Mg-perovskite on Sm/Hf, Sc/Sm, and Si/Mg ratios in a magma ocean that is initially chondritic. 10% Mg-perovskite fractionation drastically changes the Sm/Hf and Sc/Sm ratios, but only slightly changes the Si/Mg ratio. From Ringwood (1991).

Geochemistry

CHAPTER 11: THE MANTLE AND CORE

alloy. Iron meteorites contain 5-10% Ni, so the concentration of the core probably falls in this range. About 5.7% Ni in the core is required to explain the observed depletion of the mantle in Ni. Density considerations, however, require about 10% of some lighter element as well. What the light element is has been the subject of endless debate, and a resolution of the question in the near future does not seem likely. The principle candidates are S and O.

Sulfur is an obvious candidate for several reasons. First, it is severely depleted in the silicate Earth (Figure 11.10). Sulfur is volatile, and much of its depletion may be due to this. However, it is more depleted than similarly volatile elements, e.g., Zn, which strongly suggests its presence in the core. Second, iron meteorites often contain considerable amounts of FeS (troilite), indicating S was extracted into the cores of meteorite parent bodies. Finally, FeS and Fe liquids are completely miscible at low temperature, so that S would have readily dissolved into droplets of Fe liquid in the accreting Earth. Ahrens and Jeanloz (1987) found that a core with about 11% S would match the observed seismic properties of the core reasonably well. Allègre et al. (1995) estimate the core has about 2% S, based on the correlation between K/Ca and S/Ca in meteorites and an estimate of the terrestrial K/Ca value.

The idea that O might be the light element in the core is based on the observation that at high pressure FeO is miscible with Fe liquid, though it is not at low pressure. The eutectic composition in the Fe-FeO system at 16 GPa, still well below core pressures, contains about 10% FeO (e.g., Ringwood and Hibbertson, 1990). Furthermore, the Earth's mantle is depleted in oxygen compared with CI chondrites. The relevance of this last observation, however, is questionable, since the oxidation state of chondrites, and therefore presumably the solar nebula, clearly varied widely. Significant in this respect is the observation that based on O isotope ratios, the Earth appears to be more closely related to enstatite chondrites (Figure 9.23), which are highly reduced, than to other chondrites, and in particular to CI. Allègre et al. (1995) have estimated that the core contains about 4% O.

C, Si, P, Mg, and H have also been mentioned as the possible light element in the core. Allegre et al. (1995) estimated the core contains about 7% Si, which is based on the observed mantle deficit of Si and the assumption the Earth has the same Mg/Si ratio as carbonaceous chondrites. There are, however, some inconsistencies and questionable assumption in their approach. At present, the available observations simply do not allow us to specify the core composition very accurately.

The core undoubtedly contains the bulk of the Earth's inventory of other siderophile elements as well. Refractory siderophiles such as Re, Os, Ir, and Pt, among others, are probably present in chondritic relative abundance in the core, based on the same arguments made for chondritic relative abundances of lithophile elements. If the light element in the core is S, the core may also be rich in chalcophiles such as Pb.

It is generally assumed that core formation occurred very early in Earth's history, probably simultaneously with accretion. One line of evidence for this is the observation that the gravitational energy released by formation of the core from an initially homogeneous Earth is sufficient to raise the temperature of the Earth by 2000° C. Thus if core formation were catastrophic, i.e., happened rap-

PARTITION COEfficients $D^{S/M}$ $D^{S/M}$ Element Element 1×10^{-4} W 1 Au 5×10^{-4} Р 0.24 Re 5×10^{-5} Ga 0.8 Ir 8×10^{-4} 0.01

Ag

Pb

TABLE 11.4. LIQUID SILICATE-LIQUID METAL

from Jones and Drake (1986).

 2×10^{-4}

 7×10^{-3}

Mo

Ni

Co

idly, it must have occurred before formation of any presently existing crust; i.e., before 4.2-4.3 Ga ago (the age of zircons recovered from the Jack Hills of Australia). Modelling of the physical aspects of accretion also suggests core formation would have occurred simultaneously with accretion. In addition, core formation appears to have occurred vary early in meteorite parent bodies. Prima facie evidence of this is the existence of anomalies in the isotopic composition of Ag resulting from decay of shortlived ¹⁰⁷Pd. Certainly all those who have

0.15

CHAPTER 11: THE MANILE AND CORE

considered the problem agree that core formation *started* very early. There has been some question as to when it was complete, with some arguing that it continued through most of geologic time.

As we found in Chapter 7, trace elements are useful in geochemistry because geochemical processes often lead to much greater variations in their abundances than in the major elements. Equilibration between me-

 Table 11.5.
 Siderophile and Chalcophile Element

 Abundances in the Upper Mantle

Element	Concentration/CI	Element	Concentration/CI
W	0.22	Au	0.05
Re	0.007	Р	0.08
Ir	0.007	Ga	0.38
Mo	0.064	Ag	0.04
Ni	0.17	Pb	0.07
Со	0.21		

tallic (Fe-Ni) liquid and silicate liquid is an example of a process that should lead to a large fractionation of trace elements. Table 11.4 lists experimentally determined liquid silicate–liquid metal partition coefficients for a number of siderophile elements. These experiments were performed at low pressure with about 20% sulfur in the metal phase. Their relevance to the problem of core formation thus depends on 1) whether the core contains sulfur, 2) whether equilibration occurred at high or low pressure, and 3) the form (liquid or solid) of the silicate and metal phases at the time and point of equilibration. More fundamentally, it depends on whether core formation was an equilibrium or disequilibrium process.

There is as yet little consensus on any of these points. Most students of core formation prefer lowpressure equilibration. New material is added to the surface of the earth and hence the initial interaction between metal and silicate would occur near the surface at low pressures. Presumably, blobs (for lack of a better word) of metal grow and coalesce as they sink. The greater their size, the less the opportunity for equilibration with surrounding silicate. So on these grounds, low pressure partition coefficients may be more relevant than high pressure ones. Solid silicate–liquid silicate and solid metal–liquid metal partition coefficients for these elements are also available. From these data, any combination of metal-liquid equilibrium distribution coefficients can be calculated.

Accepting for the moment the relevance of the distribution coefficients in Table 11.4 to core formation, we can see that these values predict that the mantle should be highly depleted in elements such as Mo, Re, Au, and Ir. The partition coefficients can be compared with depletion factors (abundances relative to CI chondrites) in the mantle shown in Table 11.5. Although there is very rough qualitative agreement, there are significant discrepancies in detail. For example, Ir should be an order of magnitude more depleted than Re, and Ni should be a factor of 30 more depleted than Co. But the actual depletions of Ni and Co are similar, as are the depletion factors for all the platinoid metals, all of which are highly siderophile (Figure 11.10). The mismatch tells us something is wrong with the partition coefficients or a simple model of equilibrium core formation. Assuming equilibration took place between solid, rather than liquid, silicate and metal generally only makes things worse.

Possible explanations for the discrepancy between predicted and observed depletion of siderophiles in the mantle include inefficient metal extraction or disequilibrium, discrepancy between the composition of the metal phase in experiments and the core (i.e., non-S bearing, O-bearing metal). All of these possibilities must be considered, but perhaps the best hypothesis is that core formation ended before the Earth had completely accreted. To illustrate the effect of this, imagine that core formation effectively extracts all Re and Ir from the mantle, as would be the case for complete equilibrium with distribution coefficients given in Table 11.4. Then imagine new material equivalent to 1% of the mass of the mantle and with chondritic abundances of Re and Ir is added to the mantle. The mantle would be highly depleted in these elements, concentrations of these elements would be close to 0.01 times the chondritic abundances, but the Re/Ir ratio would be close to chondritic, despite the order of magnitude difference in their distribution coefficients. Yet another possible explanation relates to how metal and silicate are distributed during the hypothesized giant impact (e.g., Newsom and Taylor, 1989).

Geochemistry

CHAPTER 11: THE MANILE AND CORE

Some authors (e.g., Allègre et al., 1983) have ascribed the excess radiogenic nature of mantle Pb isotope ratios (i.e., that they lie to the right of the geochron) to continued growth of the core and extraction of Pb (and other things) from the mantle long after formation of the Earth. This would, of course produce higher U/Pb ratios, and ultimately higher ²⁰⁶Pb/²⁰⁴Pb, ²⁰⁷Pb/²⁰⁴Pb, and ²⁰⁸Pb/²⁰⁴Pb than in a closed system. Newsom et al. (1986) showed that this is unlikely by examination of the relationship between Pb isotopes and siderophile element concentrations in mantle-derived volcanic rocks. They found that the ratio of Mo, a siderophile, to Pr, a lithophile light rare earth, was uniform in the mantle and did not correlate with Pb isotope ratios. A correlation is expected if post-accretion core growth extracted Pb from the mantle because Mo would be more drastically extracted than Pb (Table 11.4), decreasing the Mo/Pr ratio and increasing the U/Pb ratio. That the Mo/Pr ratio is uni-

form implies that in the modern mantle, Mo behaves as do moderately incompatible lithophile elements, such as Pr. Apparently, no part of the mantle sampled by volcanism has experienced more siderophile depletion than any other part. In a subsequent study, Sims et al. (1990) found that the Mo/Ce ratio in mantle-derived rocks has been constant through time. Thus the core has not extracted siderophile elements from any part of the mantle sampled by volcanism.

MANTLE GEOCHEMICAL RESERVOIRS

Evidence from Oceanic Basalts

Isotopic variations in oceanic basalts provide clear evidence that the mantle is presently heterogeneous. Oceanic basalts provide better evidence of this than continental basalts because the possibility of the former being contaminated by the crust through which they pass is much reduced. This is true for 3 reasons: oceanic crust is much thinner, it has a higher solidus, and it is compositionally similar to melts of the mantle (so that when assimilation does occur, it chemical affects are minimized). Of course, not all continental basalts appear to have been contaminated. Those that have apparently not been contaminated also show evidence of mantle heterogeneity. Xenoliths in basalts also provide evidence of mantle heterogeneity.

The fundamental isotopic variation in oceanic basalts is between mid-ocean ridge basalts (MORB) and oceanic island basalts (OIB). Figures 11.14 through 11.16 illustrates the differences in isotopic composition between these two groups as well as island arc volcanics, which we shall discuss in Chapter 12. In general, MORB have lower Sr and Os and higher Nd and Hf isotope ratios than OIB.



Figure 11.14 Comparison of ⁸⁷Sr/⁸⁶Sr ratios in midocean ridge basalts (MORB), oceanic island basalts (OIB) and island arc volcanics (IAV).

Geochemistry

CHAPTER 11: THE MANILE AND CORE



Figure 11.15. Comparison of ϵ_{Nd} and ${}^{3}\text{He}/{}^{4}\text{He}$ in mid-ocean ridge basalts (MORB), oceanic island basalts (OIB) and island arc volcanics (IAV).

These differences are consistent with MORB being derived from a mantle reservoir with consistently lower relative abundances of incompatible elements than OIB. For example, low ⁸⁷Sr/⁸⁶Sr ratios imply low time-integrated Rb/Sr ratios of the MORB source, and Rb is known to be more incompatible than Sr.

The situation with Pb isotope ratios is not quite so clear. As may be seen in Figure 11.17, there is considerable overlap between MORB and OIB in ²⁰⁶Pb/²⁰⁴Pb ratios, though OIB are on average higher. We found in Chapter 8 that Pb isotope ratios plotting to the right of the geochron imply time-integrated U/Pb ratios in the mantle that are higher than in the bulk silicate Earth, and this is the opposite of what would be expected from the relative compatibilities of U and Pb.

He isotope ratios also show greater complexity than Sr, Nd, and Hf isotope ratios (for historical reasons, He isotope ratios are generally expressed with ⁴He, the radiogenic isotope, in the denominator). He isotope ratios in MORB are relatively uniform with a mean of about 8 R/R_A^* , whereas OIB have ratios both higher and lower than this. He isotope ratios in the mantle are affected by 2 factors: growth of radiogenic ⁴He from U and Th (Sm and other long-lived alpha emitters do not contribute significant ⁴He) and degassing of primordial He. The conventional interpretation of higher ³He/⁴He ratios in many OIB is that these basalts come from reservoirs are less degassed that the MORB-source reservoir.

Figure 11.17 illustrates the difference between MORB and OIB in rare earth elements. Relative to chondrites, the light rare earths in MORB are underabundant compared to the heavier rare earths. The easiest way to produce such depletion in the mantle is to extract a melt from it. The incompatible elements, including the light rare earth elements (LREE), partition preferentially into the melt leaving a residue depleted in light rare earths. When the mantle melts again, the basalt produced will inherit this light rare earth depletion (though the basalt will not be as LREE depleted as its mantle source). Thus the mantle source of midocean ridge basalts is thought to have suffered melt extraction in the past. The obvious place to store the incompatible elements extracted from the MORB source is the continental crust, a reservoir of very considerable size known to be enriched in these elements. At least in this gross sense, the evolution of the source of MORB is thought to be understood.

^{*} Recall from Chapter 8 that He isotope ratios are typically reported relative to the atmospheric ratio, $R_A = 1.39 \times 10^{-6}$, hence a ratio of 8 R/R_A is equal to $8 \times 1.39 \times 10^{-6} = 1.11 \times 10^{-5}$.



CHAPTER 11: THE MANILE AND CORE



Figure 11.16. Comparison of Pb isotope ratios in MORB, OIB, and IAV.

remainder are referred to as normal or N-MORB.

OIB are generally light rare earth-enriched to varying degrees. Part of the difference in rare earth patterns of MORB and OIB may be due to the smaller degrees of melting involved in the generation of OIB magma. Many, though by no means all, OIB are alkalic, whereas almost all MORB are tholeiitic. Alkali basalts are generally thought to be produced by smaller degrees of melting than tholeiites. Not all the differences in rare earth patterns between OIB and MORB can be attributed to degree of melting — there must also differences in the rare earth abundances in the mantle sources of these magmas.

There do seem to be some inter-oceanic differences in the composition of MORB. The Pacific has the lowest ⁸⁷Sr/⁸⁶Sr ratios, the Atlantic the highest ¹⁴³Nd/¹⁴⁴Nd ratios and the Indian the lowest ²⁰⁶Pb/²⁰⁴Pb ratios. Thus the variation is not systematic in the sense of MORB from one ocean being more depleted than those from another. Another difference is that isotopic compositions in Pacific MORB are more uniform than those from other oceans. Some MORB, particularly those from ridge segments near oceanic islands (e.g., the Mid-Atlantic Ridge near the Azores and Iceland; the Galapagos Spreading Center near the Galapagos, etc.), have geochemical characteristics similar to OIB. These MORB are referred to as E-MORB (incompatibleelement enriched) or P-MORB (mantle plume). The

Figure 11.18 compares the incompatible element abundances in MORB and typical OIB alkali basalt in a *spider diagram*, in which elements are ordered according to their incompatibility and normalized to primitive mantle values. Thus this kind of plot is also sometimes referred to as an *extended rare earth plot*. The incompatible element depletion of MORB is apparent. OIB are generally incompatible-element enriched, but the most incompatible elements (Cs, Rb, and Ba) are not as enriched as are the slightly less incompatible elements. Though part of this may be attributed to uncertainty in the choice of normalizing values for these elements (i.e., the primitive mantle abundances), this depletion in the most incompatible elements in OIB sources suggests even they have experience prior depletion in incompatible elements due to melt extraction. This is consistent with Nd isotope ratios, which in most OIB are higher than chondritic, implying time-averaged higher than chondritic Sm/Nd ratios, and depletion in Nd relative to Sm.

In the view of most mantle geochemists and geophysicists, MORB are derived from the asthenosphere, which is that part of the upper mantle that directly underlies the lithosphere and is characterized by relatively low seismic velocities (Figure 11.3). Oceanic island basalts are thought to be derived from mantle plumes that rise from a thermal boundary layer at either the 660 discontinuity or the core-mantle boundary. There are some fairly straight-forward reasons for this. MORB are by far the most abundant volcanic rocks on the Earth and they are also compositionally the most uniform. This kind of basalt is erupted, with a few exceptions, wherever plates are moving apart. This is most easily explained if the MORB source is relatively close to the surface. Furthermore, the pull of subducting lithospheric slabs has been shown to be the primary force acting on lithospheric plates. Upwelling of mantle beneath mid-ocean ridges appears to be largely a passive response to the plate

Geochemistry





Alternative models have been proposed, however. For example, D. Anderson argues that the source of OIB basalts is the uppermost mantle 200 or 300 km of mantle, which in his model is lherzolitic. According to him, this region is underlain by piclogite, which extends down to the 660 discontinuity, and which is the source of MORB. Though strong arguments can be mounted against such models, they are difficult to disprove. They also appear to violate the 'elegance axiom' in that they are more complex. Here we adopt the majority viewpoint that the upper mantle is the source of MORB and that oceanic island basalts are derived from mantle plumes, which rise from the deep mantle. Let's now consider the evolution of these mantle reservoirs.

Evolution of the Depleted Upper Mantle

Ancient mantle-derived rocks preserved in the continents provide evidence of how the mantle has evolved through time. Figure 11.19 shows initial ϵ_{Nd} values of mantle-derived igneous rocks. These data reveal the mantle had a positive ϵ_{Nd} by 3.8 Ga. The positive ϵ_{Nd} in turn implies the mantle was already LREE depleted. Indeed, since time is required before an increase in Sm/Nd ratios will be reflected in ¹⁴³Nd/¹⁴⁴Nd, the depletion in light rare earths, and presumably other incompatible ele-

motion induced by this "slab-pull". Again, this suggests that the mantle that gives rise to MORB is fairly shallow. In contrast, there is clear evidence, such as topographic swells, of dynamic mantle upwelling associated with oceanic islands. That the locus of active volcanism remains approximately stationary for periods as long as 100 million years, indicating that this upwelling rises from depths beneath the convection associated with plate motions. The thrust of the reasoning is that it is easier to poke a few plumes though MORB-source asthenoа sphere than to get a lot of



Figure 11.17. Rare earth patterns of mid-ocean ridge basalts and oceanic island basalts.

Geochemistry

CHAPTER 11: THE /MANTLE AND CORE



Figure 11.20. Comparison of the effects of fixed continental volume and increasing continental volume on the Sm/Nd ratio and ε_{Nd} evolution in the depleted mantle. A constant continental mass results in a constant Sm/Nd ratio of the depleted mantle and linear evolution of ε_{Nd} in the mantle. Growth of the continents through time leads to and increasing Sm/Nd of the mantle and a concave path for upward evolution of ε_{Nd} in the mantle.

not immediately mixed back into the upper mantle.

The idea certainly has some attractive features. First, a large amount of oceanic crust could have been created during early Archean times, and this process would certainly deplete the mantle in incompatible elements. Since heat flow was presumably higher in the early Archean than today, it is quite possible that plate velocities were higher then, implying more rapid creation and destruction of oceanic crust. Chase and Patchett (1988) calculate that if the rate of oceanic crust creation through time has decreased in proportion to radioactive heat production in the Earth, some 2×10^{23} g of oce-

anic crust would have been created by 3.5 Ga. If the percentage melting involved in the creation of oceanic crust is 10% (which is roughly the case for present oceanic crust), melt would have been extracted from some 2×10^{24} g of mantle, or about half the mass of the mantle, by 3.5 Ga. The resulting depletion in Nd relative to Sm would be sufficient to produce the ϵ_{Nd} values observed in the early Archean rocks preserved in continents.

Yet another possibility is that an early continental crust did form, but it was subsequently destroyed and recycled into the mantle. This possibility was first suggested by R. L. Armstrong in 1968, long before the isotopic data in Figure 11.19 were ments, must have occurred hundreds of millions of years before the formation of the 3.8 Ga rocks. Thus we can trace the depletion of the upper mantle back to the earliest 500 Ma of Earth history.

Incompatible elements lost from the upper mantle must be stored elsewhere. It is generally assumed that this complimentary incompatible-element enriched reservoir is the continental crust. If this was true in early Archean times, the data in Figure 11.20 implies an extensive continental crust had formed prior to 4.0 Ga. However, very little continental crust of this age has been found. It is possible that such crust simply has not been identified, but that seems unlikely. Chase and Patchett (1988) argued that the complimentary reservoir to the depleted mantle was mafic or ultramafic oceanic crust, which was



Figure 11.19. ε_{Nd} in rock suites for which there is little evidence of involvement of much older crust in their genesis (after Smith and Ludden, 1989).

Geochemistry

CHAPTER 11: THE /MANILE AND CORE

available. Armstrong argued that the continental crust reached its present mass by about 4.0 Ga and crustal mass has subsequently remained constant. Armstrong was the first to recognize the possibility that plate tectonics, which then was a new and revolutionary theory, provided a means of transporting material from the crust to the mantle. In Armstrong's model, new continental crust is continually created, but this creation is balanced by destruction of crust through erosion, deposition of the sediments on the oceanic crust, and subduction of this sediment into the mantle.

There are a number of reasons to believe that continental crust has been recycled into the mantle. As we shall see in the next chapter, there have clearly been additions to the continental crust through geologic time. If there has been no accompanying destruction of crust, the volume and mass of crust should have increased through geologic time. Armstrong pointed out that this should have led to flooding of the continents, as the oceans are confined to a smaller and smaller area. But this has not occurred. Armstrong also noted that virtually no deep ocean pelagic sediment is preserved on continents, implying that it is carried into the mantle during subduction. Though it was at one time widely believed that the sedimentary veneer on the oceanic crust is scraped off or "obducted" in subduction zones, careful study of several trenches shows little or no accumulation of sediment despite tens of millions of years of continuous subduction.

From a geochemical perspective, growth of the continental crust through time should lead to increasing incompatible element depletion of the upper mantle. This leads, for example, to an increase in the Sm/Nd ratio through time, which should result in the ε_{Nd} of the mantle following a concave upward path (Figure 11.20). In actuality, ε_{Nd} in the depleted mantle appears to follow a linear evolution, implying that the mass of the continents has not grown through time. However, there is an alternative explanation, proposed by Patchett and Chauvel (1984). They pointed out that if a growing continental mass could still result in constant Sm/Nd of the mantle if the volume of the depleted mantle grew.

Let's consider now consider the question of the volume of mantle that would have to be depleted in incompatible elements to create the continental crust. We start by assuming that the Earth consists of three reservoir: the continental crust, the volume of mantle depleted in incompatible elements as a result of formation of this crust, and an undepleted, or primitive mantle (Figure 11.21). For any radioactive decay system we can write a series of mass balance equations. For the Nd isotope system, we assume that the bulk Earth has ε_{Nd} of so, so we may write:

$$\sum_{j} \mathbf{M}^{j} \mathbf{C}^{j} \boldsymbol{\varepsilon}_{\mathrm{Nd}}^{j} = 0 \qquad \qquad 11.1$$

where M^{j} is the mass of the *j*th reservoir, C^{j} is the concentration of Nd in that reservoir, and ϵ^{j}_{Nd} is the value of ϵ_{Nd} in that reservoir. We also assume the Sm/Nd is chondritic. We'll use $f_{Sm/Nd}$ to denote the relative deviation of the Sm/Nd ratio from the chondritic value, i.e.:

$$f_{Sm/Nd} = \frac{\frac{147}{M}Sm^{144}Md^{-147}Sm^{144}Md_{CHUR}}{\frac{147}{M}Sm^{144}Md_{CHUR}}$$
11.2

Then we may write a similar mass balance for the Sm/Nd ratio for the Earth:

$$\sum_{j} M^{j} C^{j} f^{j}_{Sm/Nd} = 0$$
 11.3

The mass balance for the Nd concentration is:

$$\sum_{i} M^{j} C_{Nd}^{j} = M^{o} C_{Nd}^{o}$$
 11.4

where M^{o} is the mass of the silicate Earth and C^{o}_{Nd} in the concentration of Nd in the silicate Earth. Finally, the masses of our three reservoirs must sum to the mass of the silicate Earth:

$$\sum_{i} M^{j} = M^{o}$$
 11.5

Since the half-life of ¹⁴⁷Sm is long compared to the age of the Earth, we may use the approximation:

CHAPTER 11: THE /MANTLE AND CORE

$$e^{\lambda t} = \lambda t + 1$$

and hence:

143
Nd/ 144 Nd = 143 Nd/ 144 Nd_i + 147 Sm/ 144 Nd λ t

The equation may be transformed into epsilon notation, in which case it becomes:

$$\varepsilon_{\rm Nd} \cong \varepsilon_{\rm Nd}^{\rm l} + f_{\rm Sm/Nd} Q t \qquad 11.7$$

11.6

where $\varepsilon_{\text{Nd}}^1$ is the initial value of ε_{Nd} (i.e., at t = 0), and Q and $f_{\text{Sm}/\text{Nd}}$ are defined as:

$$Q_{Nd} = \frac{10^{4} \lambda^{147} \text{Sm}^{144} \text{Nd}_{CHUR}}{143 \text{Nd}^{144} \text{Nd}_{CHUR}}$$
 11.8

 Q_{Nd} is a constant with a value of 25.13 Ga⁻¹.

Assuming that the crust has grown from primitive mantle, then

$$\varepsilon_{Nd}^{c} = f_{Sm/Nd}^{c}QT^{c}$$
 11.9

where T^c is the average age of the crust. If the Earth consists of only three reservoirs for Nd, namely the continental crust, depleted mantle, and primitive mantle, and if the depleted mantle and crust evolved from a reservoir initially identical to primitive mantle the mass balance equation 11.1,11.3, and 11.4 must hold for crust and depleted mantle alone. In this case, these equations can be combined with 11.9 to derived a relationship between the mass of the crust and the mass of the depleted mantle:

$$\mathbf{M}^{dm}/\mathbf{M}^{c} = \left(\frac{\mathbf{C}_{Nd}^{c}}{\mathbf{C}_{Nd}^{o}} - 1\right) - \left(\frac{\mathbf{C}_{Nd}^{c}}{\mathbf{C}_{Nd}^{o}}\right) \frac{\mathbf{Q}\mathbf{f}_{Sm/Nd}^{c}\mathbf{T}^{c}}{\mathbf{\epsilon}_{Nd}^{dm}} \qquad 11.10$$

Thus the mass ratio of depleted mantle to crust can be calculated if we know the Sm/Nd ratio of the crust, the ε_{Nd} of the depleted mantle, and the concentration of Nd in the crust and in primitive mantle. Figure 11.22 is a plot showing the solutions of 11.10 as a function of T^C for various values of ε^{dm} obtained by DePaolo (1980). Most estimates of the average age of the crust are between 2 and 2.5 Ga, and ε^{dm} is about +10. Possible solutions for the ratio of depleted mantle to whole mantle are in the range of 0.3 to 0.5, i.e., somewhere between a third and a half of the mantle must be de-

pleted to the point of having ε_{Nd} of +10 to create the continental crust. A number of similar mass balance calculations that included other isotopic systems as well were published between 1979 and 1980, all of which obtained rather similar results. We may conclude from this that the volume of depleted mantle must be substantial, occupying at the least much of the upper mantle.

Interestingly, the fraction of the mantle above the 660 km seismic discontinuity is roughly 0.33. Thus these mass balance calculations are consistent with the 660 discontinuity being a chemical boundary between upper and lower mantle. In this case, the lower mantle would appear to be chondritic, at least with respect to rare earth elements.

It is important to remember the assumptions that went into this calculation, in particular, that there are only 3 reservoirs of significance for the rare earths: the crust, the depleted mantle, and primitive mantle. If some other reservoir with a significant mass of rare earths and ε_{Nd} very different from 0 exists, then the calculated volume of the depleted mantle could be significantly larger or smaller than calculated above. Two reservoirs in particular cause concern in this respect: the first is the subcontinental lithosphere and the second is the region of the deep mantle that gives rise to mantle plumes (discussed below). At present, the mass of rare earths in the subcontinental lithosphere does not seem to be large enough to invalidate the assumptions made. The mantle plume source region



Figure 11.21. The three reservoir model of the mantle. The depleted mantle is the source of MORB and has $\varepsilon_{Nd} = +10$, the lower mantle is primitive and has bulk Earth characteristics, e.g., $\varepsilon_{Nd} = 0$.

Geochemistry

CHAPTER 11: THE MANTLE AND CORE



Figure 11.22. The relationship between ratio of mass of the depleted mantle to mass of the continental crust as a function of mean age of the crust calculated from equation 18.6 using various values of ε_{Nd} for the depleted mantle. The arrows at the bottom enclose the range of probable values for the mean age of the crust.

is also often assumed to be small, but this need not be the case. The present plume flux has been estimated by Sleep (1990) to be about 2.56×10^{17} g/yr (assuming an average excess temperature for plumes of 225° C). If material resides in some plume source layer (such as D'') for 10^9 yr for rising in a mantle plume, and if the plume flux is constant, then the mass of this plume source layer must be 2.5×10^{23} kg. This is ten times the mass of the crust. Thus if Sm and Nd concentrations in plumes are as much as 10% of the crustal concentration and this is included in the mass balance calculations above, the mass of the depleted mantle could be nearly twice as large as allowed by DePaolo's calculation. Thus these mass balance calculations must be viewed with some skeptism.

Nd isotope systematics are thus consistent with a fairly simple model of the mantle evolution: generation of the continental crust has depleted a third or more of the mantle. Some recycling of continental crust to the mantle can readily account for the apparent linear growth of ε_{Nd} in the mantle through time. Although difficult to document because initial Sr isotope ratios are poorly preserved due to the mobility of Rb and Sr, Sr isotope systematics appear to be consistent with this model, as

are Hf isotope systematics (though Hf isotopes suggest higher rates of recycling). However, Pb isotope systematics of the mantle and crust are not consistent with this simple picture, and hint at a much more complex evolution.

Pb isotope ratios do not seem to be consistent with such simple evolutionary models. As we pointed out in Chapter 8, Pb isotope ratios of MORB plot to the right of the Geochron, implying μ (²³⁸U/²⁰⁴Pb) has increased in the depleted mantle. From the incompatibilities of U and Pb, we would have expected this ratio to have decreased. Furthermore, a mass balance problem may exist since both the upper crust and the mantle have more radiogenic Pb than the bulk Earth, which must lie on the Geochron, and it is unclear whether this can be balanced be unradiogenic Pb from the lower crust alone.

Th isotope ratios in mid-ocean ridge basalts may point the way to a solution of the dilemma posed by Pb isotope ratios. As we saw in Chapter 8, the ²³⁰Th/²³²Th ratio in the young basalts is directly proportional to the ²³⁸U/²³²Th (κ) ratio in the mantle at the time the basalts were generated. Since this value of κ is derived from Th isotope ratios, we'll designate it as κ_{Th} . On the other hand, the ratio of radiogenic ²⁰⁸Pb to ²⁰⁶Pb is proportional to the *time-integrated* ²³⁸U/²³²Th (κ) ratio. Let's explore this a bit further.

We can write a radiogenic growth equation for both ²⁰⁸Pb and ²⁰⁶Pb:

$${}^{208}\text{Pb}/{}^{204}\text{Pb} = {}^{208}\text{Pb}/{}^{204}\text{Pb}_{i} + {}^{232}\text{Th}/{}^{204}\text{Pb}(e^{\lambda_2 t} - 1)$$
 11.11

rearranging:

$${}^{208}\text{Pb}/{}^{204}\text{Pb} - {}^{208}\text{Pb}/{}^{204}\text{Pb}_{i} = {}^{232}\text{Th}/{}^{204}\text{Pb}(e^{\lambda_{2}t} - 1)$$
 11.12

The term on the left is the radiogenic ²⁰⁸Pb, which we will designate as ²⁰⁸Pb*. Thus:

$${}^{208}\text{Pb}*/{}^{204}\text{Pb} = {}^{232}\text{Th}(e^{\lambda_2 t} - 1)$$
 11.13

Designating the radiogenic ²⁰⁶Pb as ²⁰⁶Pb*, and analogous equation may be written for ²⁰⁶Pb:

$${}^{206}\text{Pb}^{*}/{}^{204}\text{Pb} = {}^{238}\text{U}/{}^{204}\text{Pb}(e^{\lambda_8 t} - 1)$$
 11.14

November 25, 1997



Figure 11.24. Correlation of Rb/Sr and μ (²³⁸U/²⁰⁴Pb) in MORB. Solid line is a regression line, dashed lines represent errors on the regression. The value of Rb/Sr in the depleted mantle must be lower than 0.0045, based on Sr isotope evolution. From the correlation, this implies μ < 6.33. From White (1993).

$$\frac{{}^{208}\text{Pb}^*}{{}^{206}\text{Pb}^*} = \frac{{}^{208}\text{Pb}/{}^{204}\text{Pb} - 29.476}{{}^{206}\text{Pb}/{}^{204}\text{Pb} - 9.307}$$
11.16

From the ${}^{208}\text{Pb}^{*}/{}^{206}\text{Pb}^{*}$ ratio, we can calculated a time-averaged κ , which we denote as κ_{Pb} , by rearranging equation 11.15:

$$\kappa_{Pb} = \frac{{}^{208}\text{Pb}^{*}}{{}^{206}\text{Pb}^{*}} \frac{(e^{\lambda_{gt}} - 1)}{(e^{\lambda_{gt}} - 1)}$$
 11.17

where *t* is now set to 4.55 Ga, the age of the Earth.

Galer and O'Nions (1985) compared values of κ_{Pb} with κ_{Th} in MORB. They found that the values of κ_{Th} averaged about 2.5, whereas the time-integrated ones averaged about 3.7. The value of κ in *Orgueil* (CI-1) is about 3.8, which they took as the primitive mantle value. The time-integrated value is barely lower than the primitive mantle value, and much higher than the present value. From this they concluded that Pb now in the depleted mantle could not have resided there for very long. In other words, the depleted mantle appears to be open to the flow of Pb, with both inputs and outputs. They calculated the residence time of Pb in the depleted mantle, the average time spent in the depleted mantle by an atom of Pb, to be about 600 million years.

As we saw earlier in this chapter, there is reason to believe the terrestrial value of κ may be somewhat higher than that in Orgueil, perhaps 4 or 4.2. However, with additional data, it appears the average value of κ_{Th} in MORB is lower than 2.5, perhaps about 2.2 (Figure 11.23). The residence time for Pb calculated using these parameters is about 1 Ga, which is still short compared to the age of the Earth.

Dividing 11.13 by 11.14, we have:

$$208$$
 pt * 232 Th (a λ st = 1)

Geochemistry

$$\frac{{}^{208}\text{Pb}^{*}}{{}^{206}\text{Pb}^{*}} = \frac{{}^{232}\text{Th}(e^{\lambda_{2}t} - 1)}{{}^{238}\text{U}(e^{\lambda_{8}t} - 1)} 11.15$$

For any fixed time the exponential term is a constant, so that $^{208}\text{Pb}^*/^{204}\text{Pb}^*$ is proportional to the average κ over that time. Taking the initial isotope ratios to be the meteoritic values (i.e., Canyon Diablo troilite), we can calculate $^{208}\text{Pb}^*/^{206}\text{Pb}^*$ as:





Geochemistry

CHAPTER 11: THE MANILE AND CORE

White (1993) addressed the question of the ²³⁸U/²⁰⁴Pb (μ) ratio in the depleted mantle, and concluded it was less than about 6, which is lower than the bulk Earth value of about 8. Unlike κ , μ cannot be calculated directly. It can, however, be constrained in several ways. Figure 11.24 illustrates one example. The Rb/Sr ratio in MORB is strongly correlated with μ . From the ⁸⁷Sr/⁸⁶Sr ratio in MORB, the maximum Rb/Sr ratio in the depleted mantle can be constrained to be less than 0.0045. Based on the correlation, the value of μ in the depleted mantle must be less than 6.33. The low present μ in the MORB contrasts with the time-integrated one deduced from the ²⁰⁶Pb/²⁰⁴Pb ratio, the latter being about 9. This difference also leads to the conclusion that the Pb now in the depleted mantle could not have resided there for very long.

These results suggest a very different model for evolution of the upper mantle than the one based on Nd isotope ratios discussed above. They suggest that radiogenic isotope ratios and concentrations of incompatible elements are controlled not simply by the degree of depletion resulting from partial melting, but by the ratio of the flows into and out of the mantle. In other words, the mantle appears to be a *kinetically* controlled system rather than a thermodynamically controlled one. The large scale reservoirs on the surface of the Earth, such as the oceans and the atmosphere, are known to be kinetically controlled in this same sense. So this result should not be too surprising. It appears that the compositions of all large chemical systems on the Earth are controlled, at least in part, by kinetic factors.

If there is a flux of Pb and other elements into the depleted mantle, where is this Pb coming from? The continental crust can be easily ruled out as having inappropriate Pb isotope composition. A lower "primitive mantle", originally suggested by Galer and O'Nions (1985), can also be ruled out. White (1993) concluded mantle plumes were supplying this Pb. There are several reasons to believe this is the case. First, they have the appropriate isotopic composition. Second, only about 10% of the total plume flux is required to supply the necessary Pb to the depleted mantle. Third, if the suppositions that the depleted mantle occupies the uppermost mantle and that plumes come from deeper are correct, then plumes must transit the upper mantle, providing the opportunity to deliver Pb and other incompatible elements to it. Finally, numerous studies, particularly by J. G. Schilling and coworkers, have shown that plumes do mix with the depleted mantle. For example, Figure 11.25 shows the variation in Sr isotopes along the northern Mid-Atlantic Ridge. Regions of high ⁸⁷Sr/⁸⁶Sr extend hundreds of km outward from Iceland and Azores, reflecting mixing of these mantle plumes with the asthenosphere.

It is presumably this plume flux that buffers the ε_{Nd} of the depleted mantle as well. As we shall see below, mantle plumes may contain a significant component of recycled oceanic and continental crust. In a sense then, continental recycling may contribute to the flux of material into the depleted



Figure 11.25. Sr isotope variations along the Mid-Atlantic Ridge. After White et al. (1976).

Geochemistry

CHAPTER 11: THE /MANILE AND CORE

upper mantle, although indirectly.

GEOCHEMICAL EVOLUTION OF MANTLE PLUME RESERVOIRS

As Figures 11.14 through 11.17 demonstrate, the geochemistry of OIB is more variable than that of MORB. Oceanic island volcanism is thought to be the surface manifestation of mantle plumes. Mantle plumes are columns of buoyant mantle that rise from the deep mantle. Exactly how deep is uncertain. Fluid dynamic experiments and numerical simulations indicate that they must arise from a thermal boundary layer (a region where heat is transported conductively rather than convectively). The only certain thermal boundary layer in the mantle is the base: the core-mantle boundary. If the thermodynamics of the phase change at 660 km prevents convection accross the 660 km discontinuity, then it would be a thermal boundary layer as well. Thus plume come either from the core-mantle boundary or the 660 km discontinuity. At present, most geophysicists and geochemists suspect it is the former.

There are clear systematic relationships in the geochemistry of plumes. White (1985) found that OIB could be divided into 4 distinct groups based on their Sr, Nd and Pb isotopic compositions. Thus though there are many mantle plumes and each is to some degree geochemically unique, there appear to be a much smaller number of geochemical reservoirs from which plumes are drawn. These groups, named St. Helena, Society, Kerguelen, and Hawaii for a type island, show distinctive correlations between isotope ratios or have distinctive isotopic compositions. For example, OIB of the St. Helena group, which encompasses the Austral Islands (Pacific), the Comores (Indian), and Ascension (Atlantic) as well as St. Helena (Atlantic), have Sr and Nd isotope ratios that plot below the main array (Figure 11.26), and their Pb isotopic compositions are very radiogenic. Indeed, all basalts with very radiogenic Pb (206 Pb/ 204 Pb > 20) plot within or close to the St. Helena field on a Sr-Nd isotope diagram. Similarly, islands of the Society group, which include the Azores (Atlantic), the Marguesas (Pacific), and Samoa (Pacific) as well as the Societies (Pacific), define a shallower Sr-Nd isotope correlation than do basalts from Kerguelen group islands, which include Tristan da Cunha (Atlantic), and Juan Fernandez (Pacific) as well as Kerguelen (Indian). The Kerguelen group OIB always have less radiogenic Pb than the Society group. Zindler and Hart (1986) subsequently renamed they groups as HIMU (St. Helena), EMI (Kerguelen), EMII (Societies), and these names are now more commonly used than the originals. The term HIMU comes from high- μ , as these basalts are characterized by high 206Pb/204Pb. The term EM is an abbreviation for 'Enriched Mantle'. In addition,



Figure 11.26. Isotopic classes of oceanic basalts of White (1985) and end-member compositions of Zindler and Hart (1986).

W. M White

CHAPTER 11: THE MANTLE AND CORE

White (1985) identified MORB as a distinct isotopic class, with which Zindler and Hart (1986) associated the component DMM (Depleted MORB Mantle).

Weaver (1991) found that there were systematic variations in trace elements between these classes of plumes. This is illustrated in Figure 11.27.

Determining how these distinct geochemical reservoirs have evolved is among the most vexing problems in mantle geochemistry. The principal observation to be explained is that mantle plumes invariably have less depleted isotopic signatures than MORB, and the isotopic compositions of some indicate net enrichment in incompatible elements. Initially, mantle plumes were thought to consist of primitive mantle (e.g., Schilling, 1973). Mixing between primitive and depleted mantle can explain the Sr and Nd isotopic compositions of some plumes, but virtually none of the Pb isotope data can be explained this way, nor are the trace element compositions of OIB consistent with plumes being composed of primi-Indeed, although 'primitive tive mantle. mantle' has proved to be a useful hypothetical concept, no mantle-derived basalts or xenoliths have appropriate compositions to be 'primitive mantle'. It is possible, and perhaps likely, that no part of the mantle retains its original, primitive, composition.

Hofmann and White (1982) suggested mantle plumes obtain their unique geochemical



Figure 11.27. Spider diagrams showing differences in incompatible element enrichment among the various classes of oceanic basalts. After Weaver (1991).

signature through deep recycling of oceanic crust (Figure 11.28). Partial melting at mid-ocean ridges creates oceanic crust that is less depleted in incompatible elements than the depleted upper mantle. The oceanic crust is apparently inevitably subducted as virtually none is preserved at the surface, so it clearly is recycled back into the mantle. The question is what becomes of it? Hofmann and White noted that once oceanic crust reaches depths of about 90 km it converts to eclogite as is more dense than peridotite. Because it is rich in Fe, and garnet-forming components, it remains denser than peridotite at all depths greater than 90 km (except, perhaps, just at the 660 discontinuity due to the negative Clapeyron slope). Thus it will sink to the base of the convecting region. If the mantle is chemically stratified, with a Fe-rich lower mantle, the oceanic crust would sink to a thermal boundary layer at the 660 discontinuity. If the entire mantle convects as a single unit, that is if it is not chemically stratified, ocean crust will sink to base of the mantle, becoming embedded in thermal boundary layer there (D´´). Hofmann and White originally suggested radioactive heating would ultimately cause it to become buoyant. It now seems more likely that heat conducted into it from below, from either the lower mantle or the core, may be more important. In any case, upon sufficient heating, it rises, melting near the surface to create intraplate volcanos.

Sediment appears often, if not always, to be subducted along with the oceanic crust (we will discuss the evidence for this in the next chapter). This subducted sediment would also contribute to incompatible element enrichment of plumes. Varying amounts, types, and ages of subducted sediment

W. M White

CHAPTER 11: THE MANTLE AND CORE

may be responsible for some of the geochemical variety displayed by plumes. Since sediment is ultimately derived from the continents, recycling of oceanic crust, continental crust, mantle plumes, and oceanic island basalts may all be part of a grand geochemical cycle. Tectonic erosion of continental crust in subduction zones and delamination of continental crust may be alternative mechanisms for deep recycling of continental crust.

Because the major element chemistry of OIB is often similar to that of MORB, it seems unlikely plumes could be composed entirely of recycled oceanic crust. Presumably they consist primarily of peridotite, with perhaps at most a few tens of percent oceanic crust. However, because the oceanic crust has much higher incompatible element concentrations than peridotite, it provides most of the isotopic and incompatible element "flavor" of plumes.

Trace elements provide some evi-



Figure 11.29. Pb/Ce and ²⁰⁷Pb/²⁰⁴Pb in basalts from the Societies Islands studied by White and Duncan (in press). A calculated mixing line between depleted mantle and sediment passes through the data. Also shown are estimated Pb/Ce ratios of average continental crust and bulk silicate Earth (BSE).



Figure 11.28. Cartoon illustrating the oceanic crustal recycling model of Hofmann and White (1982). Oceanic crust is transformed into eclogite and post-eclogite assemblages upon subduction. It separates from the less dense underlying lithosphere and sinks to the deep mantle where it accumulates. Eventually, it becomes sufficiently hot to form plumes that rise to the surface, producing oceanic island volcanism. After Hofmann and White (1982).

dence that some plumes contain a recycled sediment component. The Pb/Ce ratio is particularly useful indicator of the presence of sediment for several reasons. First, the Pb/Ce ratio is comuniform in MORB and paratively many OIB. Second, the Pb/Ce ratio is an order of magnitude higher in sediments than in the mantle (typically, Ce/Pb is greater than 0.3 in sediments and <0.04 in MORB). Third, sediments have two orders of magnitude higher concentrations of Pb (typically 20 ppm or more) than the mantle (less than 0.05 ppm), so that addition of even small amounts of sediment to mantle shifts the Pb/Ce ratio. Finally, the near constancy of Pb/Ce in most basalts suggests this ratio is not significantly changed in by magmatic processes such as partial melting and fractional crystallization. There is a strong correlation between isotope ratios and Pb/Ce

W. M White

CHAPTER 11: THE /MANILE AND CORE

in basalts from the Society Islands. As Figure 11.29 shows, the correlation is consistent with mixing between recycled sediment and mantle.

An alternative origin for mantle plumes was proposed by McKenzie and O'Nions (1983). They noted the common evidence for incompatible element enrichment in the subcontinental lithosphere (which we discuss in the next section) and suggested this material may occasionally become delaminated. Because it is cold, it would also sink to the deep mantle. As in the case of the Hofmann and White model, it would be stored in a thermal boundary layer, heated, and rise in the form of mantle plumes. However, recent studies have shown that the Os isotope composition of the subcontinental lithosphere is quite distinctive, and quite different from that of mantle plumes, as we shall see in the next section. This rather strongly suggests "delaminated" subcontinental that lithosphere does not contribute to mantle plumes. Because mantle plumes



Figure 11.30. Three dimension plot of ⁸⁷Sr/⁸⁶Sr, ¹⁴³Nd/¹⁴⁴Nd, and ²⁰⁶Pb/²⁰⁴Pb. Most oceanic basalt data plot within a tetrahedron defined by the composition of EMI, EMII, HIMU, and DMM components. Oceanic islands and island chains tend to form elongate isotopic arrays, many of which seem to point toward a focal zone (FOZO) at the base of the tetrahedron. Adapted from Hart et al. (1992).

come in several geochemical varieties, it is possible that both mechanisms operate. Indeed, other as yet unknown processes may be involved as well.

Most oceanic islands show some variability in their isotopic compositions, defining elongated arrays on plots of isotope ratios. Such elongated arrays suggest mixing. This raises the rather obvious question of what is mixing with what. In a few cases, the Comores are a good example, the elongate arrays seems to reflect mixing between different plume reservoirs. The Comores data defines a trend in isotopic space that appears to be the result of mixing between an EMI and a HIMU component. In other cases, such as the Galapagos, the plume is clearly mixing with the depleted upper mantle. However, in many cases, the cause of the isotopic variation is less clear.

Hart et al. (1992) plotted oceanic basalt isotope data in three dimensions, with axes of ⁸⁷Sr/⁸⁶Sr, ¹⁴³Nd/¹⁴⁴Nd, and ²⁰⁶Pb/²⁰⁴Pb (Figure 11.30). Principal component analysis confirmed that 97.5% of the variance in the oceanic basalt isotope data could be accounted for by these ratios (leaving 2.5% to be accounted for by ²⁰⁷Pb/²⁰⁴Pb, ²⁰⁸Pb/²⁰⁴Pb, and ¹⁷⁶Hf/¹⁷⁷Hf). They found that most of the data plotted within a tetrahedron defined by the hypothetical end members EM1, EM2, HIMU, and DMM. They also noticed that many arrays were elongated toward the base of this tetrahedron on the DMM-HIMU join. From this they concluded that in many, if not most cases, mantle plumes appear to mixing with a previously unidentified component, which they named "FOZO" (an acronym for Focal Zone), that has the approximate isotopic composition of ${}^{87}Sr/{}^{86}Sr = 0.7025$, $\varepsilon_{Nd} = +9$, and ${}^{206}Pb/{}^{204}Pb = 19.5$. They suggested that FOZO is the isotopic composition of the lower mantle and that plumes rising from the core mantle boundary entrain and mix with this lower mantle material. It is unclear, however, whether such a composition for the lower mantle can be fitted to reasonable isotopic mass balances for the Earth. A rather similar idea was presented by Farley et al. (1992), who point out that this additional component, which they called "PHEM", seems to be associated with high ${}^{3}\text{He}/{}^{4}\text{He}$. White (1995) concurred with these ideas, but argued that the 8^{7} Sr / 8^{6} Sr of FOZO is higher, and the $\epsilon_{\rm Nd}$

W. M White

CHAPTER 11: THE /MANILE AND CORE

lower, than estimated by Hart et al. (1992) and probably closer to the values chosen by Farley et al. (1992).

If Hart et al. (1992) are correct, this may explain why the isotopic composition of some volcanos, notably those of Hawaii and the Society Islands, change over time. Hawaiian volcanos commonly go through several evolutionary stages: a shield-building stage, during which the volcanic edifice is rapidly constructed and a relatively uniform tholeiitic basalt is erupted, followed by a late stage,

during which eruption rates drop dramatically and composition shift to alkali basalt, and, following a significant hiatus in activity, a post-erosional stage in which small volumes of basanite and nephelinite are erupted. These latter magmas types are thought to be produced by small degrees of melting. Through this sequence, isotope ratios evolve, with ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ decreasing and ϵ_{Nd} increasing (e.g., Chen and Frey, 1985).

White and Duncan (1996) observed a similar pattern in volcanos of the Society Islands. The argued that this evolution in isotopic composition reflects the passage of the volcano over a compositionally zoned mantle plume (Figure 11.31). The most incompatible element enriched material, which is also the hotest, is located in the center, and formed the core of the plume. This is surrounded by a sheath of material that is viscously entrained by the plume as it rises. It is cooler and also not as enriched in incompatible elements. When the volcano is over the core of the plume, the degree of melting is high, giving rise to tholeiitic basalt with high 87Sr/86Sr and high eruption rates during the shield building stage. When the volcano is over the sheath, the extent of melting is smaller, producing small volumes of nephelinite with low ⁸⁷Sr/⁸⁶Sr.

There is also a geographic pattern to both the distribution of mantle plumes and their isotopic compositions. Mantle plumes appear to be preferentially located within regions of slow lower mantle seismic velocities, as may be seen in Figure 11.32. There are two areas where isotopic compositions are particularly extreme (e.g., high ⁸⁷Sr/⁸⁶Sr), one in the southeastern Indian Ocean and South Atlantic, the other in the central South Pacific (Hart, 1984; Castillo, 1989). The anomaly in the Indian Ocean is called the DUPAL anomaly, while that in the South Pacific is called the SOPITA anomaly. Interestingly enough, both anomalies close to regions where lower mantle seismic velocities are particularly slow (Figure 11.32), which indicate low densities. The low density in turn implies that these are regions of high temperatures in the lower mantle. While the exact significance of this remains unclear, it does establish a connection with oceanic island volcanism and lower mantle properties, strengthening the plume hypothesis, and favoring a lower mantle origin for plumes.

There is still much to understand about the nature and evolution of the mantle. Many of the ideas presented here will undoubtedly evolve in the future and new studies pro-





Figure 11.31. Cartoon of a model that explains why isotopic signatures of magmas become more "depleted" as volcanos evolve. During the "shieldbuilding stage" the time of most vigorous growth, the volcano is located directly over the plume, and magmas are derived from the hot core of the plume, which has enriched isotopic signatures (high ${}^{87}\text{Sr}/{}^{86}\text{Sr}$, low ε_{Nd}). Because of lithopsheric plate motion, the volcano will be located over the edge of the plume during later stages, such as in the post-erosional stage. During this stage magmas are derived from the viscously entrained sheath, which has more "depleted" isotopic signatures. After White and Duncan (1996).

W. M White

CHAPTER 11: THE /MANTLE AND CORE



Figure 11.32. Map showing the distribution of mantle plumes (triangles), P-wave velocity anomalies (m/sec) averaged over the whole lower mantle (red lines), and location of the DUPAL and SOPITA isotope anomalies (pale red regions). Mantle plumes are located in regions of slow lower mantle seismic velocities, implying high temperatures. The DUPAL and SOPITA anomalies are located near seismic velocity minima. After Castillo (1989).

vide new insights into the workings of the mantle.

The Subcontinental Lithosphere

Figure 11.33a shows Sr and Nd isotopic variations in continental basalts. The data span a much larger range than oceanic basalts. Some, but not all, of this variation reflects the effects of assimilation of continental crust on the isotopic signatures of the mantle-derived magmas. Assimilation effects can be avoided by considering only the data on peridotite xenoliths in continental basalts, the data for which is shown in Figure 11.33b. As may be seen, the range of values is reduced, but nevertheless much greater than that observed in oceanic basalts. One needs be cautious in directly comparing the heterogeneity observed in xenolith data to basalt data because the two represent different scales of sampling of the mantle. Basalts are created by melting of regions that have characteristic scales of tens of kilometers, and perhaps greater in some cases. The magma generation process undoubtedly averages out very small scale heterogeneities. Xenoliths, on the other hand, have characteristic dimensions of centimeters. Thus variations in isotope ratios in basalts reflect large scale heterogeneity in the mantle, while xenoliths reflect small scale heterogeneity. Despite this, it appears that the subcontinental lithosphere is more heterogeneous, even on relatively large scales, than is the suboceanic mantle.

It appears that the subcontinental mantle can be quite old, and often has the same age as the crust it underlies. Studies of xenoliths and inclusions in diamond from South African kimberlites suggests the mantle is 3–3.5 Ga old in this region, ages similar to that of the South African craton. The greater isotopic heterogeneity of the subcontinental lithosphere probably reflects it long term stability, which allows variations in parent-daughter ratios to be expressed in variations in radiogenic isotope ratios. Convective mixing in the subcceanic mantle will tend to destroy heterogeneity in the subcceanic mantle.

Though many xenoliths have isotopic compositions indicating incompatible element enrichment, others xenoliths show parts of the subcontinental lithosphere can be extremely incompatible element depleted. ε_{Nd} values of +500 have been recorded in garnets in eclogites from the Roberts Victor mine kimberlite. These eclogites appear to be rafts of subducted oceanic crust stranded in the subcontinental lithosphere over 3 Ga ago, an interpretation supported by highly variable oxygen isotope ratios in the eclogites. They apparently suffered extreme LRE depletion around that time, perhaps by a
CHAPTER 11: THE MANTLE AND CORE



Figure 11.33. (a) top. Sr and Nd isotope ratios in continental basalts. (b) bottom. Sr and Nd isotope ratios in xenoliths in continental basalts. After Zindler and Hart (1986).

small degree of melting or dehydration after subduction. Much of the subcontinental lithosphere may consist of mantle from which partial melts have been extracted to form the continental crust. Interesting, when the upper mantle undergoes melting both the melt and residual solid will have a density that is less than the original material. This residue is less dense is because garnet, a very dense phase, is preferentially removed during melting. Thus both the crustal and mantle parts of the continental lithosphere have relatively low density, which may help to explain its stability.

If the subcontinental lithosphere is residual material from which melts have been extracted, why are xenoliths and basalts with "enriched" isotopic signatures so common? What process or processes could have produced this incompatible element enrichment of many parts of the subcontinental lithosphere? One possibility, first suggested by Brooks et al. (1976), is that partial melts from mantle plumes migrate upward into the lithosphere, where they freeze. The extent to which upwelling mantle can melting will depend on the depth to which it rises. Where continental lithosphere pre-

Geochemistry

CHAPTER 11: THE /MANTLE AND CORE

vents plumes from rising above 200 km depth or so, the degree of melting is likely to be quite small, meaning the melts would be quite incompatible element enriched. These melts could then migrate upward into the lithosphere, reacting with it and enriching it in incompatible elements. Yet another possibility is that hydrous fluids released during dehydration of subducting oceanic lithosphere may migrate into the continental lithosphere and react with it (Hawkesworth, 1990). Judging from studies of island arc magmas (Chapter 12), such fluids appear to be particularly enriched in soluble incompatible elements, such as the alkalis and alkaline earths. These processes in which lithosphere reacts with melts or fluids is known as *mantle metasomatism**. Petrographic studies of some xenoliths clearly reveal features, such as the secondary growth of hydrous minerals such as phlogopite (Mgrich mica) and richterite (an alkali-rich amphibole) indicative of such metasomatism.

Recent studies of Os isotope ratios in xenoliths from the subcontinental lithosphere have been particularly enlightening. Most xenoliths derived from below regions of old continental crust have low Os isotope ratios, which imply that low Re/Os ratios were established long ago. The low Re/Os ratios are consistent with the idea that this material undergone partial melting in the past, since Re is moderately incompatible, and would partition into the melt, while Os is highly compatible, and would remain in the solid. Despite their low ¹⁸⁷Os/¹⁸⁸Os ratios, many of these same xenoliths have quite low ε_{Nd} (Figure 11.34). The low ε_{Nd} suggests incompatible element enrichment, and hence would appear to be inconsistent with the high ¹⁸⁷Os/¹⁸⁸Os ratios (Figure 11.34). The explanation of this

paradox appears to be that Os is was not affected by the metasomatism that enriched these regions in incompatible elements and decreased Sm/Nd ratios (e.g., Carlson and Irving, 1994). Apparently, neither Re nor Os are transported effectively by metasomatic fluids. If the fluids are aqueous, this is perhaps not surprising, since these elements have low solubilities under reducing conditions. If the fluids are silicate melts, it is unclear why they do not transport Re. The answer may have to do with dependence of the Re parition coefficient on composition and oxygen fugacity.

Regardless of why it arises, these unusual Os-Nd isotope systematics provide the continental lithosphere with a distinctive isotopic signature, and geochemists a means of identifying continental lithosphere. In an earlier section, we discussed the hypothesis of



Figure 11.34. ϵ_{Nd} vs. γ_{Os} in xenoliths from the subcontinental lithosphere and oceanic island basalts. Despite low and variable ϵ_{Nd} , the subcontinental lithosphere appears to be characterized by systematically low γ_{Os} (γ_{Os} is the percent deviation of the ¹⁸⁷Os/¹⁸⁸Os ratio from the condritic value).

^{*} Metasomatism is defined in metamorphic petrology as a subsolidus process that results in a net change in the composition of the metamorphic rock. Usually this is accomplished by the flow of aqueous solutions through the rock. The term 'mantle metasomatism' is widely used to refer to reaction between rock and silicate liquid as well as between rock and aqueous solution.

Geochemistry

CHAPTER 11: THE /MANTLE AND CORE



Figure 11.35. ϵ_{Nd} of picritic basalts from the Karroo flood basalts plotted against γ_{Os} . The fall between the fields. Three lines, illustrating 3 mixing models (different concentrations of Os and Nd in the end-members), connect the fields of xenoliths from the subcontinental lithosphere and oceanic island basalts (i.e., mantle plumes). The data fall close to these lines, suggesting they are mixtures of melts of lithosphere and a mantle plume. After Ellam et al. (1992).

McKenzie and O'Nions that subcontinental lithosphere can delaminate and sink to the bottom of the mantle where it is incorporated into mantle plumes. The distinctive isotope signatures of mantle plumes on the one hand and subcontinental lithosphere on the other (Figure 11.34) is inconsistent with this hypothesis.

Continental flood basalts provide another interesting example. These are hugh outpouring of basaltic lava that apparently occurred within relatively short time intervals, a few million years and possibly less in some cases. The great oceanic plateaus, such as Ontong-Java and Kerguelen are the marine equivalents. A number of continental floods basalts can be clearly associated with mantle plumes. For example, the Deccan Traps erupted 65 million years ago when India lay directly over the Reunion mantle plume, and the Parana in Brazil and Etendeka in Namibia were erupted 130 million years ago over the Tristan da

Ellam et al. (1992). Cunha mantle plume when Africa and South America were rifting. These observations have given rise to the idea that continental flood basalts are produced when new mantle plumes arrive at the surface. Fluid dynamic experiments and simulations show that new plumes will have large bulbous heads. When the heads arrive in the upper mantle, they melt, producing a pulse of volcanism. Others, however, have argued on geochemical grounds that continental flood basalts are produced by melting of the continental lithosphere. Because mantle plumes and continental lithosphere have such different Os and Nd isotope signatures, Os-Nd systematics provide a means of discriminating between these possibilities. Because of the dif-

ficulties in determining Os isotope ratios in basalts, only one such study has been carried out thus far. In it, Ellam et al. (1992) found that the Karroo flood basalts, erupted in South Africa 190 million years ago, have Os and Nd isotope compositions that lie on mixing lines connecting mantle plume compositions and continental lithosphere compositions (Figure 11.35). Thus at least in this case, then both a mantle plume and continental lithosphere appear to have contibuted to the magmas. The data also demonstrate the assimilation of continental crust cannot explain the low ε_{Nd} observed in these basalts.

References and Suggestions for Further Reading

Ahrens, T. J. and R. Jeanloz. 1987. Pyrite: shock compression, isentropic release, and composistion of the Earth's core. *J. Geophys. Res.* 92: 10363-10375.

Albarède, F. 1995. Introduction to Geochemical Modeling. Cambridge: Cambridge Univ. Press.

- Allègre C. J., Dupré, B., and Lewin, E., 1986. The thorium/uranium ratio of the Earth, *Chem. Geol.*, 56: 219-227.
- Allègre, C. J., J. P. Poirier, E. Humler and A. W. Hofmann. 1995. The chemical composition of the Earth. *Earth Planet Sci. Lett.* 134: 515-526.
- Allègre, C. J., O. Brévart, B. Dupré, and J. F. Minster, 1980. Isotopic and chemical effects produced in a continuously differentiating convecting Earth mantle, *Phil Trans. R. Soc. Lond.* A 297, 447-477.

Anderson, D. L. 1989. Theory of the Earth. Boston: Blackwell Scientific Publishers.

Armstrong, R. L., 1968. A model for the evolution of strontium and lead isotopes in a dynamic Earth, *Rev. Geophys.*, 6, 175-199.

Geochemistry

CHAPTER 11: THE MANTLE AND CORE

- Brooks, C., D. E. James and S. R. Hart, Ancient lithosphere: its role in young continental volcanism., *Science*, 193, 1086-1094, 1976.
- Carlson, R. W. and A. J. Irving. 1994. Depletion and enrichment history of subcontinental lithospheric mantle; an. *Earth and Planetary Science Letters*. 126: 457-472.
- Castillo, P. 1989. The Dupal anomaly as a trace of the upwelling lower mantle. Nature. 336: 667-670.
- Chase, C. G. and P. J. Patchett, 1988. Stored mafic/ultramafic crust and early Archean mantle depletion, *Earth Planet. Sci. Lett.*, 91, 66-73.
- Chen, C.-Y. and F. A. Frey. 1985. Trace element and isotopic geochemistry of lavas from Haleakala Volcano, east Maui, Hawaii: implications for the origin of Hawaiian basalts. *Jour. Geophys. Res.* 90: 8743-8768.
- Chopelas, A., R. Beohler and K. T. 1994. Thermodynamics and behavior of γ-Mg₂SiO₄ at high pressure: implications for Mg₂SiO₄ phase equilibrium. *Phys. Chem. Mineral.* 21: 351-359.
- DePaolo, D. J., 1980. Crustal growth and mantle evolution: inferences from models of element transport and Nd and Sr isotopes, *Geochim Cosmochim Acta*, 44, 1185-1196.
- DePaolo, D., 1983. The mean life of continents: estimates of continent recycling rates from Nd and Hf isotopic data and implications for mantle structure, *Geophys. Res. Lett.*, 10: 705-708.
- Dickin, A. 1995. Radiogenic Isotope Geochemistry. Cambridge: Cambridge University Press.
- Ellam, R. M., R. W. Carlson and S. B. Shirey. 1992. Evidence from Re-Os isotopes for plume-lithosphere mixing in Karoo flood basalt genesis. *Nature*. 359: 718-721.
- Farley, K. A., J. H. Natland and H. Craig. 1992. Binary mixing of enriched and undegassed (primitive ?) mantle components (He, Sr, Nd, Pb) in Samoan lavas. *Earth Planet. Sci. Lett.* 111: 183-199.
- Galer, S. J. G. and R. K. O'Nions, 1985. Residence time of thorium, uranium and lead in the mantle with implications for mantle convection., *Nature*, 316, 778-782.
- Hart, S. R. 1984. The DUPAL anomaly: A large scale isotopic mantle anomaly in the Southern Hemisphere. *Nature*. 309: 753-757.
- Hart, S. R. and A. Zindler, 1986. In search of a bulk-earth composition. Chem. Geol., 57: 247-267.
- Hart, S. R., E. H. Hauri, L. A. Oschmann and J. A. Whitehead, 1992. Mantle Plumes and entrainment: isotopic evidence, *Science*, 256, 517-520.
- Hawkesworth, C. J., P. D. Kempton, N. W. Rogers, R. M. Ellam and P. W. van Calsteren. 1990. Continental mantle lithosphere, and shallow level enrichment processes in the Earth's mantle. *Earth Planet Sci. Lett.* 96: 256-268.
- Hofmann, A. W. and W. M. White, 1982. Mantle Plumes from ancient oceanic crust, *Earth Planet. Sci. Lett.*, 57, 421-436.
- Hofmann, A. W. and W. M. White. 1983. Ba, Rb, and Cs in the Earth's mantle. *Z. Naturforsch.* 38: 256-266.
- Hofmann, A. W., Chemical differentiation of the Earth: the relationship between mantle, continental crust, and oceanic crust, *Earth Plan Sci Lett*, 90, 297-314, 1988.
- Jahn, Bor-Ming, G. Gerau, and A. Y. Glikson, 1982. Komatiites of the Onverwacht Group, S. Africa: REE geochemistry, Sm/Nd age and mantle evolution. *Contrib. Mineral. Petrol.* 80: 25-40.
- Jones, J. H., and M. J. Drake, 1986. Geochemical constraints on core formation in the Earth, *Nature*, 322: 221-228.
- Kato, T., A. E. Ringwood and T. Irifune, 1988. Constraints on element partition coefficients between MgSiO₃ perovskite and liquid determined by direct measurements, *Earth Plan. Sci. Lett.*, 90, 65-68.
- Kesson, S. E., J. D. Fitz Gerald and J. M. G. Shelley. 1994. Mineral chemistry and density of subducted basaltic crust at lower mantle pressures. *Nature*. 372: 767-769.
- McDonough, W. F. and S.-S. Sun. 1995. The composition of the Earth. Chem. Geol. 120: 223-253.
- McKenzie, D. P. and R. K. O'Nions, 1983. Mantle reservoirs and ocean island basalts, *Nature*, 301, 229-231.
- Menzies, M. A. and C. J. Hawkesworth (ed.), 1987, Mantle Metasomatism, London: Academic Press.

Geochemistry

CHAPTER 11: THE MANILE AND CORE

- Newsom, H., W. M. White, K. P. Jochum, and A. W. Hofmann, 1986. Siderophile and chalcophile element abundances in oceanic basalts, Pb isotope evolution and growth of the Earth's core, *Earth. Planet. Sci. Lett.*, 80: 299-313.
- Palme, H. and K. G. Nickel, 1985. Ca/Al ratio and composition of the Earth's upper mantle, *Geochim. Cosmochim. Acta*, 49, 2123-2132.
- Patchett, P. J. and C. Chauvel, 1984. The mean life of continents is currently not constrained by Nd and Hf isotopes, *Geophys. Res. Lett.*, 11, 151-153.
- Ringwood, A. E. and W. Hibberson. 1990. The system Fe—FeO revisited. *Phys. Chem. Mineral.* 17: 313-319.
- Ringwood, A. E., 1989. Significance of the Terrestrial Mg/Si ratio, Earth Planet. Sci. Lett., 95, 1-7.
- Ringwood, A. E., 1991. Phase transformation and their bearing on the constitution and dynamics of the mantle, *Geochim. Cosmochim. Acta*, 55, 2083-2110.
- Scarfe, C. M. and E. Takahashi, 1986. Melting of garnet peridotite to 13 GPa and the early history of the upper mantle, *Nature*, 322, 354-356.
- Schilling, J.-G., 1973. Iceland mantle plume: geochemical study of the Reykjanes Ridge, *Nature*, 242, 565-571.
- Sims, K. W., H. E. Newsom and E. S. Gladney, 1990. Chemical Fractionation during formation of the Earth's core and continental crust: clues from As. Sb, W and Mo, in *Origin of the Earth,* edited by H. E. Newsom and J. H. Jones, p. 151-174, Oxford Univ. Press, Oxford.
- Sleep, N. H. 1990. Hotspots and Mantle Plumes: some phenomenology, J. Geophys. Res., 95, 6715-6736.
- Smith, A. D. and J. N. Ludden, 1989. Nd isotopic evolution of the Precambrian mantle, *Earth Planet. Sci. Lett.*, 93, 14-22.
- Sun, S. S., 1980. Lead isotope study of young volcanic rocks from mid-ocean ridges, oceanic islands and island arcs. *Phil. Trans. R. Soc. Lond. Ser. A.*, 297: 409-445.
- Tonks, W. B. and H. J. Melosh, 1990. The physics of crystal settling and suspension in a turbulent magma ocean, in *Origin of the Earth*, edited by H. E. Newsom and J. H. Jones, p. 151-174, Oxford Univ. Press, Oxford.
- Weaver, B. L., 1991. Trace element evidence for the origin of oceanic basalts, Geology, 19, 123-126.
- White, W. M., 1985. Sources of oceanic basalts: radiogenic isotope evidence, Geology, 13: 115-118.
- White, W. M., 1993. ²³⁸U/²⁰⁴Pb in MORB and open system evolution of the depleted mantle, *Earth Planet. Sci. Lett.*, 115, 211-226.
- White, W. M. and R. A. Duncan. 1996. Geochemistry and geochronology of the Society Islands: new evidence for deep mantle recycling. In *Earth Processes: Reading the Isotope Code, Geophysical Monograph Vol.* 95, S. R. Hart and A. Basu. ed., pp. 183-206. Washington: AGU.
- White, W. M., J.-G. Schilling and S. R. Hart, 1976. Evidence for the Azores mantle plume from strontium isotope geochemistry of the Central North Atlantic, *Nature*, 263, 659-663.
- Zindler, A. and S. R. Hart, 1986. Chemical Geodynamics, Ann. Rev. Earth Planet. Sci., 14: 493-571.

Problems

- 1. Assuming the Earth has a CI chondritic composition, and using the values for the mass of the core and the mass of the mantle in Table 11.1 and the partition coefficients in Table 11.4, calculate what the concentrations of Re, Ir, Mo, and Ag should be in the bulk silicate Earth if the core formed by an equilibrium process. Compare you results with the primitve mantle values in Table 11.3.
- 2. Calculate new bulk silicate Earth concentrations of Re, Ir, Mo, and Ag by adding 1% CI chondritic material to your results from Problem 1. Again compare these results with primitive mantle values of Table 11.3.
- 3. Assume that the bulk silicate Earth has a 87 Sr / 86 Sr ratio of 0.705, and Sr concentration of 20 ppm. Calculate the bulk silicate Earth Rb/Sr ratio.

Geochemistry

CHAPTER 11: THE /MANTLE AND CORE

- 4. Suppose a rising mantle plume is 220° hotter than surrounding mantle and that the γ -olivine–Mgperovskite transition occurs at 600 km depth in the surrounding mantle. Assume a homgeneous mantle density of 3.5 g/cc at 660 km and above. At what depth does the γ -olivine–Mg-perovskite transition occur within the mantle plume?
- 5. Whether the γ -olivine–Mg-perovskite transition is endothermic or exothermic will affect mantle convection (i.e., sinking of lithospheric slabs, rise of mantle plumes). Discuss why this is so, explaining the effects of both endothermic and exothermic phase transitions on rising mantle plumes and sinking lithospheric slabs.

Geochemistry

W. M White CHAPTER 10: THE /MANTLE AND CORE

CHAPTER 12: THE CRUST

CHAPTER 12: GEOCHEMISTRY OF THE SOLID EARTH II: THE CRUST

INTRODUCTION

We now turn our attention to the crust. Though the crust forms only a small fraction of the mass of the Earth (about 0.5%), it is arguably the most varied and interesting fraction. Further, it's the fraction we can examine directly and therefore know most about. The crust has formed through igneous processes from the mantle over geologic time. There are two fundamental kinds of crust: oceanic and continental. Oceanic crust, created by magmatism at mid-ocean ridges, is basaltic in composition, thin, ephemeral, and relatively uniform. It is important in a number of respects, however. First, its composition tells us much about the composition of the mantle from which it is derived. Second, oceanic crust may be, at least some times, the raw material from which continental crust is formed. Thus we begin by briefly examining the composition of oceanic crust and the factors that control this composition. We then turn our attention to the continental crust, which is much thicker, essentially permanent, and on average andesitic in composition. The continental crust is also much more varied in com-



Figure 12.1. Schematic cross-section of the oceanic crust. Numbers on the left indicate the designation of seismically identifiable layers.

position. Although it too has formed by magmatism, its evolution is far more complex than that of oceanic crust. Though we have an excellent understanding of how oceanic crust forms, our understanding of the processes that have led to the present continental crust is far from complete. Subduction-related, or 'island arc' volcanism appears to play a particularly important role in the formation of the continental crust, so we will pay special attention to processes in islands arcs. We will then consider the problem of interaction of mantle-derived magma with the crust, then the problems of the composition and evolution and the continental crust, and finally differentiation of the crust through melting and metamorphism.

THE OCEANIC CRUST

The crust beneath the oceans differs from the continental crust in a number of important respects. First, it is thinner, with a typical thickness of 6 km or, compared to an average thickness of 35 km or so for continental crust. Second, it is more mafic, i.e., richer in Mg and Fe and poorer in Si that the continental crust. Third, it is temporary, on average, the time between its creation at mid-ocean ridges and destruction at subduction zones is 100 million years or less, compared to an average age of about 2 billion years for the continental crust. Finally, it is essentially monogenetic; the vast majority of oceanic crust is created at mid-ocean ridges. As a result, it is much more uniform in composition than the continental crust. In this section, we will focus entirely on the igne-

Geochemistry

CHAPTER 12: THE CRUST

TABLE 12.1: COMPOSITION OF MORB										
EPR MORB ¹ Primitive MORB ²										
A	Average	Std Dev.								
SiO_2	50.39	1.89	49.10							
TiO_2	1.72	0.47	0.6							
$Al_2 \tilde{O}_3$	14.93	1.13	16.4							
ΣFeO	10.20	1.52	8.8							
MnO	0.18	0.04								
MgO	7.34	1.30	10.3							
CaO	11.29	1.38	12.4							
Na_2O	2.86	0.46	1.9							
K ₂ Õ	0.25	0.47	0.1							
P_2O_5	0.35	0.48								
	99.52		99.60							

¹Average of 1266 analyses of basalts from the East Pacific Rise compiled by C. Langmuir. ²A primitive MORB composition from *Basaltic Volcanism on the Terrestrial Planets*

 Table 12.2.
 Concentrations of

 Trace Elements in Average MORB

Element	ppm	Element	ppm
Κ	883.7	Nd	11.18
Sc	41.37	Sm	3.752
Со	17.07	Eu	1.335
Ni	149.5	Gd	5.077
Cu	74.4	Tb	0.885
Rb	1.262	Dy	6.304
Sr	113.2	Но	1.342
Y	35.82	Er	4.143
Zr	104.2	Tm	0.621
Nb	3.507	Yb	3.90
Sn	1.382	Lu	0.589
Cs	0.0141	Ηf	2.974
Ва	13.87	Та	0.192
La	3.895	Pb	0.489
Ce	12.00	Th	0.1871
Pr	2.074	U	0.0711

From Hofmann (1988).

ous part of the oceanic crust. The geochemistry of ocean sediments and hydrothermal interaction between seawater and oceanic crust are discussed in a Chapter 15.

In a 1962 paper that he called "an essay in geopoetry," Harry Hess summarized his radical views on seafloor spreading. He speculated that mid-ocean ridges were produced by rising mantle convention currents, and that these convection currents then moved laterally away from the midocean ridges, producing the phenomenon of continental drift. This concept now forms the basis of plate tectonics, the fundamental paradigm of geology. Hess did miss one detail, however. He thought the oceanic crust was hydrated mantle, consisting of "serpentinized peridotite, hydrated by release of water from the mantle over the rising limb of a [convection] current." However, when mantle decompresses as it rises, it does not merely dehydrate, it melts. This melting generates the basalts that form the oceanic crust. In some respects, though, Hess's mistake is a very minor indeed. Oceanic crust is very ephemeral, and for this reason, it is sometimes better to think of it as part of the mantle reservoir than the crustal one. Nevertheless, igneous processes at mid-ocean ridges have fascinated many geochemists and much has been learned about them in the past several decades.

Seismic studies show that the oceanic crust has a layered structure (Figure 12.1). The uppermost layer, which is not present at mid-ocean ridges, consists of sediments (Seismic Layer 1). Beneath this lies Layer 2, composed of basaltic lava flows and the dikes that fed their eruption (the "sheeted dike complex"), and finally gabbros (Layer 3). The gabbros apparently consist both of basaltic magmas that crystallized in place (isotropic gabbros) and accumulations of minerals that crystallized as the basaltic magma was held in crustal magma chambers (layered gabbros). Because of the latter, the gabbros are probably somewhat more mafic on average than are the basalts. Layer 2 is often divided into Layer 2A and Layer 2B, with the latter having slightly higher seismic velocities. For many years it was thought that

boundary between the two was the boundary between the lava flows and the dikes. Based on the results the Ocean Drilling Program, however, it appears the seismic boundary reflects instead a change in porosity due to filling of voids and fractures by secondary minerals in Layer 2B. It is thus a metamorphic boundary, with the transition to the sheeted dike complex actually occurring within Layer 2B.

The average and standard deviations of the major oxides in basalts from the East Pacific Rise are listed in Table 12.1. An average of trace element concentration in mid-ocean ridge basalt (MORB) is listed in Table 12.2. The incompatible trace element abundances and isotope ratios are controlled

Geochemistry

CHAPTER 12: THE CRUST



Figure 12.2. Mg# in basalts from mid-ocean ridges of differing spreading rates. The Mg# is used as an index of the extent of fractional crystallization. Because olivine is the primary crystallizing phase, and because the Mg[#] of olivine is much higher than that of the liquid, Mg# will decrease during fractional crystallization. The figure shows that basalts from slow spreading ridges experience somewhat less fractional crystallization on average than basalts from fast spreading ridges. The most fractionated basalts occur on intermediate spreading rate ridges where small discontinuous melt lens can occur. After Sinton and Detrick (1992).

mainly by mantle chemistry, which we discussed in the preceding chapter. Hence here we will focus mainly on the major element composition of MORB, which is controlled mainly by igneous processes.

There was a considerable debate in the late 1960's and throughout the 1970's whether midocean ridge basalts were 'primary' (or close to primary), i.e., whether they were direct mantle melts or whether they had experienced extensive fractionation crystallization before eruption. In the 1980's, this debate was resolved in favor of the view that most MORB had experienced extensive fractional crystallization. A critical observation is that all primary mantle-derived magmas must be in equilibrium with olivine of the composition found in the mantle. Mantle olivine is typically 90-92% forsterite and according to experimental studies such as that of Roedder and Emslie (1972) (see Chapter 4), a melt would have to have an atomic Mg/(Mg+Fe²⁺) ratio^{\dagger} of around 0.72 to be in equilibrium with such olivine. The average compo-



Figure 12.3. Na₂O vs. MgO in MORB from three areas of the mid-ocean ridge system: the Australian-Antarctic Discordance, the Tamayo Fracture Zone area of the East Pacific Rise, and the Kolbeinsey Ridge, just north of Iceland. Thick arrows show calculated fractional crystallization paths ('liquid line of descent') for each data set. Na_{8.0} values are the intersection of the fractional crystallization path with MgO concentration of 8%. This is illustrated for the Tamayo data (dashed lines), which has Na_{8.0} of 2.31. After Klein and Langmuir (1987).

⁺ This ratio, multiplied by 100, is called the 'Mg number', generally written Mg#.

Geochemistry

CHAPTER 12: THE CRUST

sition listed in Table 12.1 has an Mg# of 59, and therefore could not be in equilibrium with mantle olivine. From this we can conclude that the oceanic crust, including the gabbroic section, is probably somewhat more mafic than the average MORB listed here. The primitive MORB composition in Table 12.1 has an Mg# of 70 and is close to a possible 'primary' composition.

One of the remarkable features of basalts erupted at mid-ocean ridges is their uniform chemistry. Nevertheless, the small variations in chemistry that do occur are significant in that they reveal the details of the magmatic processes at mid-ocean ridges. As the discussion above suggests, fractional crystallization is an important control on MORB composition. Olivine, augitic clinopyroxene, calcic plagioclase, and spinel, are, with rare exceptions, the only minerals to crystallize from MORB before eruption. Spreading rate, which is closely related to magma supply rates, seems to be an important factor in the extent of fractional crystallization: basalts erupted on fast spreading ridges are generally more fractionated than those erupted on slow spreading ridges (Figure 12.2). On fast spreading ridges, such as the East Pacific Rise, magma supplies rates are generally sufficient to maintain a small steady-state magma chamber, perhaps a few hundred meters deep and 1-2 km wide, at a depth

of a few kilometers beneath the ridge axis (Sinton and Detrick, 1992). Melt trapped in this chamber will cool and crystallize. New magma rising from the mantle will mix with the fractionated magma in the chamber before eruption. In contrast, magma supply rates on slow spreading ridges, such as the Mid-Atlantic Ridge, are too low to maintain a steady-state magma chamber. As new magma injected into the crust can erupt without mixing with older, more fractionated magma.

The other factors that control the major element composition of MORB are the degree and depth of melting. The effects of these factors were investigated by Klein and Langmuir (1987). The difficulty with understanding the melting process is that all MORB have suffered some fractional crystallization, and this tends to obscure the melting effects. Klein and Langmuir found that MORB major oxide data from each locality on the mid-ocean ridge system formed a coherent and distinct array when oxide abundances or ratios of oxides were plotted against MgO (Figure 12.3). Calculated frac-



Figure 12.4. Relationship between $Na_{8.0}$ and $Fe_{8.0}$ in MORB and axial depth of mid-ocean ridges. Data from regions near hot spots are circled. Data from ridge segments near the Galápagos, Azores, and Jan Mayen hot spots, which are enclosed in the dashed field, appear to have anomalously low Fe. From Klein and Langmuir (1987).

Geochemistry

CHAPTER 12: THE CRUST

tional crystallization paths, illustrated in Figure 12.3, passed through these arrays, so they concluded that the trends observed in these plots reflected fractional crystallization. They then argued that differences between the trends reflected differences in the melting process.

Rather than attempt to solve the difficult problem of determining the composition of parental magmas, Klein and Langmuir simply corrected regionally averaged data to a common MgO concentration of 8%. They did so simply by projecting each regional array on an oxide versus MgO plot to 8% MgO. For example, a line drawn through Na₂O data from the Tamayo Fracture Zone region of the East Pacific Rise intersects 8% MgO at 2.31% Na_2O (Figure 12.3). They called this the $Na_{8,0}$ value; thus Tamayo has an $Na_{8,0}$ of 2.31. So instead of comparing primary magma compositions, they compared compositions that had experienced similar extents for fractional crystallization.

Klein and Langmuir found that the regional corrected averages correlated strongly with the depth of the ridge axis (Figure 12.4). Comparing the variations in concentrations they observed with those in experimentally produced partial melts, they found that both variations in the mean pressure of melting and in the mean extent of melting were required to explain the data. For example, sodium concentrations in partial melts appears to be controlled only by the extent of partial melting. Iron concentrations, on the other hand, are only weakly affected



Figure 12.5. Variation of MgO and FeO in partial melts of mantle peridotite. Grayed fields show the compositions of experimental produced partial melts of peridotite at 3 different pressures. Hashed fields show compositions of high MgO basalts from the AAD, Tamayo, and Kolbeinsey regions. Curves for the calculated compositions at maximum and minimum melting are also shown. Dashed arrows show the path of melt composition produced by melting of adiabatically rising mantle. Curved arrow shows how the compositional a 15 kbar melt will evolve due to fractional crystallization of olivine. After Klein and Langmuir (1987).

by degree of partial melting, but are strong functions of the pressure at which melting occurs (Figure 12.5).

Klein and Langmuir concluded mantle temperature was probably the key factor in accounting for both depth of the ridge axis and the composition of melts erupted, because mantle temperature affects both degree of melting and the mean depth of melting. Shallow segments of the mid-ocean ridge system overlie relatively hot mantle. The hot mantle intersects the solidus at greater depth and ultimately melts to a greater degree (Figure 12.6). Hotter mantle is less dense and therefore more buoyant, so that ridges overlying hotter mantle will be more elevated. Cooler mantle will not begin to melt until it reaches shallower depth, and total extent of melting will be more limited. Klein and Langmuir concluded that a range in degree of melting of 8-20% and in mean pressure of melting of 0.5 to 1.6 GPa were required to produced the range in compositions observed. The hottest regions of the mantle occur near mantle plumes such Iceland. The coolest region occurs at the Australian-Antarctic Discordance, a region where the ridge is particularly deep and isotope studies have suggested is a boundary between mantle convection cells. Overall, the data suggested a range in mantle temperature of some 250° C.

Geochemistry

CHAPTER 12: THE CRUST





Figure 12.6. a.) Cartoon of the pressure-temperature relationship of adiabatically rising mantle undergoing melting. Hotter mantle (X) intersects the solidus at higher pressure and ultimately melts to a higher degree than cooler mantle The break in slope occurs (Y). because energy is consumed by melting (enthalpy of fusion). b.) Cartoon illustrating the relationship between axial depth, crustal thickness, melting, and mantle temperature. Hotter mantle (X) maintains the ridge at higher elevation because of its buoyancy. It also has a deeper melt column and melts to a greater degree, producing thicker crust than cooler mantle (Y). After Klein and Langmuir (1989).

Processes in Subduction Zones and Island Arc Geochemistry

Though the evolution of the continental crust has undoubtedly been complex and involved a number of processes, there is virtual unanimous agreement among scientists that magmatism is primary way in which the continental crust has formed. Today, and throughout the Phanerozoic, additions to the continental crust occur primarily by magmatism associated with subduction. As we shall see, there are also good chemical arguments that subduction zone volcanism has been the most important, though not necessarily exclusive, mechanism by which the continents have formed. Volcanos of the Andes Mountains of South America, which overlie the subducting Nazca Plate, are perhaps the best example of this process. Not all subduction zones are located along continental margins; indeed, most are not. The Marianas are a good example of an intra-oceanic subduction zone. Some old island arcs, however, such as Japan, parts of Indonesia, and the Aleutians have crustal seismic structures that are intermediate between those of continental and oceanic crust. This suggests that island arcs may eventually transform to continental crust. Plate motion may eventually result in these arcs accreting to continents. Thus intra-oceanic arcs may also contribute to continental growth. Thus in attempting to understand how the continental crust has formed, it is well worth while to consider subduction zone processes.

Subduction zones are, of course, the place where oceanic crust and lithosphere (often referred to as "the slab") are returned to the mantle. However, oceanic crust is not chemically identical to the basalt produced at mid-ocean ridges. Two important things have happened to the crust between its creation and subduction. First, it has reacted with seawater at a variety of temperatures. This process, which we discuss in detail in Chapter 15, hydrates the oceanic crust, adds some elements to it from seawater (e.g., Mg, U), and extracts others. Isotopic exchange with seawater affects the isotopic composition of Li, B, and Sr. The second thing that happens is that the oceanic crust acquires a sedimentary veneer. Elements gained through hydrothermal alteration and the elements in sediments are derived almost entirely from the continental crust. Thus the subducting oceanic crust carries with it a certain amount of continental crustal material (particularly when sediment is subducted) and subduction zones are therefore the principle sites of crust—to—mantle mass transfer.

MAJOR ELEMENT COMPOSITION

Magmas found in island arcs (we'll use the term *island arc* for all subduction zone magmatism, including continental margin type) appear to be predominantly andesitic. It seems unlikely that andesite is the principle magma produced in arcs,

Geochemistry

W. M. White

CHAPTER 12: THE CRUST

however. Generally, we don't see the lower parts of arc volcanic edifices, which may be basaltic. There is also considerable doubt whether an andesite can be produced by partial melting of the mantle, particularly at depth. Most arcs sit about 100 km above the Benioff zone, and magmas may be generated close to this depth. A safer bet is that the primary magma is actually basaltic, of which andesites are fractional crystallization products. In any case, basalt is not uncommon in intraoceanic arcs.

In major element composition, island-arc volcanics (IAV) are not much different from other volcanic rocks. Compared to MORB, the principal difference is perhaps simply that siliceous compositions are much more common among the island-arc volcanics. Most IAV silica saturated or oversaturated; silica undersaturated magmas (alkali basalts) are rare. In that sense, we might call them tholeiitic. However, in the context of island arc magmas, the term tholeiite has a more restrictive meaning. Two principal magma series are recognized, one called tholeiitic, the other called calc-alkaline. The principle difference is that the tholeiites differentiate initially toward higher Fe and higher Fe/Mg than the calc-al-This is illustrated in Figure kaline lavas.

12.7. In addition, the tholeiites tend to be poorer in K and some other incompatible elements than the calc-alkaline rocks. Kay et al. (1983) argued that the difference, at least for the Aleutians, relates to tectonic environment and depth of magma stagnation and crystallization. Tholeiites occur in extensional environments in the arc where magma can ascend relatively rapidly and cool less. Fractional crystallization occurs at shallow levels. Calc-alkaline lavas ascend more slowly and undergo crystallization at greater depth.

A third series is sometimes defined, the high-alumina series. IAV in general tend to have slightly higher Al contents on average than MORB or OIB, but there is very considerable overlap with the OIB and MORB fields.

IAV also tend to be somewhat poorer in Ti than MORB and OIB, though this is not necessarily a primary feature. Perfit et al. (1980) argued that the difference in Ti content is due to early crystallization of Fe-Ti oxides (e.g., magnetite and ilmenite) in IAV, which buffers the Ti concentration. Thus the higher Ti concentrations in MORB may reflect the way the magmas differentiate rather than higher Ti concentrations of MORB primary melts or MORB sources. The same is true for iron enrichment. Fe-Ti oxide precipitation may in turn depend



Figure 12.7 AFM ($A=K_2O+Na_2O$, F=FeO+MnO, M=MgO) diagram illustrating the difference between tholeiitic and calc-alkaline lava series of island arcs. Calc-alkaline rocks plot below the heavy line, tholeiites above. PRS is Kuno's pigeonite rock series; HRS is Kuno's hypersthene rock series. I is the Thingmuli series of Iceland. After Carmichael, et al. (1973).



Figure 12.8 Correlation of $Ca_{6.0}$ and $Na_{6.0}$ with crustal thickness in island arc basalts. $Ca_{6.0}$ and $Na_{6.0}$ are the CaO and Na_2O concentrations after correction for fractional crystallization to 6.0 MgO. From Plank and Langmuir (1988).

Geochemistry

CHAPTER 12: THE CRUST

on oxygen fugacity. Oxygen fugacity is higher in IAV because of the higher water contents. IAV are enriched in water as well as other volatiles compared to MORB and at least some OIB (e.g., Ha-waii). Island arc volcanics may also have higher CO_2/H_2O ratios than MORB (Perfit et al., 1980).

Plank and Langmuir (1988) investigated the factors that control the major element composition of island arc basalts. They treated the data in a manner analogous to Klein and Langmuir (1987), correcting regional data sets to a common MgO content, but they used 6% MgO rather than the 8% used by Klein and Langmuir. They found that $Na_{6.0}$ and $Ca_{6.0}$ (i.e., Na_2O and CaO concentrations corrected to 6% MgO) correlated well with crustal thickness (Figure 12.8). They argued that crustal thickness determines the height of the mantle column available for melting. Most island arc volcanoes are located above the point where the subducting lithosphere reaches a depth of 100-120 km. This suggests that melting begins at a relatively constant depth in all island arcs. If this is so, then over which mantle can rise and undergo decompressional melting will be less if the arc crust is thick, leading to smaller extents of melting beneath arcs with thick crust, and higher $Na_{6.0}$ and $Ca_{6.0}$ in the parental magmas.

TRACE ELEMENT COMPOSITION

The differences in trace elements between island arc volcanics and those from other tectonic environments are probably more significant than the differences in major elements. Rare earths, however, are not particularly distinctive. There is a very considerable range in rare earth patterns: from LRE depleted to LRE enriched (Figure 12.9). IAV are virtually never as LRE depleted as MORB, but absolute REE concentrations are, however, often low, and it is not unusual for the middle and heavy rare earths to be present at lower concentrations than in MORB. One other aspect is of interest. Ce anomalies occur in some IAV, whereas they are never seen in MORB or OIB, though they have been observed in continental carbonatites and kimberlites. The significance of the Ce anomalies remains uncertain.



Figure 12.9. Rare earth patterns of some typical island arc volcanics. From White and Patchett (1984).

Geochemistry

CHAPTER 12: THE CRUST



Figure 12.10. Relative alkali-alkaline earth enrichment of IAV illustrated by plotting the $(Ba/La)_N$ ratio vs. the $(La/Sm)_N$ ratio. The subscript N denotes normalization to chondritic values. From Perfit et al. (1980).

To summarize, IAV MORB. may be either light-rareearth enriched or light-rareearth depleted. They generally are enriched relative to MORB in the alkalis and alkaline earths. Significantly, these are the elements most enriched in sediments, and the most soluble elements. Relative to MORB, IAV tend to be depleted in Nb, and Ta, and sometimes in Ti, Hf, and Zr. These are the elements least enriched in sediments, and are generally highly insoluble in aqueous solutions because of their high ratios of ionic charge to radius (Z/r).

Ce anomalies occur in sediment, so there is the immediate suspicion that the anomalies in IAV could be inherited from subducted sediment.

Island-arc volcanics are richer in the incompatible alkalis and alkaline earths (K, Rb, Cs, Sr, and Ba) relative to other incompatible elements when compared with MORB or OIB. This is illustrated in Figure 12.10, using the Ba/La ratio. Though both IAV and oceanic basalts can have a large range in rare earth patterns, as illustrated by the range in La/Sm ratios, the Ba/La ratios of IAV are generally higher.

Island-arc volcanics tend to be poor in the so-called high-field-strength elements (HFS), i.e., those elements with charge of +4 or +5 (Chapter 7). While Nb and Ta are almost invariably depleted in arcs, Zr, Hf, and Th are not. It would be better to speak of Nb and Ta depletion. This is illustrated in Figure 12.11.

Figure 12.12 is a spider diagram summarizing the trace element differences between island arc volcanics and



Figure 12.11. Low Nb concentrations of IAV compared to MORB and OIB illustrated by plotting Nb vs. Ba. From Perfit et al. (1980).

Geochemistry

CHAPTER 12: THE CRUST



Figure 12.12. Spider diagram comparing typical incompatible element contents of island arc volcanics and MORB. Note the relative enrichment in alkalis and alkaline earths and the depletion in Ta in the island arc volcanics. After Sun (1980).

ISOTOPIC COMPOSITION AND SEDIMENT SUBDUCTION

Island arcs overlie subduction zones, which raises the obvious question of the degree to which subducting oceanic crust and sediment might contribute to island arc magmas. These questions have been most successfully addressed through isotope geochemistry.

Sr isotope ratios are generally higher, and Nd isotope ratios generally lower than in MORB, with ⁸⁷Sr/⁸⁶Sr ratios around 0.7033 and ε_{Nd} of +8 being fairly typical of intraoceanic IAV (Hawkesworth et al., 1991). Though there is considerable overlap with oceanic basalts (MORB and



Figure 12.13. Pb isotope ratios in island arc volcanics. Fields for the South Sandwich, Lesser Antilles, Aleutians, Marianas, Philippines, Taiwan, Banda and Sunda arcs are shown and compared with fields for Atlantic and Pacific MORB (field labeled MORB) and Indian Ocean MORB (IMORB), and modern marine sediments.

Geochemistry

W. M. White

CHAPTER **12**: THE CRUST

OIB), there is a slight tendency for IAV to have higher Sr isotope ratios for a given Nd isotope ratio and hence plot to the right of the oceanic basalt array on a Nd–Sr isotope ratio plot. This shift to higher Sr isotope ratios appears to result from a contribution of subducted oceanic crust to IAV magma sources. This conclusion is based on the observation that during weathering and hydrothermal alteration of the oceanic crust, isotopic exchange with seawater shifts Sr isotope ratios of the oceanic crust to higher values. Nd isotope ratios, are virtually unaffected by these processes because of the extremely low concentration of Nd in seawater.

²⁰⁶Pb/²⁰⁴Pb isotope ratios overlap values of oceanic basalts, but generally having a more restricted range. ²⁰⁷Pb/²⁰⁴Pb ratios are typically higher IAV than most oceanic basalts. They tend to form steeper arrays on ²⁰⁷Pb/²⁰⁴Pb-²⁰⁶Pb/²⁰⁴Pb plots, and overlap the field of marine sediments (Figure 12.13). For most island arcs, Pb isotope ratios in the arc volcanics lie between sediment local to the arc and the MORB field (Karig and Kay, 1980). We pointed out in the last chapter that Indian Ocean MORB have lower ²⁰⁶Pb/²⁰⁴Pb and higher ²⁰⁷Pb/²⁰⁴Pb than do MORB from the Atlantic and Pacific. This presumably reflects a difference in the isotopic composition between the



Figure 12.14. Pb isotope ratios in volcanics of the Lesser Antilles island arc and on the Atlantic Plate subducting beneath it. Contours show the $^{206}Pb/^{204}Pb$ ratios in the sediment, which increases from north to south due to the increasing contribution of Orinoco River sediment. The average $^{206}Pb/^{204}Pb$ for each island, which is shown beside the island, also increases from south to north. From White and Dupré (1986).

upper mantle beneath the Atlantic and Pacific on the one hand and the Indian on the other. Interestingly, the Pb isotope arrays for Indian and southwestern Pacific arcs (Sunda, Banda, Philippines, and Taiwan) are elongate toward the Indian Ocean MORB field rather than Atlantic and Pacific MORB. On the whole then, Pb in island arc magmas appears to be a mixture Pb from local sediment and local upper mantle.

The Lesser Antilles arc illustrates this particularly well. The Lesser Antilles arc lies on the Caribbean Plate just to the north of South America. The Orinoco River, which drains the Archean Guiana Highland, delivers sediment that contains particularly radiogenic Pb to the front of the arc. As a result, the Pb isotope ratios in sediments of the Demerara Abyssal Plain decrease systematically from south to north (Figure 12.14). A similar pattern of decreasing Pb isotope ratios can be seen in the volcanics of the arc, apparently reflecting the changing isotopic composition of the sediment being subducted.

Pb isotope ratios are particularly sensitive to the presence of subducted sediment because the concentration of Pb in sediment is so much higher, well over two orders of magnitude, than in the mantle. In a mixture of sediment and mantle, the Pb isotope ratios of the mixture are virtually identical to those of the sediment when as little as 2 to 3% sediment is present in the mixture. This is not true of Sr and Nd isotope ratios.

¹⁰Be is an even more sensitive indicator of the presence of sediment in IAV magma sources than Pb, because it is present in young sediment, but entirely absent from the mantle. As we found in Chapter 8, ¹⁰Be is a cosmogenic isotope; it is produced in the atmosphere by cosmic ray spallation of ¹⁴N. It has a half-life of only 1.5 million years, so we would not expect to find significant amounts of ¹⁰Be in the Earth's interior; any created before the solar system formed has long since decayed away. ¹⁰Be created in the atmosphere is purged by rainfall and is strongly absorbed by clays of sediment and soil.

CHAPTER 12: THE CRUST



Figure 12.15. ¹⁰Be in arc and non-arc lavas. "A" indicates sample from active volcano, "H" a historic eruption, and "F" (fresh) indicates collected during eruption. From Tera et al., 1986.

Geochemistry

Workers at Carnegie Institution of Washington turned to ¹⁰Be when they were seeking a means of determining subduction rates. They decided to attempt to determine the amount of ¹⁰Be in island arc volcanics on the assumption that sediments containing ¹⁰Be were subducted, and by comparing the amount in arc volcanics with the amount of sediment, they hoped to determine how much had decayed away, and thereby how much time had elapsed since the sediment left the surface. The problem proved to be too complex to determine subduction rates quantitatively. However, in the course of this attempt, they produced by far the best evidence that sediment is indeed subducted in island arcs. Figure 12.15 compares the amount of ¹⁰Be in arc lavas with that in non-arc lavas. ¹⁰Be has not been found in measurable quantities (10^6) atoms per gram) in non-arc lavas, but has been found in lavas of some arcs. The interpretation is that the ¹⁰Be originates from sediment subducted to the magma genesis zone. The absence of ¹⁰Be in other arc lavas does not mean sediment is not subducted in those arcs. Only very young sediment, < 10-15 Ma, will contain appreciable amounts of ¹⁰Be, so if the uppermost sediment is removed in an accretionary wedge, no ¹⁰Be will be delivered to the magma genesis zone even though deeper, older sediment is being subducted.

Figure 8.36 showed that IAV have

higher and more variable δ^{18} O than either MORB or OIB. This may also be attributed to a contribution of subducted oceanic crust and sediment. Indeed, as we have said before, oxygen isotope ratios can be changed only at or near the surface of the Earth, so this is the only way the δ^{18} O of IAV can be readily explained.

MAGMA GENESIS IN SUBDUCTION ZONES

Now that we have an overview of the composition of arc magmas, let's consider in more detail the processes that lead to the unique geochemistry of island arc magmas. Figure 12.16 summarizes petrogenesis in the subduction zone environment.

Most geochemists and petrologists believe that arc magmas are produced primiarily within the mantle wedge[‡] (e.g., Hawkesworth et al., 1991; Kay and Kay, 1994). The evidence for this is: (a) Primary arc magmas differ only slightly in their major element chemistry from oceanic basalts, which are definitely mantle-derived. The andesitic nature of many arc magmas probably results from fractionation crystallization in a crustal or subcrustal magma chamber. It is therefore most

⁺ The *mantle wedge* is the part of the mantle overlying the subducting slab; because the slab descends at an angle, this region in triangular, or wedge-shaped.

Geochemistry

CHAPTER 12: THE CRUST



Slab dehydration

Figure 12.16. Cross-section of a subduction zone illustrating island arc magma genesis.

likely they are partial melts of peridotite rather than subducted basalt or sediment. (b) Radiogenic isotopic and trace element systematics generally allow only small fraction of sediment (generally a few percent or less) to be present in arc magma sources. Relatively high ³He/⁴He ratios in arc lavas confirm this. (c) Rare earth patterns of island arc magmas are consistent with these magmas being generated by partial melting of peridotite, but not by partial melting of eclogite, which would be the stable form of subducted basalt at 100 km depth (the subducting lithosphere is typically located at about 100 to 120 km depth beneath island arc volcanos). Because the heavy rare earths partition strongly into garnet (e.g., Figure 6.15), melts of eclogite should show steep rare earth patterns, with low concentrations of the heavy rare earths. This is not generally the case. Rare high magnesium andesites, sometimes called "adakites" (after a well documented occurrence on Adak Island in the Aleutians) with steep rare earths patterns may represent exceptions to this rule and may indeed by generated by small extents of melting of subducted oceanic crust (Kay, 1978; Defant and Drummond, 1990). It is possible that such "slab melts" were more common several billion years ago. We discuss this possibility later in the chapter. Finally, we should point out that some scientists believe arc magmas are indeed generated in the subducting slab (e.g., Brophy and Marsh, 1986).

We have seen that isotope systematics, particularly Pb and ¹⁰Be, provide evidence of subducted sediment in the sources of most arcs. Subducted sediment appears to influence the trace element compositions of arc lavas as well. Plank and Langmuir (1993) carried out careful study of the composition of volcanics from 8 arcs and the sediments being subducted beneath them. By analyzing representative samples from the sediments and considering the proportions of sediment types being carried beneath the arc, they estimated the flux of elements being carried by sediment beneath the arc. They found they could relate the degree of enrichment of most incompatible elements to the sediment flux of that element. For example, the Ba/Na and Th/Na ratio (after correction for fractional crystallization to 6% MgO) correlate strongly with the Ba and Th sediment fluxes (Figure 12.17). Different arcs are enriched to different degrees in these elements: for example, the Lesser Antilles arc has moderate Th/Na ratios but low Ba/Na ratios. The difference appears to be due to the difference in the sediment flux.

The trace element geochemistry of island arc magmas cannot, however, be explained solely by partial melting of simple bulk mixtures of mantle, sediment and subducted oceanic crust. In particular, the characteristic enrichment of the alkalis and alkaline earths cannot be accounted for in this way. This is illustrated in Figure 12.18, which compares the observed enrichment of Lesser Antilles low-K basalts in incompatible elements with the enrichment predicted assuming the source was a mixture of depleted mantle and sediment. White and Dupré (1986) calculated the fraction of sediment in the mixture from Nd isotope ratios of the basalts, and assumed simple batch melting. This simple model

Geochemistry

CHAPTER 12: THE CRUST



Figure 12.17. Relationship between $Ba/Na_{6.0}$ and $Th/Na_{6.0}$ in volcanics from 8 arcs and Ba and Th sediment flux beneath those arcs. Horizontal bars represent uncertainty in the amount of sediment subducted; vertical bars reflect the variance of the ratio in the arc volcanics. G is Guatemala, M: Mexico, J: Java, T: Tonga, Mar: Marianas, Al: Aleutians, Ant: Lesser Antilles. From Plank and Langmuir (1993).

rine basalts from the Marianas. As mentioned above, water concentrations correlated with those of mobile trace elements (as well as those of several less mobile, but incompatible trace elements, such as La). However, water concentrations correlated inversely with the concentrations of moderately incompatible elements, such as Ti, Zr, and Na. These observations at first seem contradictory. To explain them, Stolper and Newman (1994) developed a model in which a mantle source consisting of a mixture of depleted mantle (i.e., a composition that would melt to produce average MORB) and an H₂O-rich component melts to varying degrees. Their calculated H₂O-rich

predicted the concentrations of the rare earths and Th reasonably well, but the enrichment of Pb, Cs, Rb, U, K, Ba, and Sr was greater than predicted. The elements that are over-enriched are those most soluble in aqueous solution, so the excess abundance of these "mobile" elements suggests they are enriched due to preferential transport in aqueous fluids. Particularly strong enrichment of these elements is characteristic of virtual all arc magmas, not just those of the Lesser Antilles. Thus it is widely believed that water released by dehydration of the subducting oceanic crust and sediment transports mobile elements from the slab to the overlying mantle wedge. This idea is supported by the positive correlation observed between concentrations of these elements and water contents of submarine basalts from the Marianas (Stolper and Newman, 1994).

A long standing and key question is why magmas are generated at all in an area where cold lithosphere is descending. Decades of experimental work has shown that water lowers the solidus of rock and leads to enhanced melting at any given temperature compared to dry conditions, hence water released by the subducting slab may induce melting in the wedge. The best evidence of this comes from work of Stolper and Newman, 1994) on subma-



Figure 12.18. Ratio of observed enrichment of incompatible elements in Lesser Antilles low-K basalts compared to enrichment calculated from a sedimentdepleted mantle mixing model based on Nd isotopic composition. Enrichment of the rare earths and Th can be accounted for by this model, but the enrichment of Pb, Cs, Rb, U, K, Ba, and Sr is greater than predicted, probably due to preferential transport of these elements in fluids. Gray area is 1 standard deviation of the basalts used in the average. From White and Dupré (1986).

Geochemistry

CHAPTER 12: THE CRUST

component was strongly enriched in incompatible and mobile elements, including Na. To explain the inverse correlation between Na and H_2O in the basalts, they found that the extent of melting must vary inversely with the amount of H_2O -rich component in the source mixture. The smallest extents of melting (about 5%) occur in H_2O -poor sources and give rise to incompatible element-poor basalts, while the highest extents (over 20%) give rise to H_2O and incompatible element-rich basalts. The large extents of melting producing that latter leads to their being poor in elements such as Na and Ti, which are only moderately incompatible. Thus Stolper and Newman's results suggest water induces melting in the mantle wedge.

The origin of the Nb-Ta depletion remains unclear, though there are a variety of ideas: (a) In part, this depletion may reflect the low abundance of these elements in sediments, but it seems unlikely this alone can account for this depletion. (b) The high charge to radius ratios of these elements makes them relatively insoluble in the aqueous solutions, and hence they may be transported to the mantle wedge less efficiently. (c) They may be retained in specific phases, such as ilmenite, stabilized by unusual conditions or composition of the slab or the web. (d) Finally, these elements may have the same abundances in the mantle wedge as in the source of MORB, but higher degrees of melting involved in the generation of IAV compared to MORB lead to lower abundances in the former. These possibilities are not mutually exclusive; several factors may contribute to Nb-Ta depletion.

Why island arc magmas are more silicic and appear to have experience more fractional crystallization than oceanic basalts is also unclear. Again, there are several possibilities: (a) Retardation of plagioclase crystallization and early onset of oxide crystallization due to high concentrations of water in IAV may drive island arc magmas to more silicic compositions than water-poor magmas such as MORB. (b) Island arc crust is generally thicker than oceanic crust, providing a greater opportunity for stagnation and fractional crystallization. This will certainly be true of continental margins such as the Andes. It would be particularly difficult for basaltic magmas to ascend through such thick, low density continental crust. (c) Subduction zones are fundamentally compressional environments, which may make it more difficult for magmas to ascend to the surface, again causing magmas to stagnate and crystallize. (d) Primary magmas may be more silicic and less mafic, though not necessarily andesitic, due to generation under "wet" conditions. Again, these possibilities are not mutually exclusive.

CRUST-MANTLE INTERACTION: ASSIMILATION

Some of the characteristics of island arc magmas that have been ascribed to sediment subduction can also be produced when mantle-derived magmas assimilate continental crust. When mantle-derived magmas encounter continental crust, the basalt heats the crust. If temperatures are high enough, the crust may melt. The temperature reached will depend on the size of the magma body: a relatively thin dike will cause partial melting only very locally; a larger body will cause partial melting at greater distances (Figure 12.19). The melt produced can be assimilated by the intruding magma. Assimilation of crust is a pervasive phenomenon and affects most, though certainly not all, magmas erupted through continental crust, both in subduction and other tectonic environments. For this reason, we need to devote special attention to the problem.

Geochemistry of Two-Component Mixtures

In the simplest case, the resulting magma becomes a two component mixture of the crustal and mantle melts. If homogenization in not complete, there will be some compositional variation. If we plot the concentration of any two elements in different samples of this mixture against each other, they must lie on a straight line between the two end members. However, if we plot ratios of elements, or isotope ratios, they need not lie on a straight line. Indeed, in the general case they do not; rather they will define a curve whose equation is:

$$Ax + Bxy + Cy + D = 0 12.1$$

where *x* and *y* are the variables of the abscissa and ordinate respectively (Langmuir et al., 1978). If end members are designated 1 and 2 and have ratios x_1 and y_1 , and x_2 and y_2 respectively, then

CHAPTER 12: THE CRUST



Figure 12.19. Temperatures as a function of time and distance surrounding basaltic dikes of various thickness. Lines B, M, and G are the temperatures of biotite breakdown, muscovite breakdown and granite minimum melting respectively. Gray area is region of partial melting. From Patchett (1980).

more common. Note that on a ${}^{207}\text{Pb}/{}^{204}\text{Pb}$ — ${}^{206}\text{Pb}/{}^{204}\text{Pb}$ plot, mixing curves will always be straight lines because the denominators are the same (i.e., $a = b = {}^{204}\text{Pb}$).

Two component mixtures will also form straight lines on isochron plots (e.g., ⁸⁷Sr/⁸⁶Sr—⁸⁷Rb/⁸⁶Sr such as Figure 7.6), because the denominators are the same. Thus mixing lines can be mistaken for isochrons and visa versa. One way to distinguish the two is a ratio-element plot. A ratio-element plot, for example ⁸⁷Sr/⁸⁶Sr vs. Sr, will also be a curve described by equation 12.1, but a ratio plotted against the inverse of the denominator, for example ⁸⁷Sr/⁸⁶Sr—1/Sr, will be a straight line. Such a plot can be a useful discriminator between isochrons and mixing lines because only in the latter case will ⁸⁷Sr/⁸⁶Sr—1/Sr necessarily define a straight line (Figure 12.21). Again, this result is completely general, and while the general prin-

Geochemistry

$$A = a_2 b_1 y_2 - a_1 b_2 y_1 \qquad 12.2$$

$$B = a_1 b_2 - a_2 b_1 12.3$$

$$C = a_2 b_2 x_1 - a_1 b_2 x_2 \qquad 12.4$$

$$D = a_1 b_2 x_2 y_2 - a_2 b_1 x_1 y_1 \quad 12.5$$

where a_i is the denominator of y_i and b_i is the denominator of x_i . The curvature of the mixing line will depend on the ratio r:

$$r = a_1 b_2 / a_2 b_1$$
 12.6

The greater the value of r is, the greater the curvature. Only in the special case were r=1 is the line straight. This is illustrated in Figure 12.20. This result is completely general and applies to mixing of river water and seawater, etc. as well as mixing of magmas.

Taking the example of ¹⁴³Nd/¹⁴⁴Nd versus ⁸⁷Sr/⁸⁶Sr, the curvature depends on the ratio of (¹⁴⁴Nd₁⁸⁶Sr₂)/(¹⁴⁴Nd₂⁸⁶Sr₁). Since to a very good approximation the amount of ¹⁴⁴Nd and ⁸⁶Sr are proportional to total Nd and Sr respectively, *r* is approximated by Nd₁Sr₂/Nd₂Sr₁. If we express this ratio as $r = (Nd/Sr)_1/(Nd/Sr)_2$ we see that the curvature depends on the ratio of the Nd/Sr ratio in the two end members. In mantle-derived rocks Sr/Nd does not deviate greatly from 10, so mixing curves will be close to straight lines. In crustal rocks and sediments, deviations from *r* = 1 are more likely and curved mixing lines therefore



Figure 12.20. Plots of ratios of elements or isotopes, Q/a versus P/b for mixing of end members 1 and 2. The numbers along the curves are the values for *r*. From Langmuir et al. (1978).

CHAPTER 12: THE CRUST



Figure 12.21. Mixing hyperbola formed by components A and B. After Faure (1986).

cipals have been illustrated with isotope ratios, they apply equally well to ratios of concentrations.

When a magma or series of magmas appear to reflect mixing, we are often faced with having to decide whether (1) two mantlederived magmas are mixing, (2) two distinct mantle sources are mixing before melting, or (3) a mantle-derived magma is mixing with assimilated crust. In case (2), plots involving an elemental concentration will not fall on mixing lines because partial melting and fractional crystallization will change element concentrations. Isotope ratios

will not be changed by magma genesis so a plot of two isotope ratios will describe a mixing line in case (2) as well as cases (1) and (3). As we pointed out in Chapter 9, stable isotopes are particularly useful in recognizing assimilation. This is so because mantle materials have comparatively uniform stable isotope ratios and crustal rocks very often have stable isotope ratios that are different from mantle values.

Recognizing crustal assimilation in subduction zones magmas can be particularly difficult because many of geochemical effects characteristic of crustal assimilation can also result from the presence of subducted sediment component in such magmas. For example, the $\delta^{18}O_{\text{SMOW}}$ of sediment is generally greater than +15 whereas the mantle value is about +5.6. Island arc magmas derived from a source having a component of subducted sediment will thus have slightly $\delta^{18}O$ higher than mantle values. However, high $\delta^{18}O$ can also result from assimilation of crust during magma ascent. Similarly, elevated values of ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ could result from either the subducted sediment in the magma source of assimilation of continental crust.

By combining O isotope analyses with radiogenic isotope analyses, it is possible to distinguish between these two processes. Virtually all silicate rocks, including sediments, magmas, and peridotite, have similar O concentrations (generally close to 50%), whereas concentrations of the radiogenic elements Sr, Nd, Hf, Pb, etc. vary widely. In particular, since these elements (except Os) are incompatible to some degree, their



Figure 12.22. O—Sr isotope plot showing the difference is mixing curves produced by contaminating magma with crust ("crustal contamination") as opposed to contaminating the magma source with subducted material ("source contamination"). x is the fraction of end member "C" (crust or subducted sediment) in the mixture. After James (1981).

Geochemistry

CHAPTER 12: THE CRUST

concentrations in magmas will be higher than in the mantle. Many island arc magmas have Sr concentrations that are greater than those typical of continental crust, but the mantle has much lower concentrations of Sr than sediment. This means mixing curves for sediment-mantle mixing will look very different from for magma-crust mixing.

In the case of sediment-mantle mixing, the value of $r = Sr_M/O_M/Sr_S/O_S$ (where the subscripts M and S denote the concentrations in mantle and sediment respectively) will be typically much less than 1. In the case of magma-crust mixing, the value of $r = Sr_M/O_M/Sr_C/O_C$ (where the subscripts M and C denote the concentrations in magma and crust respectively) will typically be 1 or greater. On a plot of δ^{18} O vs. 87 Sr/ 86 Sr, this produces a convex curve in the case of assimilation and a concave curve in the case of subducted sediment mixing with mantle (Figure 12.22).

Assimilation-Fractional Crystallization

As we pointed out in Chapter 9, assimilation will inevitably be accompanied by fractionation crystallization, because heat released during crystallization provides the energy for melting surrounding crust. Where fractional crystallization also occurs, assimilation will not produce simple



Figure 12.23. Variation of concentration during assimilation-fractional crystallization. Short dashed red lines are for simple fractional crystallization. Numbers on the curves refer to values of C_a/C_m^o , the ratio of the concentration in the assimilant to the original concentration in the magma. After DePaolo (1981a).

mixing curves. Let's consider this in more detail. The change in concentration of an element, C,

in a magma undergoing assimilation and fractional crystallization (AFC) is given by:

$$C_m/C_m^o = F^{-z} + \left(\frac{R}{R-1}\right) \frac{C_a}{zC_m^o} (1 - F^{-z})$$
 12.7

where C_m is the concentration in the magma, C_m^0 is the original concentration in the magma, F is the mass of magma remaining as a fraction of the original magma mass, R is the ratio of mass assimilated to mass crystallized, D is the solid-liquid partition coefficient of the element, and C_a is the concentration in the assimilated material, and

$$z = \frac{R+D-1}{R-1}$$
 12.8

(this treatment follows DePaolo, 1981a). These equations are invalid if R = 1, but as we found in Chapter 9, R will generally be less than 1. The variation of concentration as a function of F for R = 0.2 and various values of D is shown in Figure 12.23.

The variation of an isotope ratio in assimilation-fractional crystallization (AFC) is given by:

$$\epsilon_{m} = \frac{\frac{R}{R-1} \frac{C_{a}}{z} (1 - F^{-z}) \epsilon_{a} + C_{m}^{0} F^{-z} \epsilon_{0}}{\frac{R}{R-1} \frac{C_{a}}{z} (1 - F^{-z}) + C_{m}^{0} F^{-z}}$$
 12.9

where ε_m is the isotope ratio in the magma, ε_a is the isotope ratio in the assimilated material, ε_0 is the isotope ratio in the original magma and other variables are as defined above. We derived an equation for the behavior δ^{18} O during assimilation in Chapter 9. That equation (8.67)

Geochemistry

CHAPTER 12: THE CRUST



Figure 12.24. Variation of δ^{18} O with 87 Sr/ 86 Sr during AFC for a magma with an initial δ^{18} O = 5.7 and 87 Sr/ 86 Sr = 0.703, and an assimilant with 87 Sr/ 86 Sr = 0.735 and δ^{18} O = +19. All curves are for R = 0.2 (5:1), except for one with D = 2 for which R = 0.11 (labeled 9:1). dashed lines are calculated Sr concentrations (ppm) assuming an initial concentration of Sr in the magma of 500 ppm. In the case where D = 1, the problem simplifies to one of simple mixing. From Taylor (1980).

is somewhat simpler because the concentration of oxygen does not vary. AFC will produce curves on plots of δ^{18} O vs. radiogenic isotope ratios that can be quite different from the simple mixing curves in Figure 12.23. Figure 12.24 shows some examples of such curves for δ^{18} O vs. ⁸⁷Sr/⁸⁶Sr.

Composition of the Continental Crust

The continental crust is extremely heterogeneous, thus task of estimating its overall composition is a difficult one. Furthermore, only the upper part of the continental crust is exposed to direct sampling: the deepest borehole, drilled by the Russians in the Kola Peninsula has reached only 12 km and the average thickness of the continental crust is about 35 km. Therefore, geochemists must rely heavily on inferences made from indirect observations to estimate the composition of the continental crust. Beginning with Clarke (1924) and Goldschmidt (1933), a number of such estimates of the composition of the continental crust have been made. These have become increasingly sophisticated with time. Among the most widely cited works are those of Taylor and McLennan (1985, 1995), Weaver and Tarney (1984), and Shaw et al. (1986). Two very recent estimates are those of Rudnick and Fountain (1995) and Wedepohl (1995). These estimates are not entirely independent. For example, Weaver and Tarney (1984) rely

on an early version of Taylor and McLennan's upper crustal estimate; Rudnick and Fountain (1995) also rely on Taylor and McLennan's upper crustal estimate. Taylor and McLennan in turn rely on the work of Shaw (1967) for many elements, as does Wedepohl (1995). In the following, we will focus particularly on the estimates of Taylor and McLennan (1985, 1995), Rudnick and Fountain (1995), and Wedepohl (1995). In doing so, we want both to learn of the composition of the crust and to understand how these estimates are made.

We can divide the problem of estimating crustal composition into two parts. The first is to estimate the composition of the upper, accessible parts of the crust. This is referred to as the "upper crust." Direct observations provide the most important constraints on the composition of this part of the crust. The second problem is the composition of the deeper, less accessible part of the crust. For this part of the crust, indirect observations, particularly geophysical ones such as seismic velocity

0			
Plutonic Rocks	Volume %	Sedimentary Rocks	Volume <u>%</u>
Granite, Granodiorite	77	Shales	72
Quartz Diorite	8	Carbonates	15
Diorite	1	Sandstones	11
Gabbros	13	Evaporites	2
Syenites, anorthosites, peridotite	es 1	1	

TABLE 12.3. Abundance of Igneous and Sedimentary Rocks in the Upper Crust

From Taylor and McLennan (1985).

Geochemistry

W. M. White

CHAPTER 12: THE CRUST

and heat flow, provide key constraints on composition. As we shall see, these observations indicate that the continental crust is compositionally stratified, with the lower part being distinctly more mafic (i.e., richer in Mg and Fe and poorer in SiO_2 and incompatible elements). Some workers divide the deep crust into a "middle" and "lower crust", while others consider only a single entity that they refer to as the "lower crust".

The Upper Crust

Historically, three approaches to estimating the composition of the upper continental crust have been used. The first is to estimate the volume of various rock types and then use typical or average compositions of each to derive a compositional estimate. Table 12.3 gives such an estimate of the relative volumes of various igneous and sedimentary rocks in the upper crust. Continental sediments constitute about 8% of the mass of the crust; if pelagic sediments are added the total sediment mass is about 11%.

A second approach is to average analyses of samples taken over a large area. An alternative is to mix sample powders to form composites of various rock types and thus reduce the number of analyses to be made (e.g., Shaw, 1967; Eade and Fahrig, 1971; Shaw et al., 1986; Wedepohl, 1995). Studies of these kinds consistently produce an average upper crustal composition similar to that of grano-This is encouraging since granodiorite. diorite is the most common igneous rock in the crust. Such estimates also tend to produce relatively similar average concentrations for minor and trace elements, as can be seen in Figure 12.25.

A third method is to let the Earth make the composites for us. A couple of kinds of such materials are available. Goldschmidt (1933) suggested the use of glacial clays in melt-water lakes adjacent to the Pleistocene ice front. An alternative but similar approach is to use *loess*, which is fine-grained aeolian material of Pleistocene age. The most readily available of this kind of natural composite, however, is simply sediments.



Figure 12.25. Comparison of estimates of the average composition of the upper crust by Eade and Fahrig (1971) and Shaw (1967). Both were based on studies of the Canadian shield, but different areas. From Taylor and McLennan (1985).



Figure 12.26. Comparison of elemental concentrations in loess with estimated upper crust. From Taylor and McLennan (1985).

CHAPTER 12: THE CRUST

One advantage of sediments over glacial material is that whereas most glacial deposits are of Pleistocene age (but there are various deposits of ages ranging up to 2.3 Ga), sediments of all ages are available so that secular variations in crustal composition can be determined.

The advantages of using geologic composites should be obvious, but there are disadvantages as well. The primary problem is that chemical fractionations are involved in producing sediments from their parents.



Figure 12.27. Rare earth patterns of Post-Archean Australian Shale (PAAS) composite, the North American Shale Composite (NASC) and the European Shale (EC) composite. From Taylor and McLennan (1985).

Weathering of rock typically produces three fractions: sands consisting of resistant minerals, clays, and a solution. These products are transported with varying degrees of efficiency away from the site of production. Since elements tend to be concentrated in one of these three fractions, none of the fractions will have a composition representative of the parent rock. Because it is produced primarily by physical, rather than chemical, action, glacial loess is less susceptible to this kind of chemical fractionation, though some fractionation nevertheless occurs. This is illustrated in Figure 12.26. Loess is enriched in SiO₂, Hf, and Zr relative to estimated concentrations in the upper crust. This reflects the enrichment of loess in mechanically and chemically stable minerals, such as quartz and zircon, because the lighter clays are carried further from their site of origin by wind and water. Loess is also depleted in Na and Ca, reflecting loss by leaching.

Numerous studies have shown that when rock weathers to produce a sediment, the rare earth pattern of the parent is usually preserved in the sediment. This is because all the rare earths are concentrated in the clay fraction, which ultimately form shales. Other Group 3 elements (Sc and Y), as well as Th, behave similarly to the rare earths during weathering. Furthermore, rare earth patterns are remarkably similar in different shales, suggesting shales are indeed good averages of crustal composition. This is illustrated in Figure 12.27, which compares three shale composites from three continents. Because of these properties of the rare earths, S. R. Taylor and colleagues at the Australian National University used them as a point of departure for estimating the composition of the upper continental crust.

Though rare earth patterns in shales are representative of their sources, their absolute concentrations are not. Because some elements are lost to the sand fraction and others to solution (and ultimately to chemical sediments such as evaporites and carbonates), the REE are enriched in shale by about 20%. Thus Taylor and McLennan (1985, 1995) reduced REE (and Sc, Y, and Th) concentrations by 20%, in making their estimate of upper crustal composition. Elements other than the REE, Sc, Th, and Y were estimated either from known ratios of elements to one of these, or were borrowed from the estimates based on analysis of large numbers of samples. For example, Taylor and McLennan (1985) obtained the K concentration from the Th concentrations by assuming a Th/U of 3.8 and a K/U ratio of 10⁴. Their Sr concentration, on the other hand, is an average of Eade and Fahrig (1968) and Shaw (1967).

CHAPTER 12: THE CRUST

Table 12.4 lists the upper crustal compositions estimated by Taylor and McLennan (1985, 1995) and Wedepohl (1995). The estimates are broadly similar, and they agree on the concentrations of many elements within 10%. Both indicate an upper crust of "granodioritic" or "tonalitic" composition. For some elements, however, the good agreement simply reflects the reliance of both papers on the previous work of Shaw (1967), but this is not true of all elements. For example, the two estimates agree closely on the concentrations of the light rare earths despite having used completely different approaches in estimating these concentrations. For a few elements, the agreement is poor. For example, Taylor and McLennan's estimate of the copper concentration is 75% greater than that of Wedepohl's.

Plank and Langmuir (in press) suggested several revisions to Taylor and McLennan's values for Cs, Nb, Ta, and TiO₂ based on their study of marine sediments. Their values are shown in parentheses in Table 12.4. Notice that these are substantial changes: Cs increases by over 50%, while Nb and Ta de-

Major Ov	(100 (100 + 07))	т <i>я</i> - М	τ <u>Λ</u> 7	adapahl		Norr	nativo Minor	logy (T & M)
	lides (wt /0)	10.101	٧V	edepoin		Nom	lauve millera	
SiO_2		66.0	$(0, \pi())$	64.9		Quar	tz	15.7
110_2		0.5 (0.76)	0.52		Orth Albit	ociase	20.1
$F_{2}O_{3}$		15.2		3 97		Dion	e sido	61
MgO		2.2		2.24		Hvne	ersthene	9.9
CaO		4.2		4.12		Il		0.95
Na ₂ O		3.9		3.46				
K ₂ O		3.4		4.04				
	T & M	Wedepohl		Т & М	Wedepohl		Т&М	Wedepohl
Li ppm	20	22	Ga ppm	17	14	Nd ppm	26	25.9
Be ppm	3	3.1	Ge ppm	1.6	1.4	Sm ppm	4.5	4.7
B ppm	15	17	As ppm	1.5	2	Eu ppm	0.88	0.95
C ppm		3240	Se ppm	0.05	0.083	Gd ppm	3.8	2.8
N ppm		83	Br ppm		1.6	Tb ppm	0.64	0.5
F ppm		611	Rb ppm	112	110	Dy ppm	3.5	2.9
Na %	2.89	2.57	Sr ppm	350	316	Ho ppm	0.8	0.62
Mg %	1.33	1.35	Y ppm	22	20.7	Er ppm	2.3	
Al %	8.04	7.74	Zr ppm	190	237	Tm ppm	0.33	
Si %	30.8	30.35	Nb ppm	25 (13.7)	26	Yb ppm	2.2	1.5
P ppm	700	665	Moppm	1.5	1.4	Luppm	0.32	0.27
S		953	Pd ppb	0.5		Hf ppm	5.8	5.8
Cl		640	Ag ppb	50	55	Ta ppm	2.2 (0.96)	1.5
Κ %	2.8	2.87	Cd ppb	98	102	Wppm	2	1.4
Ca %	3	2.94	In ppb	50	61	Re ppb	0.4	
Sc ppm	11	7	Sn ppm	5.5	2.5	Os ppb	0.05	
Ti ppm	3000 (4560)	3117	Sb ppm	0.2	0.31	Ir ppb	0.02	
V ppm	60	53	Te ppb			Auppb	1.8	
Cr ppm	35	35	Ippm		1.4	Hg ppb		56
Mn ppm	600	527	Cs ppm	3.7 (7.3)	5.8	Tl ppb	750	750
Fe %	3.5	3.09	Ba ppm	550	668	Pb ppm	20	17
Co ppm	10	11.6	La ppm	30	32.3	Bi ppm	127	123
Ni ppm	20	18.6	Ce ppm	64	65.7	Th ppm	10.7	10.3
Cuppm	25	14.3	Pr ppm	7.1	6.3	Uppm	2.8	2.5
Zn ppm	71	52	11					

 Table 12.4. Composition of the Upper Continental Crust

T&M: Taylor and McLennan (1985,1995), Wedepohl: Wedepohl (1995). Values shown in parentheses are revisions of Plank and Langmuir (in press) to Taylor and McLennan's estimates.

Geochemistry

CHAPTER 12: THE CRUST

crease by over 50%. Taylor and McLennan's values for these elements are poorly constrained so Plank and Langmuir's values should be considered superior. Wedepohl's concentrations for these elements are intermediate between the original Taylor and McLennan values and Plank and Langmuir's revised values. The values in Table 12.4 will undoubtedly be further refined in the future.

THE MIDDLE AND LOWER CRUST

Rocks metamorphosed deep within the crust are called *granulites*; an important characteristic of such rocks is that they are anhydrous, with pyroxene replacing amphibole and biotite. They are sometimes exposed at the surface by tectonic processes and hence can provide insights into the nature of the lower crust. However, these granulite terranes have often been subjected to retrograde metamorphism (metamorphism occurring while temperatures and pressure decrease), which compromises their value. Furthermore, questions have been raised as to how typical they are of lower continental crust. These questions arise because granulite terranes are generally significantly less mafic than xenoliths from the lower crust. Xenoliths perhaps provide a better direct sample of the lower crust, but they are rare. The point is, any estimate of the composition of the middle and lower crust will have to depend heavily on indirect inference and geophysical constraints. There are two principal geophysical constraints:

- Heat flow in the continental crust. A portion of the heat flowing out the crust is produced by radioactive decay of K, U, and Th within the crust (other radioactive elements do not contribute significantly to heat generation because of their long half-lives are low abundances). The concentrations of these elements can be related to rock type, as indicated in Table 12.2.
- Seismic velocities in the continental crust. Seismic velocities depend on density, compressibility and the shear modulus, which can in turn be related to composition.

Both tell us something of first order importance about the nature of the continental crust: it is vertically zoned, becoming more mafic (i.e., richer in Fe and Mg and poorer in Si and incompatible elements) with depth. Let's consider them in greater detail.

Heat is transported conductively through the lithosphere. The equation governing heat conduction is identical to that governing diffusive chemical transport (Fick's First Law, equ. 5.90):

$$Q(z) = -k \frac{\partial T}{\partial z}$$
 12.10

where Q is the heat flow at some depth, z, k is the thermal conductivity, and $\partial T/\partial z$ is the change in temperature with depth. The thermal conductivity of rocks can be measured in the laboratory and averages about 2 Wm⁻¹K⁻¹. Thus by measuring the thermal gradient, in boreholes and mines for example, the heat flow can be calculated. The average heat flow of the continents is about 60 mW/m². This heat has two components: heat conducted out of the mantle, which is about 20 mW/m², and heat generated by radioactive decay within the continents. The concentrations of K, U, and Th observed a t the surface of the crust would produce more heat than is observed to be leaving the continental crust if these concentrations were uniform through the crust. Thus, the concentrations of these elements must

TABLE 12.5. U, TH, AND K CONCENTRATIONS AND HEAT PRODUCTION IN VARIOUS ROCK Types.

Igneous Rock Type	U (ppm)	Th (ppm)	K (%)	Th/U	K/U	Density g/cm³	Heat Production 10 ⁻⁶ W/m ⁻³
Granite / Rhyolite	3.9	16.0	3.6	4.1	0.9×10^{4}	2.67	2.5
Granodiorite/Dacite	2.3	9.0	2.6	3.9	1.1×10^4	2.72	1.5
Diorite / Andesite	1.7	7.0	1.1	4.1	0.7×10^4	2.82	1.1
Gabbro/Basalt	0.5	1.6	0.4	3.2	0.8×10^4	2.98	0.3
Peridotite	0.02	0.06	0.006	3.0	0.3×10^{4}	3.28	0.01
Continental Crust	1.25	4.8	1.25	3.8	1.0×10^4		0.8

Geochemistry

CHAPTER 12: THE CRUST

decrease with depth. The concentrations of these elements are related to rock type, as is illustrated in Table 12.5. The problem is complicated, however, by variations in the "mantle" heat flow. Heat flow varies significantly with tectonic age, as is illustrated in Figure 12.28. If, as we believe, the con-

tinental crust is created by magmatism, it will be initially hot and then cool over time. Subsequent episodes of magmatism may also heat the crust. In addition, the variation seen in Figure 12.28 may represent different thickness of the lithosphere (Vitorello and Pollack, 1990; Nyblade and Pollack, 1993). The lithosphere is a conductive boundary layer, so that the thicker the lithosphere, the lower the mantle heat flow out the top of it. Nyblade and Pollack (1993) have argued that regions of old Archean crust are underlain by particularly thick mantle lithosphere, an argument supported by geochronological and thermobarometric studies of mantle xenoliths from these regions.

Seismic velocities increase with depth in the crust: P-wave velocities increase from about 6 km/sec in the upper crust to about 7 km/sec in the lower crust. The dependence of the velocity of P waves (compressional-waves) on physical properties is:

$$v_{\rm P} = \sqrt{\frac{K + \frac{4}{3}\mu}{\rho}}$$
 12.11

and that for velocity of shear waves (S-waves) is:

$$v_{\rm S} = \sqrt{\frac{\mu}{\rho}}$$
 12.12



Figure 12.28. Heat flow as a function of tectonic age. Component I is radiogenic heat produced in the crust, II is heat from a transient thermal perturbation associated with tectonism, and III is background heat from deeper sources. A is Archean, EPr is Early Proterozoic, LPr is Late Proterozoic, EPa is early Paleozoic, LPa is late Paleozoic, M is Mesozoic, and C is Cenozoic. From Vitorello and Pollack (1980).



Figure 12.29. Seismic velocity structure of the continental crust, illustrating its 3-layered nature. Velocity structure falls into 9 types. The number of profiles used to construct each type is shown below each type. From Rudnick and Fountain (1995).

Geochemistry

CHAPTER 12: THE CRUST

where K is the bulk modulus (inverse of β , compressibility), μ is the shear modulus, and ρ is density. Thus the increase in seismic velocity with depth implies increasing density and decreasing compressibility with depth. Some, but not all, of the increase in density results merely from compression. The remainder represents changing composition. The real problem is to understand precisely how seismic velocity depends on composition.

Seismic velocity profiles vary widely from place to place, as does crustal thickness. Rudnick and Fountain (1995) examined a global database of seismic cross sections and found that they can be divided into 9 classes, which are illustrated in Figure 12.29. One must next relate seismic velocity to composition. This can be done by making measurements of seismic velocity in the laborator Figure 12.30 shows the relationship between SiO₂



Figure 12.30. Correlation between measured seismic velocity (v_P) and SiO₂ concentration. From Rudnick and Fountain (1995).

measurements of seismic velocity in the laboratory on samples of known composition. For example, Figure 12.30 shows the relationship between SiO_2 and seismic velocity in a variety of rock types.

To produce an estimate of crustal composition, Rudnick and Fountain assigned an average lithology to the seismic sections shown in Figure 12.29. They then assigned a composition to each lithology using a database of the composition of lower crustal xenoliths. Then by estimating the aerial extent of each type of crustal section, they produced the compositional estimate in Table 12.6. This table shows that the composition of the lower crust corresponds to that of tholeiitic basalt; in metamorphic terminology it would be a mafic granulite. The composition of the middle crust corresponds to that of an andesite. At the prevailing pressures and temperatures this rock would be an amphibolite, consisting mainly of amphibole and plagioclase.

Wedepohl (1995) used the European Geotraverse as a model of the seismic structure of the crust. This seismic cross-section runs from northern Scandinavia to Tunisia and crosses a great variety of tectonic provinces, ranging from the Archean Fennoscandian Shield to the young fold belts to the young Alpine orogen. He assigned 3 lithologies to 3 ranges of seismic velocities: sediments, granites, and gneisses ($V_P < 6.5 \text{ km/s}$) corresponding to the upper crust, felsic granulites ($6.5 < V_P < 6.9 \text{ km/s}$), and mafic granulites ($6.9 < V_P < 7.5 \text{ km/s}$). He used a database of compositions of felsic and mafic granulites from both xenoliths and exposed terranes to calculate an average composition for each of the latter two. He then computed a lower crustal composition by weighting felsic and mafic granulites in the proportions their characteristic seismic velocities were observed in the European Geotraverse. His estimate of the composition of the lower crustal is also listed in Table 12.6.

Rare earth patterns of upper, middle and lower crust as estimated by Rudnick and Fountain (1995) are compared in Figure 12.31a. The negative Eu anomaly in the upper crust and slight positive anomalies in the middle and lower crust (such positive anomalies are typical of many granulites) are an interesting features of these patterns. Eu is strongly held in plagioclase (Chapter 7). The presence of plagioclase in the melting residue would produce a negative Eu anomaly in the melt and a positive one in the residue. Thus these anomalies suggest that crustal has differentiated to form distinct layers at least partially through partial melting, with granitic melts forming the upper crust and granulitic residues forming the deeper crust. Figure 12.31b displays the estimated composition of the middle and lower crust relative to the upper crust. The lower and middle crust are depleted in incompatible elements and enriched in compatible elements relative to the upper crust. This is also consistent with the idea that magmatic processes have been important in creating the compositional layering observed in the crust.

Geochemistry

CHAPTER 12: THE CRUST

Major Ox	ides, %						
	R & F	R & F	Wedepohl				
	Lower	Middle	Lower				
SiO ₂	52.3	60.6	58.05				
TiO ₂	0.8	0.7	0.84				
Al_2O_3	16.6	15.5	15.52				
FeO	8.4	6.4	7.34				
MnO	0.1	0.1	0.12				
MgO	7.1	3.4	5.23				
CaO	9.4	5.1	6.80				
Na ₂ O	2.6	3.2	2.86				
K ₂ O	0.6	2.01	1.85				
P_2O_5	0.1	0.1	0.20				
Trace Ele	ments, pp	m					
	R & F	R & F	Wedepohl		R & F	R & F	Wedepohl
	Lower	Middle	Lower		Lower	Middle	Lower
Li	6	7	13	Sn			2.1
Be			1.7	Sb			0.3
В			5	Ι			0.14
С			588	Cs	0.3	2.4	0.8
Ν			34	Ва	259	402	568
F			429	La	8	17	26.8
S			408	Ce	20	45	53.1
Cl			278	Pr	2.6	5.8	7.4
Sc	31	22	25.3	Nd	11	24	28.1
V	196	118	149	Sm	2.8	4.4	6.0
Cr	215	83	228	Eu	1.1	1.5	1.6
Со	38	25	38	Gd	3.1	4	5.4
Ni	88	33	99	Tb	0.48	0.58	0.81
Cu	26	20	37.4	Dy	3.1	3.8	4.7
Zn	78	70	79	Но	0.68	0.82	0.99
Ga	13	17	5.4	Er	1.9	2.3	
As			1.3	Tm			0.81
Se			0.17	Yb	1.5	2.3	2.5
Br			0.28	Lu	0.25	4.1	0.43
Rb	11	62	41	Ηf	1.9	4	4.0
Sr	348	281	352	Та	0.6	0.6	0.84
Y	16	22	27.2	W			0.6
Zr	68	125	165	Tl			0.26
Nb	5	8	11.3	Hg			0.21
Mo			0.6	Pb	4.2	15.3	12.5
Ag			0.08	Bi			0.037
Cd			0.101	Th	1.2	6.1	6.6
In			0.052	U	0.2	1.6	0.93

TAble 12.6. Composition of the Middle and Lower Continental Ci	RUST
--	------

R & F: Rudnick and Fountain (1995), Wedepohl: Wedepohl (1995).

Geochemistry

CHAPTER 12: THE CRUST



Figure 12.31. (a). Comparison of chondrite-normalized in upper, middle and lower crust. (b). Elemental enrichment or depletion of the middle and lower crust relative to the upper crust. From Rudnick and Fountain (1995).

THE TOTAL CONTINENTAL CRUST

The approach used by most workers to estimate the composition of the total continental crust is simply to calculate an average of two or three crustal sections, weighting each by its mass. This was done, for example, by Rudnick and Fountain (1995) and Weaver and Tarney (1984), both of whom divided the crust into an upper, lower, and middle section. Both Weaver and Tarney (1984) and Rudnick and Fountain (1995) relied on Taylor and McLennan's upper crustal estimate. Weaver and Tarney (1984) used average Lewisian* amphibolite as their middle crust composition and average Lewisian granulite as their lower crust composition. Shaw et al. (1986) and Wedepohl (1995) used a similar approach, but divided the crust only into upper and lower parts. An important step in this approach is estimating the thickness of the various sections. All use geophysical constraints for this.

Taylor and McLennan (1985, 1995) used an entirely different approach to estimating total crustal composition, one

based on the "andesite model" of Taylor (1967). Taylor (1967) noted the role played by subduction-related volcanism in creation of the continental crust and assumed that on average the crust consisted of island arc andesite. Thus average island arc andesite was used as the estimated composition of the continental crust. This approach was modified in subsequent work, as Taylor concluded that while post-Archean crust was created at subduction zones, Archean crust was not and is compositionally different. Taylor and McLennan (1985) essentially modify the Taylor (1967) andesite model for their estimate of Archean crustal composition.

Estimates of the major element composition of the continental crust by Weaver and Tarney (1984), Shaw et al. (1986), Taylor and McLennan (1995), Wedepohl (1995), and Rudnick and Fountain (1995) are given in Table 12.7. Also listed are estimates of trace element concentrations by Taylor and McLennan (1995), Rudnick and Fountain (1995), and Wedepohl (1995). Since Rudnick and Fountain (1995) rely on the Taylor and McLennan upper crustal estimates in calculating the total crust composition, the revisions to the Taylor and McLennan upper crustal values for TiO₂, Cs, Nb, and Ta of Plank and Langmuir (in press) affect the Rudnick and Fountain total crustal values of these elements. These revisions have been made in the Table.

The ranges of estimates for SiO₂ and Al₂O₃ in Table 12.7 are about 10% and 8% respectively; the range in Mg# (52 to 57) is similarly only about 10%. Interestingly, earlier estimates of crustal SiO₂ and Al₂O₃, going back to Goldschmidt (1933) also fall within this range. Thus we can conclude with some confidence that the continental crust on the whole is similar to that of siliceous andesite.

The details of the composition of the crust are less certain, however. Ranges for the other oxides are substantially larger: 75% for FeO, 68% for MgO, and 100% for MnO. Of these estimates, the com-

^{*} The Lewisian, which outcrops in Northwest Scotland, is perhaps the classic exposure of lower crust.

CHAPTER 12: THE CRUST

position of Taylor and McLennan is the most mafic, and that of Weaver and Tarney the least mafic (ranges for FeO and MnO decrease to 30% and 21% respectively if the estimates of Taylor and McLennan are excluded).

The Th/U ratio in both Taylor and McLennan's and Rudnick and Fountain's estimate (3.85 and 3.94 respectively) is probably too low. The Th/U ratio of the depleted mantle is well constrained by ²³⁰Th/²³²Th ratios in MORB to be <2.5. If the bulk Earth value is 3.9-4.2 (as would be the case if the terrestrial Th/U is chondritic), then mass balance requires the crustal ratio to be higher than the bulk Earth value, yet both Taylor and McLennan's and Rudnick and Fountain's Th/U ratios are close to bulk Earth. Wedepohl's Th/U ratio of 5 may be more appropriate.

Figure 12.32 shows the primitive mantle-normalized incompatible element concentrations in the total continental crust. All these estimates show the continental crust to be enriched in incompatible elements. This is consistent with the view that the crust was created by partial melting of the mantle. The elemental enrichment pattern is not smooth, however. First, Nb (and Ta, not shown), and to a lesser degree Ti, are not as enriched as elements of similar incompatibility. Second, Pb is anomalously enriched. As we found earlier in the chapter, these are features of island arc magmas and their presence here strengthens the notion that island arc volcanism has

Table 12.7. Composition of the Continental Crust

Major Oxides, wt. %							
	R&F 7	С& М	W & T	We	Shaw		
SiO ₂	59.1	57.3	63.2	61.5	63.2		
TiO ₂	0.75	0.9	0.6	0.68	0.7		
Al_2O_3	15.8	15.9	16.1	15.1	14.8		
FeO	6.6	9.1	4.9	5.67	5.60		
MnO	0.1	0.18	0.08	0.10	0.09		
MgO	4.4	5.3	2.8	3.7	3.15		
CaO	6.4	7.4	4.7	5.5	4.66		
Na ₂ O	3.2	3.1	4.2	3.2	3.29		
K ₂ O	1.88	1.1	2.1	2.4	2.34		
P_2O_5	0.2		0.19	0.18	0.14		
Trace E	lements (i	n ppm	unless oth	erwise	noted)		

		~ ~					
	R & F	Т&М	We		R & F	Т&М	We
Li	11	13	18	Sb		0.2	0.3
Be		1.5	2.4	Te, ppb			5
В		10	11	I, ppb			800
С			1990	Cs	3	1	3.4
N			60	Ва	390	250	584
F			525	La	18	16	30
S			697	Ce	42	33	60
Cl			472	Pr	5	3.9	6.7
Sc	22	30	16	Nd	20	16	27
V	151	230	98	Sm	3.9	3.5	5.3
Cr	119	185	126	Eu	1.2	1.1	1.3
Со	25	29	24	Gd	3.6	3.3	4.0
Ni	51	105	56	Tb	0.56	0.6	.65
Cu	24	75	25	Dy	3.5	3.7	3.8
Zn	73	80	65	Но	0.76	0.78	0.8
Ga	16	18	15	Er	2.2	2.2	2.1
Ge		1.6	1.4	Tm		0.32	0.3
As		1	1.7	Yb	2	2.2	2.0
Se		0.05	0.12	Lu	0.33	0.3	0.35
Br			1.0	Ηf	3.7	3	4.9
Rb	58	32	78	Та	0.7	1	1.1
Sr	325	260	333	W		1	1.0
Y	20	20	24	Re, ppb		0.4	0.4
Zr	123	100	203	Os, ppb		0.005	0.05
Nb	8.5	11	19	Ir, ppb		0.1	0.05
Мо		1	1.1	Pt, ppb			0.4
Ru, ppb			0.1	Au, ppb		3	2.5
Rh, ppb			0.06	Hg, ppb			40
Pd, ppb		1	0.4	Tl, ppb		360	520
Ag, ppb		80	70	Pb	12.6	8	14.8
Cd, ppb		98	100	Bi, ppb		60	85
In, ppb		50	50	Th	5.6	3.5	8.5
Sn		2.5	2.3	U	1.42	0.91	1.7

R & F: Rudnick and Fountain (1995) as revised by Plank and Langmuir (in press), T & M: Taylor and McLennan (1985, 1995), We: Wedepohl (1995), Shaw: Shaw et al. (1986), W & T: Weaver and Tarney (1984).

Geochemistry

CHAPTER 12: THE CRUST



Figure 12.32. Comparison of incompatible element enrichment of estimated bulk crustal composition of Taylor and McLennan (1985), Rudnick and Fountain (1995) with other estimates (Weaver and Tarney, 1984; Shaw et al., 1986; Wedepohl, 1995). From Rudnick and Fountain (1995).

played an important role in creation of the continental crust.

Rudnick and Fountain's (1995) or Wedepohl's (1995) estimated compositions are significantly more enriched in incompatible elements than the estimated composition of Taylor and McLennan (1985, 1995). Since the former two estimates are based on observations while the Taylor and McLennan estimate is based on assumptions about how the crust formed, the former estimates should probably be preferred. This strong enrichment of the continental crust in incompatible elements has important implications for the evolution of both the crust and the mantle. Based on Rudnick and Fountain's (1995) or Wedepohl's (1995) estimates, the crust is more enriched in incompatible elements than typical island arc andesite. Thus subduction-related magmatism alone is

inadequate as a mechanism for creation of the continental crust. Based on these estimates, a substantial proportion of the Earth's total inventory of lithophile incompatible elements are in the crust. For example, using the primitive mantle values listed in Chapter 11, we can calculate that 50% or 67% of the Earth's total Rb is in the Earth's crust using Rudnick and Fountain's and Wedepohl's crustal estimates respectively. Similar fractions of other highly incompatible elements such as Cs, Ba, and Th are also concentrated in the crust. This implies that the melt must have been extracted from over half the mantle to create the crust, and therefore that at least half the mantle is incompatible element depleted. Thus differentiation of

element depleted. Thus differentiation of the Earth's has been remarkably efficient.

GROWTH OF THE CONTINENTAL CRUST

The composition of the crust gives us important clues as to how it was created. For example, one possible model for the formation of the continental crust is that it was produced by late accretion of a volatile-rich veneer when the Earth formed. But the composition given in Table 12.7 is clearly inconsistent with this view: the crust is not systematically enriched in volatile elements. The crust is systematically enriched in incompatible elements; this leads to the hypothesis that the crust was created by partial melting of the mantle. However, even here there are some inconsistencies, since a partial melt of the mantle is unlikely to have as much SiO₂ as the crust ($\sim 60\%$).



Figure 12.33. Models of the rates of crustal growth. AM: Armstrong (1981a), R&S: Rymer and Schubert (1984), F: Fyfe (1978), D & W: DePaolo and Wasserburg (1979), M & T: McLennan and Taylor (1982), O'N: O'Nions and Hamilton (1981), V & J: Veizer and Jansen (1979), H & R: Hurley and Rand (1969). Adapted from Taylor and McLennan (1985).
Geochemistry

CHAPTER 12: THE CRUST

Also, the trace element composition of the crust is not like that of any single type of mantle-derived magma. This suggests the growth of the crust has been a complex process. The details of this process are not yet fully understood. In this section, we examine what is known about crustal growth.

THE PACE OF CRUSTAL GROWTH

The first question we might ask is how has the crust grown with time? A variety of answers to this question have been suggested. These are illustrated in Figure 12.33. They can be broken into three types: (1) growth rate increasing through time, e.g., curve V & J (=Vezier and Jansen) and H & R (Hurley and Rand), (2) approximately linear growth through time, such as the curve marked O'N (= O'Nions), and (3) early rapid growth followed by later slow growth or no growth, e.g., curves Am (=Armstrong) and F (Fyfe).

Early studies favored accelerating crustal growth through time. Age provinces for North America based on radiometric dates compiled by Hurley and





Rand (1969) are shown in Figure 12.34. Based on this, Hurley suggested most of the crust was produced in the last 1 or 2 Ga. However, subsequent work utilizing U-Pb dating of zircons and Sm-Nd model ages has substantially changed this view. U-Pb zircon ages are not so easily reset by metamorphism as are the Rb-Sr ages relied on by Hurley. Zircon dating has revealed large areas of the continental crust that were created more than 2.7 Ga ago. Zircon dating as also identified limited areas of crust that are old than 3.5 Ga. The oldest dates are from two localities in Australia. Zircons from Mt. Narryer, Western Australia give ages of around 4.15 Ga (Frounde et al., 1983). Slightly older ages (4.2-4.3 Ga) were determined on zircons from a second, nearby locality, the Jack Hills. These zircon analyzes were done by ion probe rather than conventional mass spectrometry and were initially controversial for that reason, but they ages are now generally accepted. In both cases, the zircons are found in metasedimentary rocks of much younger age (3–3.5 Ga), and only a handful of the zircons in these rocks are this old. While these ancient zircons have been preserved, the rocks in which they crystallized apparently have not. Nevertheless, the zircons are relatively U-rich, suggesting they come from silica-rich rocks typical of true continental crust rather than more mafic rocks.

More recently, zircons in the Acasta gneisses from the Slave Province in Canada have been dated at 3.96 Ga by this method (Bowring et al., 1989). These ages are interpreted as the age of crystallization of protoliths of these gneisses. Interestingly, their initial ε_{Nd} are negative, indicating derivation from a light rare earth-enriched source, which could be even older crust. Thus there is evidence that at least some continental crust formed very early in Earth's history.

In Chapter 11, we found that the many very ancient rocks have positive initial ε_{Nd} ratios. For example, the Isua rocks of East Greenland, which have crystallization ages around 3.8 Ga, have initial ε_{Nd} around +1 to +3; rocks of similar age from India have similar initial ε_{Nd} values. This implies their sources had high Sm/Nd ratios, and therefore were LRE-depleted for a substantial amount of

Geochemistry

W. M. White

CHAPTER 12: THE CRUST

time, hundreds of millions of years, before the crystallization of these rocks. This requires that a complimentary LRE enriched reservoir, possibly an early continental crust, formed before 4.0 Ga. How large this crust reservoir was, however, depends on how large the volume of depleted mantle was, which is not constrained (McCulloch and Bennett, 1994). Nevertheless these data provide indirect evidence that at least crust formation began well before 4.0 Ga.

Sm-Nd model ages also have the power to "see through" metamorphism and establish "crust formation ages", as we found in Chapter 8. The example of the



Figure 12.35. Isotopic provinces, based on crustal residence times (τ_{DM}) of the Western U.S. From Bennett and DePaolo (1987).

western U. S. illustrates this point. Figure 12.35 is a map of the Western U. S. showing contours of Nd crustal residence times (τ_{DM}). The data define 3 distinct provinces and suggest the existence of several others. There is a general similarity to Hurley's map (Figure 12.34), but there is greater detail, and the ages are generally older.

Figure 12.36 shows the initial ε_{Nd} values of the granites from the three numbered provinces of Figure 12.35 plotted as a function of their crystallization age. Despite the variations in crustal residence times, the crystallization ages indicate Provinces 1-3 all formed between 1.65 and 1.8 Ga. Only the ε_{Nd} from Province 3 plot close to the depleted mantle evolution curve. From this we can conclude that only Province 3 was a completely new addition to the crustal mass at that time. Initial ε_{Nd} for the remaining provinces plot below the depleted mantle evolution curve, suggesting they are mixtures of new mantle-derived material and older crust. The crustal residence ages of these provinces are older than the crystallization ages because they are mixtures of mantle and older crust.

In each province there have been subsequent episodes of magmatism. However, the initial ϵ_{Nd} lie along the same growth trajectory as the older rocks. This suggests that magmatism in these subsequent episodes simply recycled pre-existing crustal material and there were no new additions to crustal mass from the mantle. Thus Hurley's map (Figure 12.34) must reflect orogeny in which the radiogenic clocks are reset rather than new crustal additions.

At the other extreme of the continental growth question, Armstrong (1968, 1981a) argued that the mass of the crust has remained nearly constant of the past 4 Ga or so. Armstrong recognized that new crust has been continually created through time, but he argued that the rate of crustal creation was balanced by an equal rate of crustal destruction through erosion and subduction of sediment. Three lines of evidence support this view. First, Armstrong cited the absence of ancient pelagic sediment anywhere on the Earth's surface and argued that most sediment must therefore be subducted. In support of this, he cited estimates of sediment subduction rates by several workers that are sufficient to balance crustal growth. Second, he pointed out that neither sea level nor the average thickness of stable continental cratons have changed with time, and therefore that continental volume also must not have changed. Finally, he pointed out that positive ε_{Nd} in the earliest rocks requires a very early depletion of the mantle (Figure 10.19), most likely through generation of continental crust. Further-

Geochemistry

CHAPTER 12: THE CRUST



Figure 12.36. ε_{Nd} (initial) as a function of crystallization age of Western U.S. Groupings 1 2, and 3 refer to provinces shown in Figure 12.35. From Bennett and DePaolo (1987).

with their crustal residence ages. We expect that crustal residence ages will be somewhat older than the stratigraphic ages. Only when a rock is eroded into the sedimentary mass immediately after its derivation from the mantle will its stratigraphic (τ_{ST}) and crustal residence age (τ_{CR}) be equal.

The top diagram illustrates the relationships between τ_{sT} and τ_{CR} that we would expect to see for various crustal growth scenarios, assuming there is a relationship between the amount of new material added to the continents and the amount of new material added to the sedimentary mass. If the continents had been created 4.0 Ga ago and if there had been no new additions to continental crust since that time, then the crustal residence time of all sediments should be 4.0 Ga regardless of stratigraphic age. This is illustrated by the line labeled "No New Input". If, on the other hand, the rate of continent growth through time has been uniform since 4.0

more, the subsequent nearly linear growth of ε_{Nd} is consistent with constant continental mass and inconsistent with continental growth (Armstrong, 1981b, DePaolo, 1981b).

Because sediments effectively sample rocks of a wide variety of ages, isotope ratios in sediments provide a means of studying crustal evolution. By studying sediments of various ages, we should be able to draw some inferences about the rates of continental growth. Goldstein et al. (1984) found that the mean crustal residence time (τ_{DM}) of the modern river particulates was 1.7 Ga, which they interpreted as the mean age of the crust now being eroded. However, they estimated the mean crustal residence time of the entire sedimentary mass to be about 1.9 Ga. Figure 12.37 compares the stratigraphic age^{*} of sediments



Figure 12.37. Relationship between stratigraphic age of sediments and the crustal residence age of material in sediments. From Goldstein et al., 1984).

^{*} The stratigraphic age is the age of deposition of the sediment determined by conventional geochronological or geological means.

Geochemistry

CHAPTER 12: THE CRUST

Ga, then τ_{ST} and τ_{CR} of the sedimentary mass should lie along a line with slope of 1/2, which is the line labeled "Uniform Rate". The reason for this is as follows. If the sedimentary mass at any given time samples the crust in a representative fashion, then τ_{CR} of the sedimentary mass at the time of its deposition (at τ_{ST}) should be $(4.0 - \tau_{\text{ST}})/2^+$, i.e., the mean time between the start of crustal growth (which we arbitrarily assume to be 4.0 Ga) and τ_{ST} . A scenario where the rate of crustal growth decreases with time is essentially intermediate between the one-time crust creation at 4.0 and the uniform growth rate case. Therefore, we would expect the decreasing rate scenario to follow a trend intermediate between these two, for example, the line labeled 'Decreasing Rate'. On the other hand, if the rate has increased with time, the τ_{CR} of the sedimentary mass would be younger than in the case of uniform growth rate, but still must be older than τ_{ST} , so this scenario should follow a path between the uniform growth rate case and the line $\tau_{\text{ST}} = \tau_{\text{CR}}$, for example, the line labeled "Increasing Rate".

Line A in Figure 12.37b is the uniform growth rate line with a slope of 1/2. Thus the data seem to be compatible with a nearly uniform rate of growth of the continental crust, such as the line labeled "O'N" in Figure 12.33 or even with an increasing growth rate. However, Goldstein et al. noted sedimentary mass is cannibalistic: sediments are eroded and redeposited. Line B represents the evolution of the source of sediments where crustal growth is constant but erosion and re-deposition of old sediments occurs. The situation is further complicated when crust is destroyed through erosion and subduction of marine sediments. Goldstein et al. (1984) concluded their data could be consistent with either nearly constant continental growth rate or a nearly constant continental mass if the rate rate of continent-to-mantle recycling decreased through time.

In summary, formation of the continental crust began 4.0 Ga ago or earlier. The rarity of rocks of this age, however, suggests either that crustal growth was initially quite slow, or that much of this early continental crust may have been destroyed and recycled into the mantle. Since then, there have been continued additions to the continental crust. The degree to which destruction of crust through erosion and subduction may have balanced these additions is unclear and remains highly controversial. The average age of the presently existing continental crust is around 2.2 Ga. Models, such as that of Hurley, in which the rate of crustal growth is increasing cannot be correct. Reconnaissance Sm-Nd studies of the sort illustrated in Figure 12.35 show that the crust is created in large blocks, and that the rate has been somewhat episodic. Judging from these kinds of studies, the late Archean and early Proterozoic (the period between roughly 3.0 and 1.8 Ga) appears to have been a time of relatively rapid crustal growth (e.g., Patchett and Arndt, 1986; McCulloch, 1987). On the other hand, crustal growth rates in the Phanerozoic (the last 550 Ma) appear to be lower than during earlier times.

Mechanisms of Crustal Growth

We have seen how the composition of the continental crust is qualitatively consistent with the crust having been created by partial melting of the mantle. But what are the details of this process? In what tectonic setting has the crust been produced? We can identify a number of possible mechanisms for the creation of the continental crust. These are as follows:

$$\tau_{\rm CR} = \frac{4.0 + \tau_{\rm ST}}{2}$$

At time of deposition, its crustal residence age would have been: $\tau_{CR} = \frac{4.0 + \tau_{ST}}{2} - \tau_{ST} = \frac{4.0 - \tau_{ST}}{2}$.

⁺ One way to rationalized this equation is to think of newly deposited sediment at τ_{sT} as a 50-50 mixture of material derived from the mantle at 4.0 Ga and τ_{sT} . The equation for the τ_{CR} of this mixture would be:

You could satisfy yourself that a mixture of material having τ_{CR} of all ages between 4.0 Ga and τ_{ST} would have the same τ_{CR} as given by this equation.

Geochemistry

CHAPTER 17: THE CRUST

	12.0. 00		<u> </u>				
<u>Continental Crust with Andesite</u>							
	Continental	Average					
	Crust	Andesite					
SiO ₂	57.3-63.2	59.7					
TiO ₂	0.6-0.9	0.7					
Al_2O_3	14.8-16.1	17.1					
FeO	5.6-9.1	6.3					
MnO	0.08-0.18	0.12					
MgO	2.8-5.3	3.2					
CaO	4.7-7.4	6.6					
Na ₂ O	3.1-3.2	3.3					
K ₂ O	1.1-2.4	1.5					
P_2O_5	0.14-0.20	0.19					

TADLE Contin	12.8. ental Cri	Co ust	mparis with A	on Andes	lо siте
	Continer	ntal	Avera	age	
	Crush		A made	- : +-	

• Subduction-related volcanism. The Andes Mountains with their magnificent and very active andesitic volcanos represent a modern example of how the crust grows above a subduction zone.

• Accretion of oceanic crust and oceanic plateaus. Oceanic crust is usually subducted and returned to the mantle, but anomalously thick crust, such as that of the Ontong-Java Plateau of the Western Pacific may be difficult to subduct. The alternative is that is accretes to continental margins, becoming part of the continental mass.

• Continental volcanism. Rifting and mantle plumes are two causes of volcanism unrelated to subduction. Flood basalts, probably produced by the start-up phase of mantle plumes, have occasionally been erupted in tremendous volumes.

Continental Crust from Table 12.7. Average medium-K orogenic andesite from Gill (1981).

 Underplating. Magmas erupted on continents may represent only the tip of the iceberg so to speak. Because of the low density of the continental crust, basalt magma may stagnate at the crust-mantle boundary. In this scenario, the crust primarily grows from the bottom down.

 Intrusion by small-degree melts. Small degree melts of shallow mantle under the continents that are highly enriched in incompatible elements, such as kimberlites and lamprophyres, may not contribute significant volumes to the continents, but may contribute a disproportionate fraction of the incompatible element inventory.

Of course, all these mechanisms have contributed to crustal growth to varying degrees. The real question is which mechanism is most important?

There are a number of cogent arguments favoring subduction-related magmatism as the principal mechanism of crustal growth. First, in a qualitative way, the continental crust has a major element composition similar to andesites erupted on continental margins and in island arcs. This point is made in Table 12.8, which compares average "orogenic" (i.e., island arc or continental margin) andesite to the estimates of crustal composition in Table 12.7. The resemblance is strong; only the Al concentration lies outside the range of estimated concentrations in the crust.

The similarity in composition also extends to incompatible elements. This point is made in Figure 12.38, which compares the composition of a siliceous andesite from the Banda arc (Indonesia) with the range of estimates of crustal composition. While the match is not an exact one, the continental crust shares the strong enrichment in highly incompatible elements, the relative depletion in Nb and Ta, and the excess enrichment in Pb observed in most island arc magmas. We should point out, however, that to some extent, the incompatible element enrichment of island arc magmas results from a "continental component" in them, originating either through sediment subduction or assimilation of crust. Thus in trying to understand the incompatible element enrichment of the crust, postulating an origin by subduction zone magmatism begs the question in a certain sense.

Finally, as we noted earlier, subduction-related magmatism is the most important mechanism of crustal growth at present, and probably throughout the Phanerozoic as well. Many igneous and metaigneous rocks of Proterozoic age also have chemical features suggesting they were produced in subduction settings. Bennett and DePaolo (1987) concluded the provinces in the Southwest U.S. shown in Figure 12.35 formed by successive accretion or growth of island arcs on the edge of the pre-existing Archean Wyoming Craton to the north. Judging from their older Sm-Nd model ages, the northeastern arcs contain a substantial component of older crust derived from the craton. This could have occurred through erosion and subduction, or, if the volcanos were built directly on the continent, through assimilation of crust. As new Proterozoic crust was built outward from the continent, it screened subse-

Geochemistry

CHAPTER 12: THE CRUST



Figure 12.38. Comparison of incompatible element concentrations in a siliceous andesite from the Banda arc (red triangles) with the range of estimated concentrations in the continental crust (gray field). Both share a relative depletion in Nb and Ta and a relative enrichment in Pb. quent arcs from the contribution of material from the Archean crust. A similar effect has been observed in the Proterozoic provinces of Canada.

Accretion of intra-oceanic arcs, modern examples of which include the Marianas and the Aleutians, may also contribute to crustal growth. Indeed, major segments of continental crust appear to have grown so rapidly that the could not have been created in a single arc (Rymer and Schubert, 1986); they therefore may have originated as several simultaneously active arcs subsequently accreted to a continental margin. Even in this case, however, global magma production rates at subduction zones during at least some parts of

the Proterozoic must have been 6 to 10 times higher than during the Phanerozoic (Rymer and Schubert, 1986). These high growth rates may mean some mechanism other than subduction-related volcanism was responsible for growth of much of the crust.

Though subduction-related volcanism may have been the dominant mode of crustal growth in the Proterozoic and Phanerozoic, other mechanisms have played a role. The Wrangalia Terrane in NW British Columbia and Alaska, is widely considered to consist in part of oceanic plateaus. The plateaus were produced over a mantle plume in Paleozoic time and later accreted to the North American continent by plate tectonic processes. The Coast Ranges of Oregon represent an example of more recently accreted oceanic crust. Mantle plumes surfacing beneath continents also produce magmas that add mass to the continents. The most voluminous eruptions occur in the initial stages of the plume, when the large buoyant plume head approaches the surface. Under these circumstances, enormous volumes of basalt erupt. Examples of such flood basalts include the Siberian Traps, the Karoo of South Africa, the Deccan of India, the Parana of Brazil, and the Columbia River of the NW U. S. Gravity anomalies suggest even greater volumes of basaltic magma were trapped at deep crustal levels.

Continental rifts can also be sites of voluminous eruption of basaltic magma. A well-documented example is the Proterozoic Keweenawan or Mid-Continent Rift of the U. S., which formed some 1 to 1.2 Ga ago. Though now mostly covered by Phanerozoic sediments, where it is exposed the rift consists of a trough 150 km wide and 1500 km long filled with up to 15 km of volcanics, primarily basalt, and clastic sediments derived from them. Modern examples of continental rifts include the Rio Grande Rift of New Mexico and the East African Rift.

Crustal Growth in the Archean

The case for crustal growth through subduction-related volcanism is much less strong in the Archean. While andesites dominate island-arc volcanism at present, andesites seem to be relatively rare in the Archean. Taylor and McLennan (1995) remarked that "the composition of the Archean upper crust stands in marked contrast to that of the post-Archean crust." This suggests the Archean-

Geochemistry

CHAPTER 12: THE CRUST

Proterozoic boundary marks a change in the manner in which crust was generated. Since nearly half the continental crust was created by the end of the Archean, understanding the genesis of Archean crust is important.

There are two principle types of Archean terranes. The first is greenstone belts, of which the Abitibi Belt in the Superior Province of Canada is a classic example. They consist of thick sequences of volcanic and sedimentary rocks in elongate basins punctuated by circular or elongate granitic batholiths. They have experienced only low grade metamorphism (hence the term greenstone: chlorite-serpentinite is the typical metamorphic grade).



Figure 12.39. REE patterns of typical Archean tonalites and trondhjemites. From Taylor and McLennan (1985).

Tholeiitic basalt and komatiite are the predominant volcanics, but more siliceous ones are also common. The combination of mafic volcanics and granitic intrusives makes greenstone belts distinctly bimodal in composition.

The presence of komatiites is significant. They are ultramafic lavas erupted at temperatures of 1400 to 1600° C, much hotter than basalts, whose eruption temperatures are typically 1100 to 1200° C. Their chemistry indicates they are products of large extents of melting, perhaps up to perhaps 40%. Komatiites are largely restricted to the Archean; there are a few early Proterozoic occurrences and only one documented Phaner-ozoic occurrence. The absence of komatiites in the latter part of Earth's history undoubtedly reflects secular cooling of the mantle.

The other type of Archean crust, the high-grade gneiss terranes, such as West Greenland, is much different. They typically consist of felsic gneisses and sedimentary and volcanic rocks metamorphosed at amphibolite to granulite grade. These terranes do not seem to be simply a highly metamorphosed version of the greenstone belt terranes as they differ in structure and in sedimentary facies. The felsic gneisses consist of metamorphosed plutonic rocks of the so-called "TTG" (Tonalite*– Trondhjemite[†]–Granodiorite) suite, characterized by higher Na/K ratios than are found in most post-Archean granitoids. Many have steep REE patterns (Figure 12.39) and no Eu anomalies. These rare earth characteristics indicate they formed by partial melting at great depth (>60 km), where garnet, rather than plagioclase, was present in the residua. The presence of garnet would account for their steep heavy rare-earth patterns and the absence of plagioclase would explain the lack of Eu-anomalies. Positive initial ϵ_{Nd} values of many of these gneisses, particularly the oldest of them, indicates they formed directly from the mantle, or more likely, by partial melting of basalts that had them-

^{*} Tonalite, or quartz diorite, is a plutonic rock of more or less andesitic composition, having a predominance of plagioclase over orthoclase (i.e., K-poor diorite).

⁺ Trondhjemite is a plutonic rock of more or less granitic composition, but is poorer in orthoclase and richer in plagioclase than granite *senso stricto* (i.e., K-poor granite).

Geochemistry

CHAPTER 12: THE CRUST

selves formed only shortly before. The oldest preserved crust (>3.5 Ga) invariably consists of high-grade gneiss terranes.

A number of tectonic models have been suggested for the origin of these two terranes. The greenstones themselves are clearly of mantle derivation, though there is evidence in a few cases of crustal contamination. Hypotheses for their origin include continental rifts, island-arcs or continental margins, back-arc environments, and oceanic plateaus formed over mantle plumes. Many have argued that the presence of komatiites precludes formation of greenstone belts in island arcs. Because komatiites require extremely high temperatures for their generation, the relatively cool subduction zone environment seems an unlikely setting for their production. It seems more likely they formed over a zone of mantle upwelling, such as a mid-ocean ridge or mantle plume. In the Abitibi Belt, which formed between 2730 Ma and 2690 Ma, the volcanics lack Nb-Ta anomalies characteristic of subduction related magmas. This, and the rapid production of great volumes of magma, led Vervoort et al. (1993) to argue this belt originally formed as an oceanic plateau produced by a mantle plume. The granitic intrusions, which are generally slightly younger, however, do possess Nb-Ta anomalies, and hence may have formed near a subduction zone. Vervoort et al. (1993) suggested these granites were generated over a subduction zone and intruded the greenstone belt as it was accreted to the growing Superior craton. Early Proterozoic greenstone belts of the West African Craton may also represent accreted oceanic plateaus (Abouchami et al., 1990). Volcanics of other greenstone belts do, however, have Nb-Ta anomalies, suggesting they formed in an island arc setting (Condie, 1989).

The origin of the high-grade gneiss terranes is perhaps more problematic. Some have argued that the TTG suites are mantle melts. While it is possible to create siliceous melts of peridotite under conditions of shallow depth and high water pressure, there are a number of problems with this idea, and it seems more likely the TTG suite represents melts of a basaltic precursor (Taylor and McLennan, 1985; Ellam and Hawkesworth, 1988; Rudnick, 1995). There are two basic variants of this idea. First, they could represent melts of basaltic lower crust. The basaltic precursor could have formed in any number of tectonic environments. Alternatively, they could represent melts of subducted oceanic crust (Martin, 1986; Drummond and Defant, 1990). Higher temperatures and more rapid plate movements in the Archean may have meant that melting of subducting oceanic crust, which is rare today, was common then. In either case, melting must have occurred under sufficient pressure that the basaltic source was metamorphosed to amphibolite or eclogite.

Refining the Continental Crust

We can summarize by saying that the continental crust has been created by partial melting of the mantle. Subduction-related volcanism appears to have been the principal environment in which new crust has formed. Rudnick (1995) estimates that 65 to 90% of the crust has been produced in this manner. Archean crust is compositionally distinct. This may reflect a greater proportion of crust produced by some other mechanism (e.g., melting of mantle plumes), a difference in the composition of subduction-related magmas due to higher temperatures, or both.

One serious problem remains, however, namely that the composition of the continental crust does not match that of a *primary* mantle melt. Primary mantle melts are, with rare exceptions, basaltic. If the continental crust were simply a melt of the mantle, it should have a basaltic composition, as does the oceanic crust. Compared to basalt, which is the principal melting product of the mantle, it is too rich in SiO₂, too poor in MgO and CaO (and probably FeO), has too low a Mg# and too high a Rb/Sr ratio, and is generally too rich in incompatible elements. This is true irrespective of the environment in which these melts are produced. Compared to typical island arc basalts, its Al_2O_3 is also too low, as is its Sr/Nd ratio. As we stated earlier, basalt, not andesite, is the primary melt produced in subduction zones, and intra-oceanic arcs have basaltic bulk compositions. Andesites, while abundant in island arcs and even dominant in continental margin settings such as the Andes, are products of intracrustal fractional crystallization of basaltic parents, which in the case of continental margins, is often accompanied by assimilation of pre-existing crust (e.g., Hawkesworth et al., 1982). Furthermore, andesites produced in this way have concentrations of highly compatible elements, such as N i

Geochemistry

CHAPTER 12: THE CRUST

and Cr, that are much lower than those of the continental crust. Similarly, mantle-derived magmas produced at rifts or over mantle plumes are predominantly basaltic. While some andesitic magmas can be produced by melting of mantle peridotite or by reaction of basaltic magma with peridotite (e.g., boninites), such melts are rare and require very special conditions for their generation. Furthermore, their compositions do not match that of continental crust. These observations have led to the suggestion that the continental crust has been "refined" in some way; that is, some other process or processes have operated to transform the composition of the crust.

One possible explanation for this problem is that lower continental crust, consisting of mafic cumulates or partial melting residues, detaches, or "delaminates", along with underlying subcontinental mantle lithosphere (Kay and Kay, 1991). The idea is that when the crust is thickened the 50 km or more in collisional zones such as the Andes or the Himalayas, mafic lower crust will be transformed to garnet granulite and eclogite, greatly increasing its density. Under these circumstances, it may simply detach from the overlying crust and be swept away in asthenospheric circulation along with the underlying mantle lithosphere. The lithosphere would be replaced by hotter, lower density asthenosphere, which would produce uplift and basaltic volcanism. Kay and Kay argued that precisely these phenomena are observed in the Puna of Argentina, where the crust was thickened to as much as 65 km in the Miocene. As we noted in the previous chapter, McKenzie and O'Nions (1983) suggested that detached subcontinental lithosphere could later form the source of mantle plumes.

Since the lower continental crust is clearly more mafic and poorer in incompatible elements than the bulk crust (compare Tables 12.6 and 12.7), loss of lower crust to the mantle clearly provides a mechanism for adjusting the composition of the crust in the desired direction. On the other hand, the great age of many xenoliths from the subcontinental lithosphere attests to it stability (it would be difficult to detach the lower crust without also detaching the underlying mantle lithosphere). As we noted in Chapter 11, peridotite xenoliths derived from the mantle lithosphere often have the same age as the overlying crust (however, none of these ancient xenoliths have been found in areas where lithospheric detachment is thought to have occurred). Furthermore, the distinctive Os–Nd isotope systematics of subcontinental lithosphere have not been found in any mantle plumes yet. Nevertheless, detachment of the lower crust is a viable and interesting hypothesis, and one that deserves further study.

Yet another possibility is that the composition of the crust has been refined through erosion and subduction of weathering products. During weathering, certain elements, most notably Mg, preferentially go into solution. Mg is carried to the oceans where it is removed from solution by reaction with the basaltic oceanic crust during hydrothermal activity at mid-ocean ridges (discussed in Chapter 15). Depending on the fluxes assumed, it would appear that this process has the potential for removing substantial amounts of Mg from the crust and hence make it less mafic (see Problem 12.8 at the end of the chapter). In the same way, other soluble elements, such as Ca and Na, would be removed from the crust.

While this process probably contributes to refining the composition of the crust, its effect is probably quite limited. If it were important, we would expect the major element chemistry of the crust to be enriched in the least soluble elements (Al, Si, and Fe) and systematically depleted in the most soluble ones (Na, Mg, Ca) compared to typical igneous rocks. This is not the case. Among trace elements, it would not increase the extent of LRE enrichment, nor is it clear that this mechanism could increase the Rb/Sr ratio of the crust as required (though it would decrease Sr/Nd).

O'Nions and McKenzie (1988) pointed out that the incompatible element abundances in the continental crust resemble that of a very small (~1%) degree melt of the mantle. They argued that such small degree melts will be generated during episodes of continental extension. The melts would migrate into the subcontinental lithospheric mantle or lower crust and freeze there. Subsequent melting of this lower crust would produce strongly incompatible element enriched granites, which are common components of the upper crust. Interestingly, if this mechanism does indeed have a significant influence on the incompatible element budget of the continents, it would mean that Sm-Nd model ages significantly underestimate the age of the continents. Relatively little testing of this interesting hy-

Geochemistry

CHAPTER 12: THE CRUST

pothesis has been carried out, so it is difficult to judge the importance of this process. Even if the incompatible element abundances in the crust can be explained this way, they hypothesis does not address the problems with major element composition listed above.

References and Suggestions for Further Reading

- Abouchami, W., M. Boher, A. Michard and F. Albarède. 1990. A major 2.1 Ga event of mafic magmatism in West Africa — an early stage of crustal accretion. *J. Geophys. Res.* 95: 17605-17629.
- Armstrong, R. L. 1968. A model for the evolution of strontium and lead isotopes in a dynamic Earth. *Rev. Geophys.* 6: 175-199.
- Armstrong, R. L. and J. A. Cooper. 1971. Lead isotopes in island arcs. Earth Planet. Sci. Lett. 35: 27-37.
- Armstrong, R. L.1981a. Radiogenic isotopes: the case for crustal recycling on a near-steady-state nocontinental-growth Earth. in *The Origin and Evolution of the Earth's Continental Crust,* ed. S. M. and B. F. Windley. 259-287. London: The Royal Society.
- Armstrong, R. L. 1981b. Comment on "Crustal growth and mantle evolution: inferences from models of element transport and Nd and Sr isotopes". *Geochim. Cosmochim. Acta*. 45: 1251.
- Bennett, V. C. and D. J. DePaolo, Proterozoic crustal history of the western United States as determined by neodymium isotope mapping. *Bull Geol. Soc. Amer.*, 99: 674-685, 1987.
- Bowring, S. A., I. S. Williams, and W. Compston, 1989. 3.96 Ga gneisses from the Slave province, Northwest Territories, Canada. *Geology*, 17: 971-975.
- Bowring, S. A., J. E. King, T. B. Housh, C. E. Isachsen, and F. A. Podosek, 1989. Neodymium and lead isotopic evidence for enriched early Archaean crust in North America. *Nature*, 340: 222-225.
- Brophy, J. G. and B. D. Marsh. 1986. On the origin of high-alumina arc basalt and the mechanics of melt extraction. *J. Petrol.* 27: 763-789.
- Carmichael, I. S. E., F. J. Turner and J. Verhoogen. 1973. Igneous Petrology. New York: McGraw-Hill.
- Clarke, F. W. 1924. The Data of Geochemistry. U. S. Geol. Surv. Bull. 770:
- Condie, K. C. 1989. Plate Tectonics and Crustal Evolution. Oxford: Pergamon.
- Defant, M. J. and M. S. Drummond. 1990. Derivation of some modern arc magmas by melting of young subducted lithosphere. *Nature*. 347: 662-665.
- DePaolo, D. J. 1980. Crustal growth and mantle evolution: inferences from models of element transport and Nd and Sr isotopes. *Geochim Cosmochim Acta*. 44: 1185-1196.
- DePaolo, D. J. 1981a. Trace element and isotopic effects of combined wallrock assimilation and fractional crystallization, *Earth. Planet. Sci. Lett.*, 53: 189-202.
- DePaolo, D. J. 1981b. Crustal growth and mantle evolution: inferences from models of element transport and Nd and Sr isotopes (reply to a comment by R. L. Armstrong). *Geochim. Cosmochim. Acta*. 45: 1253-1254.
- DePaolo, D. J. and G. J. Wasserburg. 1979. Sm-Nd age of the Stillwater complex and the mantle evolution curve for neodymium. *Geochim Cosmochim Acta*. 43: 999-1008.
- Drummond, M. S. and M. J. Defant. 1990. A model for trondhjemite-tonalite-dacite genesis and crustal growth via slab melting: Archean to modern comparisons. *J. Geophys. Res.* 95: 21503-21521.
- Eade, K. E. and W. F. Fahirg, Chemical evolution trends of continental plates—a preliminary study of the Canadian Shield, *Geol. Serv. Canada Bull.*, 179, 1971.
- Ellam, R. M. and C. J. Hawkesworth. 1988. Is average continental crust generated at subduction zones? *Geology*. 16: 314-317.
- Faure, G., 1986. Principles of Isotope Geology, 2nd ed., Wiley & Sons, New York, 589p.
- Frounde, D. O., T. R. Ireland, P. D. Kinney, R. S. Williams, W. Compston, A. R. Williams, and J. S. Myers, 1983. Ion microprobe identification of 4,100-4,200 Myr-old detrital zircons. *Nature*, 304: 616-618.
- Fyfe, W. S. 1978. Evolution of the Earth's crust: modern plate tectonics to ancient hot spot tectonics? *Chem. Geol.* 23: 89.
- Gill, J. B. 1981. Orogenic Andesites and Plate Tectonics. Berlin: Springer Verlag.

Geochemistry

CHAPTER 12: THE CRUST

- Goldschmidt, V. M. 1933. Grundlagen der quantitativen Geochemie. Fortschr. Mineral. Kristall. Petrolgr. 17: 112-156.
- Goldstein, S. J., R. K. O'Nions and P. J. Hamilton, A Sm-Nd isotopic study of atmospheric dusts and particulates from major river systems. *Earth Planet. Sci. Lett.*, 70, 221-236, 1984.
- Hawkesworth, C. J., M. Hammill, A. R. Gledhill, P. van Calsteren and G. Rogers. 1982. Isotope and trace element evidence for late-stage intra-crustal melting in the High Andes. *Earth Planet. Sci. Lett.* 58: 240-254.
- Hawkesworth, C. J., J. M. Hergt, R. M. Ellam and F. McDermott. 1991. Element fluxes associated with subduction related magmatism. *Phil. Trans. R. Soc. Lond. A.* 335: 393-405.
- Hofmann, A. W. 1988. Chemical differentiation of the Earth: the relationship between mantle, continental crust, and oceanic crust. *Earth Plan Sci Lett*. 90: 297-314.
- Hurley, P. M., and J. R. Rand, Pre-drift continental nuclei, *Science*, 164, 1229-1242, 1969.Gill, J. B., 1981, *Orogenic Andesites and Plate Tectonics*, Springer-Verlag, Berlin, 390 p.
- James, D. E., 1981. The combined use of oxygen and radiogenic isotopes as indicators of crustal contamination, *Ann. Rev. Earth. Planet. Sci.*, 9: 311-344.
- Karig, D. E. and R. W. Kay, Fate of sediments on the descending plate at convergent margins, *Phil. Trans. R. Soc. Lond. Ser. A*, 301, 233-251, 1981.
- Kay, R. W. 1978. Aleutian magnesian andesites: metls from subducted Pacific Ocean crust. J. Volcanol. *Geotherm. Res.* 4: 117-132.
- Kay, R. W. and S. M. Kay. 1991. Creation and destruction of lower continental crust. *Geol. Rund.* 80: 259-278.
- Kay, S. M. and R. W. Kay. 1994. Aleutian magmas in space and time. In *The Geology of Alaska, The Geology of North America Vol. G-1*, G. Plafker and H. C. Berg. ed., pp. 687-722. Boulder: Geol. Soc. Am.
- Kay, S. M., R. W. Kay, and G. P. Citron. 1983. Tectonic controls on tholeiitic and calc-alkaline magmatism. J. Geophys. Res. 87: 4051-4072.
- Klein, E. M. and C. H. Langmuir. 1987. Ocean ridge basalt chemistry, axial depth, crustal thickness and temperature variations in the mantle. *J. Geophys. Res.* 92, 8089-8115.
- Klein, E. M. and C. H. Langmuir. 1989. Local versus global variations in ocean ridge basalt composition: a reply. J. Geophys. Res. 94: 4241-4252.
- Langmuir, C. H., R. D. Vocke, G. N. Hanson, and S. R. Hart, 1978. A general mixing equation with applications to Icelandic basalts, *Earth Planet. Sci. Lett.*, 37: 380-392.
- Martin, H. 1986. Effect of steeper Archean geothermal gradient on geochemistry of subduction-zone magmas. *Geology*. 14: 753-756.
- McCulloch, M. T. 1987. Sm-Nd isotopic constraints on the evolution of Precambrian crust in the Australian continent. In *Proterozoic Lithosphere Evolution, Geodyn. Ser. Vol.* 17, A. Kröner. ed., pp. 115-130. Washington: AGU.
- McCulloch, M. T. and V. C. Bennett. 1994. Progressive growth of the Earth's continental crust and depleted mantle: geochemical constraints. *Geochim. Cosmochim. Acta*. 58: 4717-4738.
- McKenzie, D. P. and R. K. O'Nions. 1983. Mantle reservoirs and ocean island basalts. *Nature*. 301: 229-231.
- Nyblade, A. A. and H. N. Pollack. 1993. A global analysis of heat flow from Precambrian terrains: implications for the thermal structure of Archean and Proterozoic lithosphere. *J. Geophys. Res.* 98: 12207-12218.
- O'Nions, R. K. and P. J. Hamilton. 1981. Isotope and trace element models of crustal evolution. *Phil. Trans. Roy. Soc. Lond. A.* 301: 473-487.
- O'Nions, R. K. and D. P. McKenzie. 1988. Melting and continent generation. *Earth Plan Sci Lett*. 90: 449-456.
- Patchett, P. J. 1980. Thermal effects of basalt on continental crust and crustal contamination of magmas. *Nature*. 283, 559-561.

Geochemistry

CHAPTER 12: THE CRUST

- Patchett, P. J. and N. T. Arndt. 1986. Nd isotopes and tectonics of 1.9-1.7 Ga crustal genesis. Earth Planet. Sci. Lett. 78: 329-338.
- Perfit, M. R., D. A. Gust, A. E. Bence, R. J. Arculus and S. R. Taylor. 1980. Chemical characteristics of island-arc basalts: implications for mantle sources. Chem. Geol. 30: 227-256.
- Plank, T. and C. H. Langmuir. 1988. An evaluation of the global variations in the major element chemistry of arc basalts. Earth Planet. Sci. Lett. 90, 349-370.
- Plank, T. and C. H. Langmuir. 1993. Tracing trace elements from sediment input ot volcanic output at subduction zones. Nature. 362: 739-742.
- Plank, T. and C. H. Langmuir. in press. The geochemical composition of subducting sediment and its consequences for the crust and mantle. Chem. Geol.
- Reymer, A. and G. Schubert. 1984. Phanerozoic addition rates to the cotinental crust and crustal growth. Tectonics. 3: 63-77.
- Reymer, A. and G. Schubert. 1986. Rapid growth of some major segments of continental crust. *Geology*. 14: 299-302.
- Roeder, P. L. and R. F. Emslie. 1970. Olivine-liquid equilibrium. Contrib. Mineral. Petrol. 29: 275-289. Rudnick, R. 1995. Making continental crust. Nature. 378: 571-578.
- Rudnick, R. L. and D. M. Fountain. 1995. Nature and composition of the continental crust: a lower crustal perspective. Rev. Geophys. 33: 267-309.
- Shaw, D. M. 1967. An estimate of of the chemical composition of the Canadian Shield, Can. J. Earth Sci. 4, 82.
- Shaw, D. M., J. J. Gramer, M. D. Higgins and M. G. Truscott. 1986. Composition of the Candadian Precambrian shield and the continental crust of the earth. In The Nature of the Lower Continental Crust, Vol. J. B. Dawson, C. D. A., J. Hall and K. H. Wedepohl. ed., pp. 275-282. Oxford: Blackwell Scientific Publ.
- Sinton, J. M. and R. Detrick. 1992. Mid-ocean ridge magma chambers, J. Geophys. Res. 97, 197-216.
- Stolper, E. and S. Newman. 1994. the role of water in the petrogenesis of Mariana trough basalts. Earth Planet. Sci. Lett. 121: 293-325.
- Sun, S. S. 1980. Lead isotope study of young volcanic rocks from mid-ocean ridges, oceanic islands, and island arcs, Philos. Trans. R. Soc. London, Ser. A, 297: 409-445.
- Taylor, H. P. 1980. The effects of assimilation of country rocks by magmas on ¹⁸O/¹⁶O and ⁸⁷Sr/⁸⁶Sr systematics in igneous rocks. Earth Planet. Sci. Lett. 47, 243-254.
- Taylor, S. R. 1967. The origin and growth of continents, *Tectonophysics*. 4, 17-34.
- Taylor, S. R. and S. M. McLennan. 1985. The Continental Crust: its composition and evolution. Oxford: Blackwell Scientific Publishers.
- Taylor, S. R. and S. M. McLennan. 1995. The geochemical evolution of the continental crust. Rev. Geophys. 33: 241-265.
- Tera, F., L. Brown, J. Morris, I. S. Sacks, J. Klein, R. Klein and R. Middleton. 1986. sediment incorporation in island-arc magmas: inferences from ¹⁰Be, Geochim. Cosmochim. Acta. 50: 535-550.
- Thorpe, R.S. (ed.). 1982. Orogenic Andesites and Related Rocks. Chichester: Wiley & Sons.
- Veizer, J. and S. L. Jansen. 1979. Basement and sedimentary recycling and continental evolution. J. Geol. 87: 341-370.
- Vervoort, J. D., W. M. White, R. I. Thorpe and J. M. Franklin. 1993. Postmagmatic thermal activity in the Abitibi Greenstone Belt, Noranda and Matagami Districts: evidence from whole rock Pb isotope data. Econ. Geol. 88: 1598-1614.
- Vitorello, I. and H. N. Pollack, On the variation of continental heat flow with age and the thermal evolution of continents, J. Geophys. Res., 85, 983-995, 1980.
- Weaver, B. L. and J. Tarney. 1984. Major and trace element composition of the continental lithosphere. In Structure and Evolution of the Continental Lithosphere, Physics and Chemistry of the Earth, vol 15, Vol. H. N. Pollack and V. R. Murthy. ed., pp. 39-68. Oxford: Pergamon Press.

CHAPTER 12: THE CRUST

- Wedepohl, K. H. 1995. The composition of the continental crust. *Geochim. Cosmochim. Acta*. 59: 1217-1232.
- White, W. M. and B. Dupré, Sediment subduction and magma genesis in the Lesser Antilles: Isotopic and trace element constraints, *J. Geophys. Res.*, 91, 5927-5941, 1986.
- White, W. M. and P. J. Patchett, 1984, Hf-Nd-Sr isotopes and incompatible element abundances in island arcs: implications for magma origins and crust-mantle evolution, *Earth Planet. Sci. Lett.*, 67: 167-185.

Problems

1. The adjacent table shows trace element data from three volcanic rocks, each erupted in a different tectonic environment. Plot the data on a "spider diagram". Make the y-axis logarithmic and order the elements as in Figure 12.32. "Normalize" the data to "Primitive Mantle" using the concentrations given in Table 10.3. Based on your plot, identify the tectonic setting in which each erupted.

2. In a sedimentary core, the surface sample had a 10 Be concentration of 100×10^6 atoms/g. Assuming steady state accumulation of 10 Be at the surface, what concentration of 10 B would you predict for a 10 million year old sample from lower in the same core? (*HINT: use the decay constant given in Table 7.5*).

3. The following Sr and Nd isotope data from St. Lucia show that these volcanics are mixtures of two components. Assuming that component A has ⁸⁷Sr/⁸⁶Sr, ε_{Nd} , Sr, and Nd concentrations of 0.706, -0.5, 200 and 8.5 ppm respectively, and component B has ⁸⁷Sr/⁸⁶Sr and ε_{Nd} of 0.710 and -13.4, estimate the Sr and Nd concentrations of component B. (*HINT: Plot the data and superimpose a 2 component mixing model on the plot; adjust the concentrations by trial and error until you find the best fit*).

4. The adjacent table below shows concentrations of Sr and ⁸⁷Sr/⁸⁶Sr measured on river samples. These samples were taken just below the point where two major rivers join. Are data consistent with simple mixing between the waters of these two rivers? Why or why not.

5. Plot oxygen and strontium isotope ratios measured on lavas from Galeras Volcano (Colombia) in the following table, then decide

whether mixing curve they define a result of sediment subduction ("source contamination") or assimilation ("crustal contamination")? Justify your answer.

	А	В	С
Cs		1.07	0.015
Rb	34.00	17.96	1.33
Ва	389.47	353.83	12.30
Th	7.80	14.50	0.29
U	1.89	4.99	0.11
Nb	45.67	9.65	5.93
Та	2.51	0.45	0.35
Κ	25921.00	5130.00	1160.00
La	42.78	31.38	6.93
Ce	93.76	66.73	23.10
Pb	4.14	7.86	0.68
Pr	11.96	7.41	4.31
Sr	717.87	798.75	92.60
Nd	48.20	28.21	22.80
Sm	10.36	5.29	7.89
Zr	300.91	121.92	227.00
Ηf	6.97	3.86	7.15
Eu	3.06	1.45	2.50
Gd	8.90	4.83	10.90
Tb	1.36	0.69	
Dy	6.37	3.51	13.00
Y	30.34	19.73	69.10
Но	1.06	0.68	2.89
Er	2.33	1.71	8.46
Tm	0.30	0.27	

1.63

0.22

SR AND NO ISOTOPE DATA FROM ST. LUCIA

TROWI ST. LUCIA				
Sample	⁸⁷ Sr/ ⁸⁶ Sr	ε _{Nd}	DATA FOR	PRODLEM 4
STL220	0.70655	-3.02	Sr, μg/l	⁸⁷ Sr/ ⁸⁶ Sr
SL124	0.70933	-11.76	45	0.71571
SL125	0.70782	-8.74	60	0.71305
SL51	0.70845	-10.26	78	0.71162
SL52	0.70850	-10.09	100	0.71073
SL121	0.70907	-11.26	123	0.71013
SL48	0.70885	-10.90	136	0.70969
SL42	0.70829	-9.71	161	0.70935
SL40	0.70837	-9.99	179	0.70909
			_, ,	00707

Yb

Lu

7.94

1.19

1.58

0.23

Geochemistry

DATA FOR PRODLEM 12.1

Geochemistry

CHAPTER 12: THE CRUST

DATA FROM GALERAS VOLCAN

⁸⁷ Sr/ ⁸⁰ Sr	δ ¹⁰ O	Sr, ppm
0.70414	6.79	503
0.70409	6.65	504
0.70453	7.65	545
0.70423	6.99	521
0.70434	7.51	521
0.70435	7.11	534
0.70419	6.72	514

6. Construct an AFC model of the data from Galeras Volcano in the accompanying table. Assume that the ⁸⁷Sr/⁸⁶Sr of the assimilant and original magma are 0.712 and 0.70375, respectively, the δ^{18} O of the assimilant and original magma are +10.5 and +5.8, respectively, and Sr concentrations of the assimilant and original magma were 100 and 470 ppm respectively. Find values of R (ratio of mass assimilated to mass crystallized) and D_{Sr} (bulk partition coefficient for Sr) that best fit the data (assume the partition coefficient for O is 1). (*HINT: plot both* ⁸⁷Sr/⁸⁶Sr vs. Sr and δ^{18} O vs ⁸⁷Sr/⁸⁶Sr, and adjust values of R and D until the line describing your model

passes through the data).

7. (a.) Assuming the Bulk Silicate Earth (BSE) has the "primitive mantle" Rb concentration given in Table 10.3, what fraction of the Rb in the BSE is in the crust according to the Taylor and McLennan, Rudnick and Fountain, and Wedepohl crustal compositions (Table 12.7)? (*HINT: use the volumes of crust and mantle given in the Appendix*).

(b.) Assume that the silicate Earth consists of only three reservoirs: continental crust, depleted mantle and primitive mantle, that the depleted mantle has a Rb concentration that is 10 times lower than the average MORB concentration (Table 12.2), and that the primitive mantle has the Rb concentration given in Table 10.3. What fraction of the mantle would consist of depleted mantle when you use the Taylor and McLennan, Rudnick and Fountain, and Wedepohl crustal Rb concentrations (Table 12.7)?

8. Weathering and erosion remove Mg from the continents and transport it in dissolved form to the oceans. Dissolved Mg is removed from the oceans by hydrothermal activity at mid-ocean ridges. Assuming that the flux of Mg into the oceanic crust given in Chapter 15 has been constant through time over the past 2.5 Ga and that all this Mg is derived from the continental crust, how much has the crustal Mg concentration decreased as a result of this process? How might this result change if we make other "reasonable" assumptions about the flux of Mg into the oceanic crust through time?

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

Chapter 13: Reactions at the Earth's Surface: Weathering, Soils, and Stream Chemistry

INTRODUCTION

he geochemistry of the Earth's surface is dominated by aqueous solutions and their interactions with rock. We saw in the last chapter that the upper continental crust has the approximate average composition of granodiorite, and that the oceanic crust consists of basalt. But a random sample of rock from the crust is unlikely to be either; indeed it many not be an igneous rock at all. At the very surface of the Earth, sediments and soils predominate. Both are ultimately produced by the interaction of water with "crystalline rock" (by which we mean igneous and metamorphic rocks). Clearly, to fully understand the evolution of the Earth, we need to understand the role of geochemical processes involving water.

Beyond that, water is essential to life and central to human activity. We use water for drinking, cooking, agriculture, heating, cooling, resource recovery, industrial processing, waste disposal, transportation, fisheries, etc. Water chemistry, i.e., the nature of solutes dissolved in it, is the primary factor in the suitability of water for human use. "Polluted" water is unsuitable for drinking and cooking; saline water is unsuitable for these uses as well as agriculture and many industrial uses, etc. We have been particularly concerned with water pollution in the past few decades; that is with the impact of human activity on water chemistry. Both our advancing technology and our exponentially increasing numbers have made pollution problems progressively worse, particularly over the past century. However, we have also become more aware of the adverse impact of poor water quality on human health and the quality of life, and perhaps less tolerant of it as well.

Understanding and addressing problems of water pollution requires an understanding of the behavior of natural aqueous systems for at least two reasons. First, to identify pollution, we need to know the characteristics of natural systems. For example, Pb can be highly toxic, and high concentrations of Pb in the blood have been associated with learning disabilities and other serious problems. However, essentially all waters have some finite concentration of Pb; we should be concerned only when Pb concentrations exceed natural levels. Second, natural processes affect pollutants in the same way they affect their natural counterparts. For example, cadmium leached from landfills will be subject to the same adsorption/desorption reactions as natural Cd. To predict the fate of pollutants, we need to understand those processes.

In this chapter, we focus on water and its interaction with solids at the Earth's surface. We can broadly distinguish two kinds of aqueous solutions: continental waters and seawater. Continental waters by this definition include ground water, fresh surface waters (river, stream and lake waters), and saline lake waters. The compositions of these fluids are obviously quite diverse. Seawater, on the other hand, is reasonably uniform, and it is by far the dominant fluid on the earth's surface. Hy-drothermal fluids are third class of water produced when water is heated and undergoes accelerated interactions with rock and often carry a much higher concentration of dissolved constituents than fresh water. Our focus in this chapter will be on the chemistry of continental waters and how they interact with rock. We we consider seawater in Chapter 15.

Redox in Natural Waters

The surface of the Earth represents a boundary between regions of very different redox state. The atmosphere contains free oxygen and therefore is highly oxidizing. In the Earth's interior, however, there is no free oxygen, Fe is almost entirely in the 2+ valance state, reduced species such as CH_4 , CO, and S_2 exist, and conditions are quite reducing. Natural waters exist in this boundary region and their redox state, perhaps not surprisingly, is highly variable. Biological activity is the principal cause of this variability. Plants (autotrophs) use solar energy to drive thermodynamically unfavorable

Geochemistry

CHAPTER 17: WEATHERING, SOILS, AND STREAM CHEMISTRY

photosynthetic reactions that produce free O_{2} , the ultimate oxidant, on the one hand and organic matter, the ultimate reductant, on the other. Indeed it is photosynthesis that is responsible for the oxidizing nature of the atmosphere and the redox imbalance between the Earth's exterior and interior. Both plants and animals (heterotrophs) liberate stored chemical energy by catalyzing the oxidation of organic matter in a process called *respiration*. The redox state of solutions and solids at the Earth's surface is largely governed by the balance between photosynthesis and respiration. By this we mean that most waters are in a fairly oxidized state because of photosynthesis and exchange with the atmosphere. When they become reducing, it is most often because respiration exceeds photosynthesis and they have been isolated from the atmosphere. Water may also become reducing as a result of reaction with sediments deposited in ancient reduced environments, but the reducing nature of those ancient environments resulted from biological activity. Weathering of reduced primary igneous rocks also consumes oxygen, and this process governs the redox state of some systems, mid-ocean ridge hydrothermal solutions for example. On a global scale, however, these processes are of secondary importance for the redox state of natural waters.

The predominant participants in redox cycles are C, O, N, S, Fe, and Mn. There are a number of other elements, for example, Cr, V, As, and Ce, that have variable redox states; these elements, however, are always present in trace quantities and their valance states reflect, rather than control, the redox state of the system. Although phosphorus has only one valance state (+V) under natural conditions, its concentration in solution is closely linked to redox state because the biological reactions that control redox state also control phosphorus concentration, and because it is so readily adsorbed on Fe oxide surfaces.

Water in equilibrium with atmospheric oxygen has a pc of +13.6 (at pH = 7). At this pc, thermodynamics tells us that all carbon should be present as CO2 (or related carbonate species), all nitrogen as NO₃, all S as SO₄⁻¹, all Fe as Fe³⁺, and all Mn as Mn⁴⁺. This is clearly not the case and this disequilibrium reflects the kinetic sluggishness of many, though not all, redox reactions^f. Given the dise-

quilibrium we observe, the applicability of thermodynamics to redox systems would appear to be limited. Thermodynamics may nevertheless be used to develop partial equilibrium models. In such models, we can make use of redox couples that might reasonably be at equilibrium to describe the redox state of the system. In Chapter 3, we introduced the tools needed to deal with redox reactions: E_{H} , the hydrogen scale potential (the potential developed in a standard hydrogen electrode cell) and pε, or electron activity. We found that both may in turn

	Reaction	pε°	$p\epsilon_w$
1	$\frac{1}{4}O_{2(g)} + H^+ + e^- \rightleftharpoons \frac{1}{2}H_2O$	+20.75	+13.75
2	$\frac{1}{5}$ NO $\frac{1}{3}$ + $\frac{6}{5}$ H ⁺ + e ⁻ $\rightleftharpoons \frac{1}{10}$ N _{2(g)} + $\frac{3}{5}$ H ₂ O	+21.05	+12.65
3	$\frac{1}{2}MnO_{2(s)} + 2H^{+} + e^{-} \rightleftharpoons \frac{1}{2}Mn^{2+} + H_2O$	+20.8	+9.8 ⁺
4	$\frac{5}{4} \text{NO}_{-3}^{-} + \frac{6}{5} \text{H}^{+} + \text{e}^{-} \rightleftharpoons \frac{1}{2} \text{NH}_{+}^{+} + \frac{3}{8} \text{H}_2\text{O}$	+14.9	+6.15
5	$Fe(OH)_{3(s)} + 3H^+ + e^- \rightleftharpoons Fe^{2+} + 3H_2O$	+16.0	$+1.0^{+}$
6	$\frac{1}{2}$ CH ₂ O* + H ⁺ + e ⁻ $\rightleftharpoons \frac{1}{2}$ CH ₃ OH	+4.01	-3.01
7	$\frac{1}{8}SO_{4}^{2-} + \frac{5}{4}H^{+} + e^{-} \rightleftharpoons \frac{1}{8}H_{2}S + \frac{1}{2}H_{2}O$	+5.25	-3.5
8	$\frac{1}{8}SO_4^{2-} + \frac{9}{8}H^+ + e^- \rightleftharpoons \frac{1}{8}HS^- + \frac{1}{2}H_2O$	+4.25	-3.6
9	$\frac{1}{8}CO_{2(g)} + H^+ + e^- \rightleftharpoons \frac{1}{8}CH_{4(g)} + \frac{1}{4}H_2O$	+2.9	-4.1
10	$\frac{1}{6}N_{2(g)} + \frac{4}{3}H^{+} + e^{-} \rightleftharpoons \frac{1}{3}NH \frac{4}{4}$	+4.65	-4.7
11	$\frac{1}{4}CO_{2(g)} + H^+ + e^- \rightleftharpoons \frac{1}{4}CH_2O^* + \frac{1}{4}H_2O$	-0.2	-7.2

|--|

+ The concentration of Mn^{2+} and Fe^{2+} are set to 1 μM .

be related to the Gibbs Free * We are using " CH_2O ", which is formally formaldahyde, as an abbreviation for Energy of reaction through organic matter generally (for example, glucose is $C_6H_{12}O_6$).

f While this may make life difficult for geochemists, it is also what makes it possible in the first place. We, like all other organisms, consist of a collection of reduced organic species that manage to persist in an oxidizing environment!

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

the Nernst Equation (equ. 3.121). These are all the tools we need; in this section, we will see how we can apply them to understanding redox in aqueous systems.

Table 13.1 lists the $p\epsilon^{\circ}$ of the most important redox half reactions in aqueous systems. Also listed are $p\epsilon_{W}$ values. $p\epsilon_{W}$ is the $p\epsilon^{\circ}$ when the concentration of H⁺ is set to 10^{-7} (pH = 7). The relation between $p\epsilon^{\circ}$ and $p\epsilon_{W}$ is simply:

$$p\varepsilon_w = p\varepsilon^\circ + \log [H^+]^v = p\varepsilon^\circ - v \times 7$$

Reactions are ordered by decreasing p_{E_W} from strong oxidants at the top to strong reductants at the bottom. In this order, each reactant can oxidize any product below it in the list, but not above it. Thus sulfate can oxidize methane to CO_2 , but not ferrous iron to ferric iron. Redox reactions in aqueous systems are often biologically mediated. In the following section, we briefly explore the role of the biota in controlling the redox state of aqueous systems.

Biogeochemical Redox Reactions

As we noted above, photosynthesis and atmospheric exchange maintains a high pe in surface waters. Water does not transmit light well, so there is an exponential decrease in light intensity with depth. As a result, photosynthesis is not possible below depths of 200 m even in the clearest waters. In murky waters, photosynthesis can be restricted to the upper few meters or less. Below this "photic zone", biologic activity and respiration continue, sustained by falling organic matter from the photic zone. In the deep waters of lakes and seas where the rate of respiration exceeds downward advection of oxygenated surface water, respiration will consume all available oxygen. Once oxygen is consumed, a variety of specialized bacteria continue to consume organic matter and respire utilizing oxidants other than oxygen. Thus pe will continue to decrease.

Since bateria exploit first the most energetically favorable reactions, Table 13.1 provides a guide the sequence in which oxidants are consumed as pe decreases. From it, we can infer that once all molecular oxygen is consumed, reduction of nitrate to molecular nitrogen will occur (reaction 2). This

processes, known as *denitrification*, is carried out by bacteria, which use the oxygen liberated to oxidize organic matter and the net energy liberated to sustain themselves. At lower levels of pɛ, other bacteria reduce nitrate to ammonia (reaction 4), a process called *nitrate reduction*, again using the oxygen liberated to oxidize organic matter. At about this pɛ level, Mn^{4+} will be reduced to Mn^{2+} . At lower pɛ, ferric iron is reduced to ferrous iron. The reduction of both Mn and Fe may also be biologically mediated in whole or in part.

From Table 13.1, we can expect that *fermenta*tion (reaction 6 in Table 13.1) will follow reduction of Fe. Fermentation can involve any of a number of reactions, only one of which, reduction of organic matter (carbohydrate) to methanol, is represented in Table 13.1. In fermentation reactions, further reduction of some of the organic carbon provides a sink of electrons, allowing oxidation of the remaining organic carbon; for example in glucose, which has 6 carbons, some are oxidized to CO_2 while others are reduced to alcohol or acetic acid. While these kinds of reactions can be carried out by many organisms, it is bacterial-me-



Figure 13.1. Important biogeochemical redox couples in natural waters.

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

diated fermenation that is of geochemical interest.

At lower pɛ, sulfate is used as the oxidant by sulfate-reducing bacteria to oxidize organic matter, and at even lower pɛ, nitrogen is reduced to ammonia (reaction 9), a process known as *nitrogen fixa-tion*, with the nitrogen serving an the electron acceptor for the oxidization of organic matter. This reaction is of great biological importance, as nitrogen is an essential ingredient of key biological compounds such as proteins and DNA (see Chapter 14), and hence is essential to life; all plants must therefore take up inorganic nitrogen. While a few plants, blue-green algae (cyanobacteria) and legumes, can utilize N_2 , most require "fixed" nitrogen (ammonium, nitrate, or nitrite). Hence nitrogen-fixing bacteria play an essential role in sustaining life on the planet.

To summarize, in a water, soil, or sediment column where downward flux of oxygen is less than the downward flux of organic matter, we would expect to see oxygen consumed first, followed by reduction of nitrate, manganese, iron, sulfur, and finally nitrogen. This sequence is illustrated on a pɛ-pH dia-gram in Figure 13.1. We would expect to see a similar sequence with depth in a column of sediment where the supply of organic matter exceeds the supply of oxygen and other oxidants.

Eutrophication

The extent to which the redox sequence described above proceeds in a body of water depends on a several factors. The first of these is temperature structure, because this governs the advection of oxygen to deep waters. As mentioned above, light (and other forms of electromagnetic energy) is not transmitted well by water. Thus only surface waters are heated by the Sun. As the temperature of surface water rises, its density decreases (fresh water reaches it greatest density at 4° C). These warmer surface waters, known in lakes as the *epilimnion*, generally overlie a zone where temperature decreases rapidly, known as the *thermocline* or *metalimnion*, and a deeper zone of cooler water, known as the *hypolimnion*. This temperature stratification produces a stable density stratification which prohibits vertical advection of water and dissolved constitutents, including oxygen and nutrients. In tropical lakes and seas, this stratification is permanent. In temperate regions, however, there is an annual cycle in which stratification develops in the spring and summer. As the surface water cools in the fall and winter, its density decreases below that of the deep water and vertical mixing occurs. The second important factor governing the extent to which reduction in deep water occurs is nutrient levels. Nutrient levels limit the amount of production of organic carbon by photosynthesizers (in lakes, phosphorus concentrations are usually limiting; in the oceans, nitrate and micronutrients such as iron appear to be limiting). The availability of organic carbon in turn controls *biological oxygen demand* (BOD) In water with high nutrient levels there is a high flux of organic carbon to deep waters and hence higher BOD.

In lakes with high nutrient levels, the temperature stratification described above can lead to a situation where dissolved oxygen is present in the epilimnion and absent in the hypolimnion. Regions where dissolved oxygen is present are termed *oxic*, those where sulfide or methane are present are called *anoxic*. Regions of intermediate pe are called *suboxic*. Lakes are where suboxic or anoxic conditions exist as a result of high biological producitivy are said to be *eutrophic*. This occurs naturally in many bodies of water, particularly in the tropics where stratification is permanent. It can also occur, however, as a direct result addition of pollutants such as sewage to the water, and an indirect result of pollutants such as phosphate and nitrate. Addition of the latter enhances productivity and availability of organic carbon, and ultimately BOD. When all oxygen is consumed conditions become anaerobic and the body of water becomes eutrophic. Where this occurs naturally, ecosystems have adapted to this circumstance and only anaerobic bacteria are found in the hyperlimnion. When it results from pollution, it can be catastrophic for macrofauna such as fish that cannot tolerate anaerobic conditions.

Geochemistry

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

Redox and Biological Primary Production

The biota is capable of oxidations as well as reductions. The most familiar of these reactions is photosynthesis. Most organisms capable of photosynthesis, which includes both higher plants and a variety of single-celled organisms, produce oxygen as a biproduct of photosynthesis:

$$\dot{CO}_2 + H_2O \rightleftharpoons CH_2O + \dot{O}_2$$
 13.1

However, there are also photosynthesis pathways that do not produce O_2 . Green and purple sulfur bacteria are phototrophs that oxidize sulfide to sulfur in the course of photosynthesis:

$$CO_2 + 2H_2S \rightleftharpoons CH_2O + \frac{1}{4}S_8 + H_2O$$
 13.2

This reaction requires considerably less light energy (77.6 vs. 476 kJ/mol) than oxygenic photosynthesis, enabling these bacteria survive at lower light levels than green plants.

While photosynthesis is far and away the primary way in which organic carbon is produced or "fixed", chemical energy rather than light energy may also be used to fix organic carbon in processes collectively known as *chemosynthesis*. In chemosynthesis, the energy liberated in oxidizing reduced inorganic species is used to reduce CO_2 to organic carbon. For example, nitrifying bacteria oxidize ammonium to nitrite in a process known as *nitrification*:

$$CO_2 + \frac{2}{3}NH_4 + \frac{1}{3}H_2O \rightleftharpoons CH_2O + \frac{2}{3}NO_2^- + \frac{4}{3}H^+$$
 13.3

Colorless sulfur bacteria oxidize sulfide to sulfate in fixing organic carbon:

$$CO_2 + H_2S + O_2 + H_2O \rightleftharpoons CH_2O + SO_4^{2-} + 2H^+$$
 13.4

Redox Buffers and Transition Metal Chemistry

The behavior of transition metals in aqueous solutions and solids in equilibrium with them is particularly dependent on redox state. Many transition metals have more than one valence state within the range of pe of water. In a number of cases, the metal is much more soluble in one valence state than in others. The best examples of this behavior are provided by iron and manganese, both of which are much more soluble in their reduced (Fe^{2+} , Mn^{2+}) than in oxidized (Fe^{3+} , Mn^{4+}) forms. Redox conditions thus influence a strong control on the concentrations of these elements in natural waters.

Because of the low solubility of their oxidized formes, the concentrations of Fe, Mn, and similar metals are quite low under "normal" conditions, i.e., high pɛ and nearneutral pH. There two common circumstances where higher Fe and Mn concentrations in water occur. The first is when sulfide ores are exposed by mining and oxidized to sulfate, e.g.:

$$2\text{FeS}_{2(s)} + 2\text{H}_2\text{O} + 7\text{O}_2 \rightleftharpoons 4\text{H}^+ + 4\text{SO}_4^{2-} + 2\text{Fe}^{2+}$$

This can dramatically lower the pH of streams draining such areas. The lower pH in turn allows higher concentrations of dissolved metals (e.g., Figure 13.2), even under oxidizing conditions. The second circumstance where higher Fe and Mn concentrations occur is under suboxic or anoxic conditions that may occur in deep waters of lakes and seas as well as sediment pore waters. Under these circumstances Fe and Mn are



Figure 13.2. Contours of dissolved Fe activity as a function of pe and pH.

Geochemistry

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

reduced to their soluble forms, allowing much higher concentrations.

In cases where precipitation or solution involves a change in valence or oxidation state, the solubility product must include pe or some other redox couple, e.g.:

$$Fe_2O_3 + 6H^+ + 2e^- \rightleftharpoons 2Fe^{2+} + 3H_2O$$
 $K = \frac{a_{Fe^{2+}}}{a_{H^+}^6 a_{e^-}^2}$ 13.5

In Chapter 3, we noted that pc is often difficult to determine. One approach to the problem is to assume the redox state in the solution is controlled by a specific reaction. The controlling redox reactions will be those involving the most abundant species; very often this is sulfate reduction:

$$\frac{1}{8}SO_{4}^{2-} + \frac{5}{4}H^{+} + e^{-} \rightleftharpoons \frac{1}{8}H_{2}S + \frac{1}{2}H_{2}O$$
 13.6

in which case pɛ is given by:

$$p\varepsilon = p\varepsilon^{\circ} - \frac{1}{8} \ln \frac{[H_2 S]}{[SO_4^{2^-}]} - \frac{4}{5} pH$$
 13.7

Under the assumption that this reaction controls the redox state of the solution, electrons may be eliminated from other redox reactions by substituting the above expression. For example, iron redox equilibrium may be written as:

$$\frac{1}{3}SO_{4}^{2-} + \frac{5}{4}H^{+} + Fe^{2+} \rightleftharpoons \frac{1}{8}H_{2}S + \frac{1}{2}H_{2}O + Fe^{3+}$$
13.8

In this sense, the pɛ of most natural waters will be controlled by a *redox buffer*, a concept we considered in Chapter 3. Example 13.1 illustrates this approach.

Example 13.1. Redox State of Lake Water

Consider water of the hypolimnion of a lake in which all oxygen has been consumed and with the following initial composition: $SO_4 = 2 \times 10^{-4} \text{ M}$, $\Sigma Fe^{3+} = 10^{-6} \text{ M}$, $Alk = 4 \times 10^{-4} \text{ eq}/\text{L}$, $\Sigma CO_2 = 1.0 \times 10^{-3} \text{ M}$, $\Sigma''CH_2O'' = 2 \times 10^{-4} \text{ M}$, pH = 6.3. Determine the pH, p ϵ , and speciation of sulfur and iron when all organic matter is consumed and redox equilibrium is achieved. The first dissociation constant of H_2S is 10^{-7} .

Answer: Let's first consider the redox reactions involved. The species involved in redox reactions will be those of carbon, sulfur, and iron. The concentration of iron is small, so its oxidation state will reflect, rather than control, that of the solution. Thus oxidation of organic matter will occur though reduction of sulfate. We can express this by combining reactions 7 and 11 in Table 13.1 (we see from the dissociation constant that H_2S will be the dominant sulfate species at pH below 7, so we chose reaction 7 rather than reaction 8):

$$\frac{1}{2}SO_{4}^{2-} + H^{+} + CH_{2}O \rightleftharpoons \frac{1}{2}H_{2}S + H_{2}O + CO_{2} \qquad K = 10^{10.54} \quad 13.9$$

From the magnitude of the equilibrium constant (obtained from the $p\epsilon^{\circ's}$ in Table 13.1), we can see that right side of this reaction is strongly favored. Since sulfate is present in excess of organic matter, this means sulfate will be reduced until all organic matter is consumed, which will leave equilmolar concentrations of sulfate and sulfide (10^4 M each). Thus the redox state of the system will be governed by that of sulfur. The redox state of iron can then be related to that of sulfur using reaction 13.8, for which we calculate an equilibrium constant of $10^{-10.75}$ from Table 13.1.

The next problem we face is that of chosing components. As usual, we chose H^+ as one component (and implicitly H_2O as another). We will also want to chose a sulfur, carbon, and iron species as a component, but which ones? We could chose the electron as a component, but consistent with our conclusion above that the redox state of the system is governed by that of sulfur, a better choice is to chose both sulfate and sulfide, specifically H_2S , as components. We can also see from Table 13.1 that Fe should be largely reduced, so we chose Fe^{2+} as the iron species. pH will be largely controlled by carbonate species, since these are more than an order of magnitude more abundant than sulfate species; oxidation of organic matter will increase the concentration of CO_2 , which will lower pH slightly. In

Geochemistry

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

Figure 6.1, we can see that at pH below 6.4, H_2CO_3 will be the dominant carbonate species, so we chose this as the carbonate species. Our components are therefore H⁺, SO₄, H_2CO_3 , Fe²⁺, and H_2S . The species of interest will include H⁺, OH⁻, SO₄²⁻, H₂S, HS⁻, H₂CO₃, HCO₃, as well as the various species of Fe (Fe²⁺, Fe³⁺, Fe(OH)²⁺, Fe(OH)⁺₂, Fe(OH)₃ (we assume that the concentrations of CO₃⁻, HSO₄⁻ and S²⁻ are neglible at this pH; we shall neglect them throughout).

Our next step is to determine pH. For *TOT*H we have:

$$TOTH = [H^+] - [OH^-] - [HCO_3^-] - [HS^-] + \frac{5}{4}\Sigma Fe^{3+}$$
 13.10

The presence of the Fe^{3+} term may at first be confusing. To understand why it occurs, we can use equation 13.7 to express Fe^{3+} as the algebraic sum of our components:

$$Fe^{3+} = \frac{1}{8}SO_4^{2-} + \frac{5}{4}H^+ + Fe^{2+} - \frac{1}{8}H_2S + \frac{1}{2}H_2O$$
13.11

The first 4 terms on the right hand side of equation 13.10 are simply alkalinity plus additional CO_2 produced by oxidation of organic matter, so 13.10 may be rewritten as:

$$TOTH = \frac{5}{4}\Sigma Fe^{3+} - Alk - [CH_2O]$$
 13.12

Inspecting equation 13.10, we see that HCO $_3$ is by far the largest term. Furthermore, the Fe term in equation 13.12 is neglible, so we have:

$$TOTH \simeq -[HCO_3] \simeq 6 \times 10^{-4}$$
 13.13

The conservation equation for carbonate is:

$$\Sigma H_2 CO_3 = [H_2 CO_3] + [HCO_3^-] = \Sigma CO_2 + \Sigma [CH_2 O] = 1.3 \times 10^{-3}$$

Hence:

$$H_2CO_3 = \Sigma H_2CO_3 - HCO_3^- = (1.3 - 0.6) \times 10^{-3}$$

We can use this to calculate pH since: $K = \left[\frac{HCO_3^-}{[H_2CO_3]}\right] = 10^{-6.35}$

Solving for $[H^+]$ and substitutiong values, we find that pH = 6.28. For the conservation equation for sulfate, we will have to include terms for both Fe³⁺ (equation 13.11) and organic matter. Writing organic matter as the algebraic sum of our components we have:

$$CH_2O = \frac{1}{2}H_2S + H_2O + CO_2 - \frac{1}{2}SO_4^{2-} - H^+$$

The amount of sulfate present will be that originally present less that used to oxidize organic matter. The only other oxidant present in the system is ferric iron, so the amount of sulfide used to oxidize organic matter will be the total organic matter less the amount of ferric iron initially present. The sulfate conservation equation is then:

$$\Sigma SO_4 = [SO_4^{2-}] - \frac{1}{2}\Sigma CH_2O + \frac{1}{8}\Sigma Fe^{3+} \simeq 1.0 \times 10^{-4} M$$

(the Fe term is again neglible). The amount of sulfide present will be the amount created by oxidation of organic matter, less the amount of organic matter oxidized by iron, so the sulfide conservation equation is:

$$\Sigma H_2 S = H_2 S + HS^- = +\frac{1}{2}\Sigma C H_2 O - \frac{1}{8}\Sigma F e^{3+} = 1 \times 10^{-4} M$$
 13.14

We now want to calculate the speciation of sulfide. We have

 $\Sigma H_2 S = H_2 S + H S^- = 1 \times 10^{-4} M$ and $K_{1H_2S} = \frac{[H^+][HS^-]}{[H_2S]} = 10^{-7}$

Solving these two equations, we have:

$$[\mathrm{HS}^{-}] = \frac{10^{-4}}{10^{7} 10^{-6.23}} = 1.92 \times 10^{-5}$$

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

The concentration of H_2S is then easily calculated as 8.08×10^5 M. We can now calculate the p ϵ of the solution by substituting the above values and the p ϵ° in Table 13.1 for reaction 7 into equation 13.7. Doing so, we find p ϵ is -2.72. Finally, for iron we have:

$$\Sigma Fe = [Fe^{2+}] + \Sigma [Fe^{3+}] = 10^{-6}$$
 and $\log \frac{[Fe^{2+}]}{[Fe^{3+}]} = p\epsilon - p\epsilon^{\circ} = -2.72 - 13.0 = 15.78$ 13.15

so that $Fe^{2+} = 10^{15.78}Fe^{3+}$. The equilibrium concentrations of the hydrolysis species of Fe^{3+} can be calculated from equations 6.73a through 6.73d. We find the most abundant species will be $Fe(OH)^+_2$, which is $10^{6.8}$ times more abundant than Fe^{3+} . However, Fe^{2+} remains $10^{15.78} \times 10^{-6.9}$ more abundant that $Fe(OH)^+_2$, so for all practical purposes, all iron is present as Fe^{2+} .

The development of anoxic conditions leads to an interesting cycling of iron and manganese within the water column. Below the oxic-anoxic boundary, Mn and Fe in particulates are reduced and dissolved. The metals then diffuse upward to the oxic-anoxic boundary where they again are oxidized and precipitate. The particulates then migrate downward, are reduced, and the cycle begins again.

A related phenomena can occur within sediments. Even where anoxic conditions are not achieved within the water column, they can be achieved within the underlying sediment. Indeed, this will occur where burial rate of organic matter is high enough to exceed the supply of oxygen. Figure 13.3 shows an example, namely a sediment core from southern Lake Michigan studied by Robbins and Callender (1975). The sediment contains about 2% organic carbon in the upper few centimeters, which decreases by a factor of 3 down core. The concentration of acid-extractable Mn in the solid phase (Figure 13.3a), presumably surface-bound Mn and Mn oxides, is constant at about 540 ppm in the upper 6 cm, but decreases rapidly to about 400 ppm by 12 cm. The concentration of dissolved Mn in the pore water in-



Figure 13.3. (a) Concentration of acid-leachable Mn in Lake Michigan sediment as a function of depth. (b) (b) Dissolved Mn in pore water from the same sediment core. Solid line shows the dissolution-diffusion-reprecipitation model of Robbins and Calender (1975) constrained to pass through 0 concentration at 0 depth. Dashed line shows the model when this constraint is removed. (c) Dissolution rate of solid Mn calculated from rate of change of concentration of acid-leachable Mn and used to produce the model in (b). From Robbins and Calender (1975).

creases from about 0.5 ppm to a maximum of 1.35 ppm at 5 cm and then subsuquently decreases (Figure 13.3) to a constant value of about 0.6 ppm in the bottom half of the core.

Because the Lake Michigan region is heavily populated, it is tempting to interpret the data in Figure 13.3, particularly the increase in acid-extractable Mn near the core top, as being a result of recent pollution. However, Robbins and Callender (1975) demonstrated that the data could be explained with a simple steady-state diagentic model involving Mn reduction, diffusion, and reprecipitation (as MnCO₃). In Chapter 5, we derived the Diagenetic Equation:

$$\left(\frac{\partial \mathbf{c}}{\partial t}\right)_{\mathbf{x}} = \left(\frac{\partial F}{\partial \mathbf{x}}\right)_{\mathbf{t}} + \sum \mathbf{R}_{\mathbf{i}} \quad (5.171)$$

The first term on the right is the change in total vertical flux with depth, the second is the sum of rates of all reactions occuring. There are two potential flux

Geochemistry

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

terms in this case, pore water advection (due of compaction) and diffusion. There are also several reaction occurring: dissolution or desorption associated with reduction and a precipitation reaction. If the system is at steady state, then $\partial c/\partial x = 0$. Assuming steady-state, Robbins and Callender (1975) derived the following verision of the diagenetic equation:

$$\phi D \frac{d^2 c}{dx^2} - v \frac{dc}{dx} - \phi k_1 (c - c_f) + \phi k_0(z) = 0$$
 13.16

where ϕ is porosity (assumed to be 0.8), D is the diffusion coefficient, v is the advective velocity (-0.2 cm/yr), and k₁ is the rate constant for reprecipitation reactions, and k₀ is the dissolution rate (expressed as a function of depth). The first term is the diffusive term, the second the advective, the third the rate of reprecipitation, and the fourth is the dissolution rate. The dissolution rate must be related to the change in concentration of acid-extractable Mn. Thus the last term may be written as:

$$\phi k_0(z) = \phi \frac{R}{\phi} \frac{\partial c_s}{\partial x}$$

where R is the sedimentation rate $(g/cm^2/yr)$ and c_s is the concentration in the solid. Using least squares, Robbins and Callender found that the parameters that best fit the data were $D = 9 \times 10^{-7}$ cm²/sec $(30 \text{ cm}^2/\text{yr}) \text{ k}_1 = 1 \text{ yr}^{-1}$ and $c_f = 0.5 \text{ ppm}$. The solid line in Figure 12.44b represents the prediction of equation 13.16 using this values and assuming c_0 (porewater concentration at the surface) is 0. The dashed line in Figure 13.3b assumes $c_0 = 0.6 \text{ ppm}$. The latter is too high, as c_0 should be the same concentration as lake water. Robbins and Callender (1975) speculated that the top cm or so of the core had been lost, resulting in an artifically high c_0 .

Redox cycling, both in water and sediment can effect the concentrations of other a number of other elements. For example, Cu and Ni form highly insoluble sulfides. Once pe decreases to levels where sulfate is reduced to sulfide, dissolved concentrations of Cu and Ni decrease dramatically due to sulfide precipitation. The dissolved concentrations of elements that are strongly absorbed onto particulate Mn and Fe oxihydroxide surfaces, such as the rare earths and P, often show significant increases when these particulates dissolved as Mn and Fe are reduced. The effect of Fe redox cycling on P is particularly significant because P is most often the nutrient whose availability limites biological productivity in freshwater ecosystems. Under oxic conditions, a fraction of the P released by decomposition of organic matter in deep water or sediment. If conditions become anoxic, iron dissolves and adsorbed P is released into solution, where it can again become available to the biota. As a result, lakes that become eutrophic due to P pollution can remain so long after the the pollution ceases because P is simply internally recycled under the prevailing anoxic conditions. Worse yet, once conditions become anoxic, nonanthropogenic P can be released from the sediment, leading to higher biological production and more severe anoxia.

WEATHERING, Soils, and Biogeochemical Cycling

Weathering is the process by which rock is physically and chemically broken down into relatively fine solids (soil or sediment particles) and dissolved components. The chemical component of weathering, which will be our focus, could be more precisely described as the process by which rocks originally formed at higher temperatures come to equilibrium with water at temperatures prevailing at the the surface of the Earth.

Weathering plays a key role in the *exogenic* geochemical cycle (i.e., the cycle operating at the surface of the Earth). Chemical weathering supplies both dissolved and suspended matter to rivers and seas. It is the principal reason that the ocean is salty. Weathering also supplies nutrients to the biota in form of dissolved components in the soil solution; without weathering terrestrial life would be far different and far more limited. Weathering can be an important source of ores. The Al ore bauxite is the product of extreme weathering that leaves a soil residue containing very high concentrations of aluminum oxides and hydroxides. Weathering, together with erosion, transforms the surface of the Earth, smoothing out the roughness created by volcanism and tectonism.

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

Geochemists are particularly concerned with question of what controls weathering rates. This concern arises from both the role weathering plays as a sink for atmospheric CO_2 and the variety of anthropogenic activities that influence weathering rates. Agriculture and harvesting of forests have had a clear impact on erosion, and probably chemical weathering rates as well. Combustion products of fossil fuels released to the atmosphere include nitrates and sulfates that acidify precipitation ("acid rain"). The resulting decrease in pH has had a clear adverse effect on the biota in some localities and may also affect chemical weathering rates. On the other hand, weathering consumes H⁺ and thereby increases pH; in some localities this buffering effect of weathering reactions is sufficient to entirely overcome the effects of acid rain. Understanding the impact of acid rain and the degree to which its effects are mitigated by weathering is an important goal of many weathering studies.

Weathering, and the subsequent precipitation of carbonate in the ocean, also consumes CO_2 in reactions such as:

 $CaAl_2Si_2O_8 + H_2CO_3 + 3H_2O \rightleftharpoons CaCO_3 + Al_2Si_2O_5(OH)_4$

Weathering thus appears to be an important control on the concentration of atmospheric CO_2 , which is in turn an important control on global temperature and climate. Hence whatever factors control the weathering rate may also control climate. What are these factors? In the BLAG (Berner Lasaga and Garrels) model of atmospheric CO_2 levels over geological time, Berner et al. (1983) and Berner (1991) assumed that temperature is a strong control on weathering rate, and therefore that there is a negative feedback that controls atmospheric CO_2 levels (the higher the atmospheric CO_2 , the higher the global temperature, the higher the weathering rate, and therefore the higher the consumption of atmospheric CO_2). Others, including Edmond et al. (1995), have argued that despite the odvious temperature effect on reaction rates, global temperature exerts little control on weathering rates in nature, and that tectonic uplift and exposure of fresh rock has a much stronger influence on weathering rate, and ultimately on global climate.

In previous chapters and sections, we examined many important weathering reactions from thermodynamic and kinetic perpectives. In this section, we will step back to look at weathering on a broader scale and examine weathering in nature and the interrelationships among chemical weathering, biological processes, and soil formation. We then discuss in some detail the question of what controls on weathering rates. Finally, we look at the composition of rivers and streams.

Most rock at the surface of the Earth is overlain by a thin veneer of a mixture of weathering products and organic matter that we refer to as soil. Most weathering reactions occur within the soil, or a t the interface between the soil and bedrock, so it is worth briefly considering soil and its development. Soil typically consists of a sequence of layers that constitute the soil profile (Figure 13.4). The nature of these layers varies, depending on climate (temperature, amount of precipitation, etc.), vegetation (which in turn depends largely on climate), time, and the nature of the underlying rock. Consequently, no two soil profiles will be identical. What follows is a general description of an idealized soil profile. Real profiles, as we shall see, are likely to differ in some respects from this.

Soil Profiles

The uppermost soil layer, referred to as the *O* horizon, consists entirely, or nearly so, of organic material whose state of decomposition increases downward. This layer is best developed in forested regions or waterlogged soils depleted in O_2 where decomposition is slow. In other regions it may be incompletely developed or entirely absent.

Below this organic layer lies the upper mineral soil, designated as the *A horizon* or the zone of removal, which ranges in thickness from several centimeters to a meter or more. In addition to a variety of minerals, this layer contains a substantial organic fraction, which is dominated by an amorphous mixture of unsoluble, refractory organic substances collectively called *humus*. Weathering reactions in this layer produce a soil solution rich in silica and alkali and alkaline earth cations that percolates downward into the underlying layer. In temperate forested regions, where rainfall is high and organic decomposition slow, Fe and Al released by weathering reactions are complexed by organic

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY



Figure 13.4. Soil profile, illustrating the O, A, B, and C horizons described in the text. Not all soils comform to this pattern.

(fulvic) acids and carried downward into the underlying layer. The downward transport of Fe and Al is known as *podzolization*. In tropical regions, organic decomposition is sufficiently rapid and complete that there is little available soluble organic acid to complex and transport Fe and Al, so podzolization may not occur. As a result, Fe and Al accumlate in the A horizon as hydrous oxides and hydroxides. Where leaching is particularly strong, the A horizon may be underlain by a thin whitish highly leached, or eluviated, layer known as the *E horizon*, enriched in highly resistant minerals such as quartz. In grasslands and deserts, the production of soluble organic acids is restricted by the availability of water, hence the podzolization is limited.

The *B horizon*, or zone of accumulation or deposition, underlies the A horizon. This horizon is richer in clays and poorer in organic matter than the overlying A horizon. Substances leached from the A horizon are deposited in this B horizon. Fe and Al carried downward as organic complexes precipitate here as hydrous oxides and hydroxides, and may react with other components in the soil solution to form other secondary minerals

such as clays. In arid regions where evaporation exceeds precipitation, relatively soluble salts such as calcite, gypsum, and halite precipitate within the soil. Downward percolating water leaches these from the A horizon and concentrates them within in the B horizon. Such calcite layers are known as *caliche*, and are typically found at a depth of 30 to 70 cm. The clay-rich nature of the B horizon, particularly when cemented by precipitated calcite or Fe oxides, can lead to greatly restrict permeability of this layer. Such impermeable layers are sometimes referred to as *hardpan*. In tropical regions, where weathering is intense and has continued for millions of years in the absence of disturbances such as glaciation, the soil profile maybe up to 100 m thick. Because of the absence of podzolization, there is often little distinction between the A and B horizons. Base cations are nearly completely leached in tropical soils, leaving a soil dominated by minerals such as kaolinite, gibbsite, and Fe oxides. Such soils are called *laterites*. In the most extreme cases, SiO_2 maybe nearly completely leached as well, leaving a soil dominated by gibbsite. The ratio of SiO_2 to $Al_2O_3 + Fe_2O_3$ (collectively called the sesquioxides) is a useful index of the intensity of weathering within the soil as well as to the extent of podzolization. Typical values for the A and B horizons of several climatic regions are summarized in Table 13.2. Often, seperate layers can be recognized within the B and the other horizons and these are designated B_1 , B_2 , etc. downward.

The *C* horizon underlies the B horizon weathered rock, often only coarsely fragmented, and its direct weathering products. Very often it consists of *saprolite*, rock in which readily weatherable silicates have been largely or wholy replaced *in situ* by clays and oxides, but textures and structures are often sufficiently well preserved that the nature of the original parent maybe recognized. This

The *C* horizon underlies the B horizon and directly overlies the bedrock. It consists of partly

TABLE 13.2. SiO₂/(Al₂O₃ + Fe₂O₃) Ratios of Soils

L L		
	A horizon	B horizon
Boreal	9.3	6.7
Cool-temperate	4.07	2.28
Warm-temperate	3.77	3.15
Tropical	1.47	1.61

From Schlesinger (1991).

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

layer has relatively little organic matter. In soils that develop directly from local materials, it is mineralogically related to the underlying bedrock. Alternatively, soil may develop on volanic ash, glacial till, or material transported by wind (loess) or water (alluvium). Loess up to 100 m thick was deposited in some northern regions during the Pliestocene glaciations. Soils in deserts surrounded by mountains and in river floodplains generally develop from alluvium rather than underlying bedrock. In such cases, the C layer will be unrelated to underlying bedrock or may be underlain by fossil soils (paleosols). While organic acids dominate weathering reactions in the upper part of the soil profile, carbonic acid, produced by respiration within the soil and transported downward by percolating water, is largely responsible for weathering in the C horizon.

The full development of an "equilibrium" soil profile requires time, thousands to hundreds of thousands of years. How much time depends largely on climate: soil development is most rapid in areas of warm temperature and high rainfall. Regions that have been disturbed in the recent geologic past may show incomplete soil formation, or soils uncharacteristic of the present climate. For example, many desert soils contain clay layers developed during wetter Pleistocene times. Soils in recently glaciated areas of North America and Europe remain thin and immature, even though 12,000 years have past since the last glacial retreat. Floods and landslides are other natural processes that can interupt soil development. Soil profiles are also be disturbed by agriculture and the resulting increase in erosion and other anthropogenic effects such as acid rain. Thus actual soil profiles often deviate from that illustrated in Figure 13.4.

Chemical Cycling in Soils

An example of chemical variations in a soil profile is shown in Figure 13.5. This soil, developed on a beach terrace in Mendocino County, California, was studied by Brimhall and Dietrich (1987) and illustrates processes occuring in a podzol (a soil that has experienced podzolization). The parent material is Pleistocene beach sand and is poor in most elements other than Si compared to common rocks. The soil profile consists of an A horizon, the upper part of which is densely rooted and rich in organic matter, unlain by a transitional layer that Brimhall and Dietrich (1987) labelled Bmir. This is in turn underlain by a white and red banded E horizon, which is in turn unlain by the B horizon, which Brimhall and Dietrich (1987) divided into B1 and B2 layers based on appearance.

Figure 12.46 shows that Fe and Al are slightly, though not uniformly, depleted from the surface through the E horizon, and strongly enriched in the B horizon. This reflects the downward transport of these element as organic complexes. Si is present in roughly the same concentrations as in the parent through the E horizon, and relatively depleted in the underlying B horizons. Though Si has undoubtedly been leached throughout the profile, the leaching has been more severe for other elements than for Si in the upper profile, leaving the Si concentration nearly unchanged. The lower concentrations in the lower horizon in part reflects dilution by Fe and Al. The alkalis and alkaline earths are depleted thoughout the profile relative to the parent, but while Na and Ca are uniformly depleted (as is Sr, not shown in the Figure), K, Rb, and Ba are enriched in the B horizon. These elements are all readily soluble and hence easily leached in the upper part of the soil profile, but K, Rb, and B a are readily accepted in the interlayer sites of clays. Thus clay formation probably explains the enrichment of these elements in the B horizon. The profiles of Pb, Cu (not shown), and Ga resemble that of K (though they are also somewhat enriched at the base of the A horizon). Brimhall and Dietrich (1987) speculated that this is due to oxidation of sulfides in the upper horizon, downward transport as organic complexes, and reprecipitation as sulfides in the B horizon. Ti and Nb, as well as Zr (not shown) are present in concentrations greater than that of the parent throughout most of the profile. These elements are not soluble and do not form soluble complexes hence they are virtually immobile in soil profiles. Their apparent enrichment is due entirely to the loss of other elements. Indeed, Brimhall and Dietrich use Nb and Ti concentrations to calculate that there is been a mass loss, through leaching, of 60% in the upper part of the soil profile.

Geochemistry

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

From the preceeding discussion, we can infer that considerable chemical cycling occurs within soils. The initial stages of weathering occur within the C horizon, where the most unstable primary miner-



Figure 13.5. Concentration profiles in a podzol developed on a Pleistocene beach terrace in Mendicino, California. Dashed line shows the concentration in the parent beach sand. Soil horizons are simplified from those identified by Brimhall and Dietrich (1987). Data from Brimhall and Dietrich (1987).

Geochemistry

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

als (e.g., sulfides, carbonates, feldspars, ferromagnesian silicates such as olivine, pyroxenes, amphiboles) undergo reaction. More resistant minerals such as quartz and Fe-Ti oxides may remain largely intact at this stage. A substantial fraction of the more soluble elements, such as the alkalis and alkaline earths, may be lost to solution at this stage. As erosion slowly lowers the surface, minerals in the C horizon are, in effect, transported into the B horizon. At this stage, new secondary minerals form by precipitation from and reaction with downward percolating solutions. In areas of high rainfall, the most important of these are Fe and Al oxihydroxides; in arid regions, carbonates and sulfates may be the dominant new minerals forming in the B horizon. Further leaching and loss of readily soluble elements, as well as SiO_2 , will continue in this stage in humid climates. As erosion continues to lower the surface, the material will eventually reach the A horizon. Here, many of the secondary minerals formed in the C and B horizons breakdown, releasing their components to solution. Some of the material leached from this horizon will reprecipitate in the B horizon and some will be lost to groundwater and streams. In addition, plants actively take up dissolved consitutents from the soil solution and store it within their tissues. These constituents will be returned to the soil when the tissue dies and is broken down by bacteria. Thus in principle, an element might be cycled many times between the biota, and the O, A, and B horizons before being carried away in groundwater flow.

Biogeochemical Cycling

The biota plays a substantial role in the weathering process and in controlling the composition of streams. As we saw earlier in this chapter, P_{CO_2} is substantially higher in soil than in the atmosphere; this is a direct result of respiration of organisms in the soil. This higher P_{CO_2} lowers pH and hence accelerates weathering reactions. Organic acids produced by plants and bacteria have the same effect. In addition, many weathering reactions, including, but not limited to, redox reactions, may be directly catalyzed by soil bateria. Plants also take up a host of elements released by rock weathering as nutrients. The biota will thus influence both soil and water chemistry within a water-shed and can be a significant reservoir for some elements.

Living organisms consist of a bewildering variety of organic compounds (we will discuss some of these is Chapter 13). From a geochemical perspective, it is often satisfactory to approximate the composition of the biomass as CH₂O (for example the composition of glucose, a simple sugar, is $C_6H_{12}O_6$). A better approximation for the composition of land plants would be C1200H1900O900N25P2S1 (Berner and Berner, 1996). A great many other nutrients^{*}, however, are essential for life (for example, Mg and Fe are essential photosynthesis; Mo is essential for nitrate reduction) and are taken up by plants in smaller amounts. These can be divided up into macronutrients, which occur in plants at concentrations in excess of 500 ppm and include N, P, K, Ca, Mg, and S, and micronutrients, which occur at lower concentrations and include B, Fe, Mn, Cu, Zn, Mo, Co, and Cl. Other elements are also taken up by plants and stored in tissue even though they play no biochemical role (so far as we know), simply because plants cannot discriminate sufficiently against them. Table 13.3 lists the concentrations of the elements in dried plant matter. One should be aware, however, that the actual composition of plants varies widely; grasses, for example, can contain over 1% SiO₂ (>4600 ppm Si).

sition of Dried Plant Matter	TABLE 13.3. ELEMENTAL COMPO-
	sition of Dried Plant Matter

Element	Percent	Element	ppm
С	49.65	V	1
Ν	0.92	Mn	400
0	43.2	Cr	2.4
Total	93.77	Fe	500
Element	ppm	Со	0.4
Li	0.1	Ni	3
В	5	Cu	9
Na	200	Zn	70
Mg	700	Se	0.1
Al	20	Rb	2
Si	1500	Sr	20
Р	700	Мо	0.65
S	500	Ag	0.05
Κ	3000	Ba	30
Ca	5000	U	0.05
Ti	2		

From Brooks (1972).

^{*} A nutrient as an element or compound essential to life that cannot be synthesized by the organism and therefore must be obtained from an external source.

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

There are four sources of nutrients in an ecosystem: the atmosphere, dead organic matter, water, and rock. The atmosphere is the obvious direct source of CO_2 and O_2 and the indirect source of N and H_2O in terrestrial ecosystems. However, it may also be the direct or indirect source of a number of other nutrients, which arrive either in atmospheric dust or dissolved in rain. Plants are able to take up some of these atmosphere-delivered nutrients directly through foliage; most, however, cycle through the soil solution and are taken up by roots, which is the primary source of nutrients. Plants cannot take up nutrients from dead organic matter or from rocks directly: nutrients from these sources must first be dissolved in the soil solution. Because equilibrium between surface adsorbed and dissolved species is achieved relatively quickly, elements adsorbed on the surfaces of oxides, clays, and organic solids represent an intermediate reservoir of nutrients. For example, phosphrous, often the growth-limiting nutrient, is readily adsorbed on the surface of iron oxides and hydroxides. For this reason, the surface properties of soil particles are an important influence on soil fertility. Even in relatively fertile soils, however, the concentration of key nutrients such as phosphorous may be effectively zero in the soil immediately adjacent roots, and the rate of delivery to plant may be limited by diffusion.

In most ecosystems, particularly mature ones, detritus, that is dead organic matter, is the the most important source of nutrients. In the Hubbard Brook Experimental Forest, for example, this recycling supplies over 80% of the required P, K, Ca, and Mg (Schlesinger, 1991). For the most part, this recycling occurs as leaf tissue dies, falls to the ground, and decomposes. However, some fraction of nutrients are recycled more directly. Nutrients may be leached from leaves by precipitation, a process called *translocation*. Translocation is particularly important for K, which is highly soluble and concentrated in cells near the leaf surface, but it can be important for other elements as well. Nutrient loss from leaves by leaching increases in the order K>>P>N>Ca. In addition to nutrient recycling through translocation and detritus, plants also recycle nutrients internally by withdrawing them from leafs and stems before the annual loss of this material and storing them for use in the following season. For this reason, the concentration of nutrients in litterfall is lower than in living tissue. Not surprisingly, the fraction of nutrients recycled in this way, and overall nutrient use efficiency, is higher in plants living on nutrient-poor soils (Schlesinger, 1991).

Rainwater passing through the vegetation canopy will carry not only nutrients leached from foliage, but also species dissolved from dust and aerosols (together called dry deposition) deposited on leaves. Fog and mist will also deposit solutes on plant leaves. The term *occult* deposition refers to both dry deposition and deposition from mist and fog. The total flux of solutes dissolved from leaf surfaces, including both the occult deposition and translocation fluxes, and carried to the soil by precipitation is called *throughfall*, and can be quite significant in regions where there is a high aerosol flux. Such regions may be either those downwind from heavily populated areas, where the atmosphere contains high levels of nitrate and sulfate from fossil fuel burning, or arid regions, where there is abundant dust in the atmosphere. Table 13.4 compares the concentration of nutrients measured in Table 13.4. Concentrations in Bulk Precipitation and Throughfall in the Vosace

	NH_4	Na	K	Mg	Ca	H^{+}	Cl	NO_3	SO_4
Concentration (μ eq/L)									
Bulk precipitation	19.1	10.0	2.8	4.5	11.9	33.9	12.5	24.1	41.5
Throughfall	36.9	46.4	52.7	17.8	65.5	114.8	63.4	48.3	185.0
Fluxes (moles/ha/y)									
Bulk precipitation	270.0	142.0	39.0	32.0	84.0	480.0	177.0	340.0	290.0
Throughfall	385.0	484.0	550.0	93.0	642.0	1197.0	661.0	817.0	966.0
Difference	115.0	342.0	511.0	61.0	558.0	717.0	484.0	477.0	676.0
Occult precipitation	115.0	342.0	102.0	31.0	206.0	1282.0	484.0	477.0	676.0
Translocation	0.0	0.0	409.0	30.0	352.0	-565.0	0.0	0.0	0.0
\mathbf{D}	0.0)								

Table 13.4. Concentrations in Bulk Precipitation and Throughfall in the Vosage, France

Data from Probst et al. (1990).

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

throughfall and bulk precipitation, and demonstrates the importance of translocation and occult deposition.

The biota affects the composition of soil and stream water in another way as well. Some fraction of soil water taken up by roots is ultimately lost from the plant to the atmosphere through leaf stomata, the opening designed to allow CO_2 into the leaf. This loss of water is called *transpiration*. Water lost through transpiration and that lost by direct evaporation from the ground sur-

TADLE 13.5.	Residence	Times of	Organic	MATTER
AND NUTRIEN	fs in <u>Fores</u>	t Litter	-	

Region	Organic Matter	Ν	Р	K	Ca	Mg
Boreal forest	353	230	324	94	149	455
Temperate forest						
coniferous	17	17.9	15.3	2.2	5.9	12.9
diciduous	4	5.5	5.8	1.3	3	3.4
Meditteranean	3.8	4.2	3.6	1.4	5	2.8
Tropical rainforest	0.4	2	1.6	0.7	1.5	1.1

From Schlesinger (1991).

face are often collectively called *evapotranspiration*. As one might expect, transpiration varies seasonally: transpiration is high in spring and summer when plants are actively growing and stomata are open and minimal in winter. Transpiration also depends on climatic factors such as temperature and relative humidity, as does evaporation. Evaptranspiration concentrates dissolved solids in soil and stream water.

Dead vegegtation lying above the mineral soil (the O soil horizon) is called *litter*. The rate at which litter decomposes, and hence "turns over" depends strongly on climate. Table 13.5 lists the mean residence times of bulk organic matter and nutrients in the surface litter of forest ecosystems. The great range of times, from hundreds of years in boreal forests to a year or less in tropical rainforests, is particularly interesting. K is recycled more rapidly than bulk organic matter, but recycling times for other nutrients are generally comparable to that of bulk organic matter. Though animals, particularly those living in the soil such as termites and worms, play a role in organic decomposition, most of it is carried out by soil fungi and bacteria. These soil microbes can comprise up to 5% of the organic matter is acomplished by extracellular enzymes released by these organisms. Because soil microbes concentrate them, a particularly high fraction of organicly bound N and P in soils is contained in the microbial biomass.

As microbes decompose organic matter, they preferentially oxidize the most labile, energy-rich compounds, such as sugars, and synthesize humus from refractory compounds such as lignins and tannins. Soil humus together with humic and fulvic acids, which are closely related (see Chapter 13), accumulate within the soil. As we have mentioned, the cation exchange capacity of this soil organic matter is important, both in providing a reservoir of nutrients to plants, and also in downward transport of Al and Fe in soils. In most cases, the mass of soil humus exceeds the combined mass or living vegetation and litter. The residence time humus in the soil may exceed that of litter by several orders of magnitude, with measured mean ¹⁴C ages ranging upward to thousands of years.

WEATHERING RATES AND REACTIONS

The Watershed Approach

As we stated at the beginning of this section, weathering produces two products: secondary minerals and dissolved components. The process may be studied in a variety of ways and on a variety of scales. Perhaps the most basic study is simply to observe the phase that replace original ones as weathering of exposed rock proceeds. These observations providing the starting point for laboratory experiments from which thermodynamic and kinetic data may be deduced. A third approach, and the one we explore in this section, looks at the problem on a much larger scale: that of a watershed. This approach relies on the observation that mass is conserved in weathering reactions (like any other chemical reaction). Mathematically, we may write the following mass balance equation:

rock + rain = altered rock + solution

Geochemistry

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

Thus if the composition of the original rock, input composition (bulk precipitation and occult precipitation), and the final water composition (the dissolved component) are known, then the gross composition of the secondary phases can be calculated. This point is illustrated in Figure 13.6. Such an approach ignores considerable complexity and details, such as the geochemical and biogeochemical cycling that occurs in the soil and biota. Often, a key assumption in such studies is that the system is at steady-state; if so, any internal cycling will not affect the net output of the system. It is, however, not always clear that the is valid, Nevertheless, such studies can be enor-



assumption of steady-state Figure 13.6. Illustration of the mass balance approach to weathering on is valid, Nevertheless, a watershed scale.

mously useful in understanding weathering, particularly when it is combined with thermodynamics and kinetics to deduce the nature of weathering reactions occurring.

Spring Waters of the Sierra Nevada

The classic example of this type of study of is work of Feth et al. (1964) and Garrels and McKenzie (1967) on springs in the granitic terrane of the Sierra Nevada. Feth et al. (1964) measured concentrations of the principal constituents of perennial (those that always flowed) and ephemeral springs (springs that flowed only seasonally or after rain), and precipitation (atmospheric input). The data is summarized in Table 13.6. Garrels and McKenzie (1967) showed that composition of these waters could be explained by weathering of the local bedrock.

The granitic bedrock consists primarily of quartz, alkali feldspar, and andesine plagioclase with lesser amounts of hornblende and biotite. The primary weathering product observed was kaolinite. Garrels and McKenzie (1967) first subtracted from spring water the concentrations of ions found in snow. They then subtracted Na, Ca, HCO_3 , and SiO_2 in proportions to convert kaolinite to plagioclase. Next all Mg and enough K were subtracted to make biotite, and the remaining K, HCO_3 , and SiO_2 used to convert kaolinite to K-feldspar. For the ephemeral springs, this procedure produced mass balance within analytical error. In other words, the various ions were present in solution in the proportions expected from the reaction considered by Garrels and McKenzie.

The mass balance for the perennial springs was less satisfactory. They had higher total dissolved solids, higher pH and higher HCO_3 , all of which indicated deeper circulation and more extensive interaction with rock, which is what one would expect. However, cation concentrations did not increase in the proportion expected if kaolinite were the only weathering product. In particular, the data suggested a more siliceous residual phase, as the increase in Na and SiO₂ in the perennial springs over

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

Ephemeral Springs	Peren	nial Springs		
<i>ppm</i>	molality ×10) ⁴ ppm	molality ×10 ⁴	
SiO ₂	16.4	2.73	24.6	4.1
Al	0.03	_	0.018	_
Fe	0.03	_	0.031	_
Ca	3.11	0.78	10.4	2.6
Mg	0.70	0.29	1.70	0.71
Na	3.03	1.34	5.95	2.59
К	1.09	0.28	1.57	0.40
HCO ₃	20.0	3.28	54.6	8.95
SO_4	1.00	0.10	2.38	0.25
Cl	0.50	0.14	1.06	0.30
F	0.70		0.09	_
NO ₃	0.02		0.28	_
Σ Dissolved solids	36.0		75.0	
pH (median)	6.2		6.8	

Table 13.6. Average concentrations of dissolved constituents in springs of the Sierra Nevada (from Garrels and McKenzie, 1967).

the ephemeral springs was 1:1 whereas weathering of plagioclase to kaolinite releases Na and SiO₂ in 1:2 proportions. Smectite was a likely candidate.

The question of whether groundwater composition was controlled by a reaction between kaolinite and smectite (Ca-beidellite) can be addressed from a thermodynamic perspective. The reaction is:

 $3Ca_{0.33}Al_{4.67}Si_{7.33}O_{20}(OH)_4 + 2H^+ + 7H_2O \rightleftharpoons 7Al_2Si_2O_5(OH)_4 + Ca^{2+} + 8SiO_{2(aq)}$ 13.17 The equilibrium constant for the reaction is:

$$K = \frac{a_{Ca^{2+}}a_{SiO_2(aq)}}{a_{H^+}^2}$$
 13.18

A plot of the "reaction quotient", i.e., the right hand side of equation 13.18, against Na concentration is shown in Figure 13.7. As weathering proceeds, the composition of the solution should evolve along path ABD. On path ABC, Na and Ca are released in same proportions that they are found in the plagioclase being weathered, namely of 0.62:0.38, if kaolinite is the weathering product. When Ca concentrations reach the point that smectite is stable, we would expect no further increase in Ca concentration, but Na should continue to increase. This is the path BD. Though the data show considerable scatter, and some springs are oversaturated with respect to smectite, the data generally support the conclusion that smectite is also a weathering product.

Garrels (1967) found that many springs in various other terranes also showed a similar cutoff in the kaolinite-smectite reaction quotient, suggesting formation of smectite is an important control on water chemistry. We should point out here that the controlling reaction need not be kaolinite to smectite. Direct weathering of plagioclase to smectite occurring when the solution becomes saturated with respect to smectite produces the same pattern. Also, the equilibrium constant for this reaction has not been directly measured. Indeed, the free energy of formation of smectite, which does not form crystals large enough to make thermodynamic measurements on, is deduced from this plot.

Coweeta Basin, Southern Appalachians

Velbel (1985a, 1985b) used a mass balance approach in his study of the Coweeta Basin to calculate rates of mineral weathering in a natural environment. The Coweeta Basin, an area of 1625 hectares located in southwestern-most South Carolina, had been subject of the a number of ecological studies and intensive hydrologic monitoring by the U. S. Forest Service and others for decades. Thus data on

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY



biomass uptake and stream water compositions and fluxes were available. already There had been no anthropogenic activity in the area for at least 50 years before Velbel's work; disturbance before that was limited to controlled selection logging, so the biomass was close to steady-state. Annual rainfall is high, ranging from 170 cm at lower elevations (670 m) to 250 at higher elevations (up to 1600 m). Bedrock consists of a variety of metasediments and metavolcanics consisting of muscovite, biquartz, plagioclase otite, (oligoclase) and almandine garnet along with a variety of accessory minerals. This is

Figure 13.7. Reaction quotient for the Ca-beidellite—kaolinite reaction vs. sodium concentration. Data for the Sierra Nevada Springs is shown, as well as predicted evolution path for spring water during the weathering of plagioclase. From Garrels and McKenzie (1967).

overlain by soil averaging about 6 m in depth; 95% of this thickness is saprolite (C horizon).

Velbel's (1985a) found from petrographic, electron microprobe, and x-ray diffraction study that the primary weathering reactions were the breakdown of biotite, garnet, and feldspar; muscovite and quartz were not appreaciably weathered and abundances of other minerals were too small to affect the mass balance. Biotite weathers initially to form hydrobiotite, a mixed-layer biotite-vermiculite. The primary lattice structure is preserved in this process, which involves loss of K (and some Mg), oxidation of Fe, and uptake of dissolved Fe and Al, and some Ca. The reaction may be written approximately as:

$$\begin{aligned} \mathsf{K}(\mathsf{Mg}_{12}\,\mathsf{Fe}_{1.3}^{2+}\,\mathsf{Al}_{05})(\mathsf{AlSi}_3)\mathsf{O}_{10}(\mathsf{OH})_2 + 0.2\mathsf{O}_2 + 0.1\mathsf{H}^+ + 0.3\mathsf{H}_2\mathsf{O} + 0.02\mathsf{Ca}^{2+} + 0.3\mathsf{Al}(\mathsf{OH})_{2(\mathsf{aq})}^+ + 0.3\mathsf{Fe}(\mathsf{OH})_{2(\mathsf{aq})}^+ \\ \Leftrightarrow \mathsf{K}_{025}\mathsf{Ca}_{002}(\mathsf{Mg}_{1.1}\,\mathsf{Fe}_{0.5}^{2+}\,\mathsf{Fe}_{1.1}^{3+})(\mathsf{AlSi}_3)\mathsf{O}_{10}(\mathsf{OH})_2 \bullet 0.14\mathsf{Al}_6(\mathsf{OH})_{15} + 0.75\mathsf{K}^+ + 0.1\mathsf{Mg}^{2+} \end{aligned}$$

(some dissolved Na⁺ is also consumed in this reaction, but we have ignored it for clarity). Upon further weathering, the hydrobiotite is transformed to vermiculite and/or "pedogenic chlorite".

Almandine garnet weathers congruently. Within the C horizon, local reprecipitation of the iron as goethite and some of the aluminum as gibbsite produces a protective surface layer on almandine and weathering reactions are limited by the rate of transport of reactants and products across the layer. In higher soil horizons, organic chelating agents remove the iron and aluminim and weathering is limited only by the rate of surface reactions.

Plagioclase weathers by selective attack at defects in the lattice. In early stages of weathering, components are removed in solution and reprecipitated elsewhere. The weathering reaction may be described as:

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

$Ca_{0.25}Na_{0.75}(Al_{1.25}Si_{2.75})O_8 + 2.5H^+ + 5.5H_2O \rightleftharpoons 0.25Ca^{2+} + 0.75Na^+ + 2.75H_4SiO_{460} + 1.25Al(OH)^+_{2(ac)}$

In the second stage of weathering, precipitation of gibbsite and kaolinite occurs, in some cases close to the site of original dissolution, forming claymineral pseudomorphs of the original feldspar.

The concentrations of major cations in stream water, well water from the base of the saprolite

TABLE 13.7. Composition of Water fromTHE COWEETA WATERSHED

	K N	Va (Ca N	⁄lg p	эН
Soil	0.92	0.27	2.48	1.27	6.12
Well	0.60	1.08	1.12	0.65	5.10
Stream	0.59	1.08	1.06	0.64	6.64
From Velbel (1985b),					

(saprolite-bedrock interface), and soil water (sampled at 25 cm depth) are shown in Table 13.7. It is apparent from the Table that well water is nearly identical in composition to stream water. This indicates that stream water chemistry is determined entirely by reactions occurring as it percolates through the saprolite, with the exception that subsequent to leaving the subsurface, the water equilibrates with the atmosphere, resulting in a loss of CO₂ and increase in pH. On a plot of log (a_{K^+}/a_{H^+}) vs. log (a_{H2SiO4}) , the composistion of stream waters plot in the kaolinite stability field,. Other aspects of stream chemistry indicate that both gibbsite and kaolinite form, consistent with the observation that both minerals occur as weathering products. Velbel's summary of the weathering profile is shown in Figure 13.8. In the upper part of the soil profile, rapid flushing keeps dissolved silica concentrations low so that kaolinite stability is not attained and aluminium released by plagioclase weathering precipitates as gibbsite, or is consumed in the production of vermiculite and chlorite from biotite. Deeper in the soil profile, water is in prolonged contact with rock and acquires enough aluminum and silica to reach kaolinite saturation. Because the bedrock is highly impermeable, most water is eventually shunted laterally downslope and does not penetrate the bedrock. What little water does penetrate forms smectite in voids and fractures. These reactions, however, have little effect on stream chemistry.

To calculate weathering rates, Velbel (1985a) used a system of simultaneous mass balance equations. For each element, c, the net flux of the element out of the watershed (steam output minus rain



13.8. Schematic diagram of weathering profile and hydrology in the Coweeta Watershed. From Velbel (1985b).

input) can be expressed as the sum of its production or consumption in each weathering reaction as well as by the biomass, i.e.:

$$\Delta m_{\rm c} = \sum_{\rm i} \alpha_{\rm i} \beta_{\rm c,i} \qquad 13.19$$

where Δm_c is the net flux of element c out of the watershed, $\beta_{c,i}$ is the stoichiometric coefficient of element c in reaction i, and α_i is the number of moles produced by weathering reaction i. Both β and Δm are known; the α 's are the unknowns. For the elements K, Na, Ca, and Mg there are then 4 equations and 4 unknowns (net biomass uptake and the rates of plagio-clase, biotite, and garnet weathering), allowing Velbel to simultaneously solve for the 4 α terms.

Velbel calculated that 40,000 years were required to weather all the garnet, 140,000 years to react all the biotite, and 160,000 to weather all the plagioclase in the soil horizon. The calculated "saprolitization rate" was 3.8 to 15 cm/1000 yrs, the lower rate applying to complete destruction of primary minerals. This rate is nearly equal to the long term denudation, or erosion, rate (rate at which rock is removed from the surface) for the southern Appalachians. This agreement suggests the system is in

Geochemistry

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

steady-state, i.e., weathering penetrates bedrock at same rate weathering products are removed by erosion, maintaining a constant thickness weathering profile. However, based on sediment export rates, the short-term denudation rates for the region were much slower, by as much as a factor of 20. This indicates that up to 96% of erosion occurs not by steady-state removal of soil, but by infrequent catestrophic events such as landslides and severe storms. This is entirely consistent with convention geologic views as well as other studies of erosion in the Southern Appalachians.

When normalized to estimated mineral surface areas, Velbel (1985a) found that calculated reactions rates of garnet and plagioclase were one to two orders of magnitude slower than measured laboratory rates. Drever and Clow (1995) point out that more recent laboratory experiments have produced lower dissolution rates: the discrpancy for plagioclase is a factor of two, that for biotite a factor of 8, and that for garnet a factor

of 3.

It appears to be generally the case that weathering rates in nature are significantly slower than laboratory-determined rates (White, 1995; Drever and Clow, 1995). Velbel (1985a) suggested the difficulty in estimates of mineral surface area in natural systems as one possible cause of the discrepancy. Drever and Clow (1995) discuss several others possibililites, including aging and formation of protective surface layers on natural surfaces, the possible inhibitory effect of dissolved Al, and local approach to equilibrium in natural systems, which reduces the reaction affinity (i.e., ΔG , see Chapter 5) and slows the rate. In reviewing data from a number of studies, Velbel (1993) found that while laboratory rates are inevitably faster, the ratio of rates of dissolution of minerals determined from field and laboratory studies are similar. For example, the ratio of dissolution rate for olivine and plagioclase in the Filson Creek Watershed in Minnesota was 22, while the ratio of laboratory cation release rates from these minerals is 25. Velbel (1993) argued this implies that a physical, rather than chemical, factor is the cause of the discrepancy. He suggested the discrepancy arises because, in natural systems, water flow is heterogeneous and not all mineral surfaces are in contact with pore fluid and participate in reactions.



Figure 13.9. Stability diagram for the system K, OH, and H_2SiO_4 showing the solution path as weathering proceeds. (Illite is a clay mineral that is compositionally and structurally similar to the muscovite, KAl₃Si₃O₁₀(OH)₂, which most commonly forms in metamorphic rocks. Illite, however, is compositionally and structurally more variable. In particular, Si often substitutes for some Al and it has a deficit of K. In this analysis, we assume that illite is compositionally identical to muscovite.) Crosses show measured compositions of Coweeta stream water (Velbel, 1985b). Rainwater is presumably much more dilute, for example, the composition marked by point A. As weathering proceeds, the composition proceeds toward point C. Paths E-J and K-M are hypothetical paths of other possible solutions. See text for discussion.

Geochemistry

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

THERMODYNAMIC AND KINETIC ASSESSMENT OF STREAM COMPOSITIONS

As we found in the discussion of the Sierra Nevada spring water data, thermodynamics can provide insights as to what reactions are occurring in the weathering process. Let's consider the thermodynamics and kinetics of weathering in more detail, using the Coweeta study as an example. Figure 13.9 shows the stability diagram for the system $K_2O-Al_2O_3$ -SiO₂-H₂O with the expected paths a solution would take in weathering of K-feldspar. Stream composition data from the Coweeta Watershed (Velbel, 1985b) are plotted on the diagram as crosses. The data plot within the kaolinite stability field, consistent with Velbel's observation that kaolinite is forming within the saprolite. The rainwater, however, is presumably much more dilute. Let's arbitrarily assume it plots at point A in Figure 13.9. This point plots within the gibbsite stability field. Thus from thermodynamics, we expect the initial weathering of feldspar will produce gibbsite. The reaction (considering only the Kcomponent in the solid solution) is:

 $KAlSi_{3}O_{8} + 7H_{2}O + H^{+} \rightleftharpoons K^{+} + 3H_{4}SiO_{4(aq)} + Al(OH)_{3(s)}$ 13.20

This reaction produces dissolved K^+ and H_2SiO_4 and consumes H^+ , so the composition of the water will evolve up and to the right toward point B on Figure 13.9 (the exact path depends on solution alkalinity because species such as H_2CO_3 can dissociate to partially replace the H^+ consumed in reaction 13.20).

When point B is reached, both kaolinite and gibbsite are stable, and any additional H_4SiO_4 produced by weathering of feldspar reacts with gibbsite to produce kaolinite:

 $KAlSi_{3}O_{8} + 3Al(OH)_{3(s)} + H_{4}SiO_{4(aq)} + H^{+} \rightleftharpoons K^{+} + 2Al_{2}Si_{2}O_{5}(OH)_{4} + 3H_{2}O$ 13.21 The path is thus vertical along B–C until all gibbsite is consumed. Once it is consumed, further weathering produces kaolinite and dissolved K and $H_{2}SiO_{4}$ by the reaction:

$$KAlSi_{3}O_{8} + 3H_{2}O \rightleftharpoons K^{+} + 2H_{4}SiO_{4(aq)} + \frac{1}{2}Al_{2}Si_{2}O_{5}(OH)_{4}$$
 13.22

The path (C–D) is steeper because less H_2SiO_4 is produced in weathering to kaolinite than to gibbsite. Eventually, the K⁺ and H_2SiO_4 concentrations reach the point (D) where feldspar is stable, at which point no further weathering occurs because the solution is in equilibrium with K-feldspar.

Depending on the initial solution composition, other reaction paths are also possible. For example, a solution starting a point E in the gibbsite field would initially evolve in a similar manner to one starting a point A: producing first gibbsite then kaolinite. However, the solution starting a point E would eventually reach the illite stability field at point H. At this point, kaolinite is converted to illite through the reaction:

$$KAlSi_{3}O_{8} + Al_{2}Si_{2}O_{5}(OH)_{4} + 3H_{2}O \rightleftharpoons KAl_{3}Si_{3}O_{10}(OH)_{2} + 2H_{4}SiO_{4(ao)}$$
 13.23

The K/H ratio is unaffected in this reaction, but the H_2SiO_4 concentration continues to increase. Once all kaolinite is consumed, further weathering of K-feldspar produces additional illite and the concentrations of H⁺ decreases and K⁺ and H_2SiO_4 increase again in the reaction:

$$3$$
KAlSi₃O₈ + 12 H₂O + 2 H⁺ $\rightleftharpoons 2$ K⁺ + KAl₃Si₃O₁₀(OH)₂ + 6 H₄SiO_{4(aq)} 13.24

This continues until the stability field of K-feldspar is reached. Yet other paths, such as K-N, may miss the kaolinite stability fields altogether.

This purely thermodynamic analysis does not predict the phases actually found in the weathering profile by Velbel (Figure 13.8), as both kaolinite and gibbsite occur together, while the water composition plots well within the kaolinite-only stability field. The problem arises because we have ignored kinetics. In essence, we have assumed that the dissolution of K-feldspar is slow, but that the solution quickly comes to equilibrium with secondary minerals such as gibbsite and kaolinite. Lasaga et al. (1994) have pointed out that such an assumption is naive. Let's consider the progress of the reaction along path A'-F' from a kinetic perspective, but we follow the reaction with a reaction progress variable, ξ , which we define as the number of moles of feldspar consumed. Figure 13.10 is a reaction progress diagram showing the number of secondary mineral produced as a function of ξ , under the assumption that equilibrium between solution and secondary minerals is fast. Gibbsite ini-
CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

tially increases, then begins to decrease just as the stability boundary is reached as kaolinite begins to appear. There is a only limited region where both gibbsite and kaolinite are present (corresponding to path B-C on Figure 13.9). Once the boundary is past, only kaolinite is present.

Now let's assume that reaction rates are not infinitely fast, but that the rate of each reaction, *i*, depends on the extent of disequilibrium and can be described by the equation:

$$\Re_{i} = k_{i} \frac{\Delta G_{i}}{RT} \qquad 13.25$$

where \Re is the reaction rate, *k* is the rate constant that includes a mineral surface area per mass term, and ΔG is the affinity of the reaction. This would be the case if each reaction behaved as an elementary one (compare equation 5.78). We further assume that the value of *k* is 10 times as large for the precipitation of gibbsite and kaolinte as for feldpar dissolution. The activity-activity and reaction progress diagrams computed by Lasaga et al. (1994) under



Figure13.10. Reaction progress diagram showing the number of moles of gibbsite and kaolinite produced as a function of ξ , the number of moles of feldspar consumed. This is the equilibrium case assuming infinitely fast equilibrium between the solution, gibbsite, and kaolinite. After Lasaga et al. (1994).

these assumptions are shown in Figure 13.11. The activity-activity diagram is similar to that in Figure 13.9, although there is almost no vertical path along the gibbsite-kaolinite boundary. The reaction progress diagram, however, is quite different. We see that gibbsite and kaolinite now coexist over a wide region. This is a simple consequence of assuming finite rates for the precipitation reactions. Thus it is not surprising that Velbel (1985a,b) found that gibbsite and kaolinite coexisted in the weathering profile even though the stream compositions plot within the kaolinite-only field.

FACTORS CONTROlling WEATHERING RATES

As we stated at the outset of this section, understanding the controls on weathering rates is of great interest to geochemists, primarily because of the role weathering plays as a sink for CO_2 . Let's now consider the factors that control weathering rates. Lasaga et al. (1994) have proposed the following general form for the net rate law for weathering reactions:

$$\Re = A_{\min}k_0 e^{-E_A/RT} a_{H^+}^n \prod_i a_i^{m_i} f(\Delta G_r)$$
13.26

where A_{\min} is the mineral surface area, $ke^{-E_A/RT}$ expresses the usual dependence of the rate on temperature and activation energy (E_A), a_{H^+} expresses the dependence on pH to some power *n*, the terms $a_{i^-}^{-r_i}$ represent possible catalytic or inhibitory effects of other ions, and $f(\Delta G_r)$ expresses the dependence of the reaction rate on the extent of dissequilibrium. As Lasaga et al. (1994) and the preceeding disucssion emphasize, any analysis of weathering rates must take account of the rates of formation of secondary minerals as well as the dissolution of primary ones. There is, however, less data on the former than the latter. This equation, which can be applied to both, provides a useful point of depar-

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY



Figure 13.11. Stability and reaction progress diagrams computed assuming the rate constants for precipitation of gibbsite and kaolinite are only 10 times faster than that for the dissolution of feldspar. After Lasaga et al. (1994).

ture for our discussion. Let's consider each of these terms in the natural, rather than laboratory, context.

The mineral surface area, or more accurately, the area of mineral surface in contact with reactive solution (ground, soil water), enters the equation as a simple linear term: the greater the surface area, the greater the reaction rate. Two factors in particular will control surface area: (1) rainfall and (2) the rate of physical weathering and erosion.

Where rainfall is insufficient to continually wet all grain surfaces, the rate of chemical weathering Thus in arid regions, chemical will be lower. weathering is slow. In some cases, this also results in low erosion rates. In others, where high rates of erosion result from high relief, glaciation, or some other factor, much of the material being removed will be fresh rock rather than weathering products. The relationship between precipitation and weathering is not simple, however. Bluth and Kump (1994) examined chemical weathering rates (using concentrations of bicarbonate and SiO₂ in streams as proxies for these rates) and found that the fluxes of SiO₂ and bicarbonate in streams from a given region remain constant over a large range of runoff, indicating that weathering rates increase with increasing precipitation. But bicarbonate and SiO₂ concentrations level off and even drop when runoff exceed 100 cm/yr, indicating, that additional precipitation is acting to merely dilute weathering products rather than increase weathering rates.

Rates of chemical weathering can also be slow even in humid tropical areas if the rate of erosion is sufficiently slow to allow thick soils to develop. This results is what Stallard and Edmond (1983) called a "transport-limited" regime, where thick lateritic soil (up to 100 m) insulates the underlying bedrock from chemical attack. Edmond et al. (1995) make the point that such tranport-limited regimes can occur even in areas of high elevation, such a the

Guiana Shield in South America: "elevation per se is not the determing variable, but rather the mechanism by which it is produced." In tectonically active areas, such as the Andes, faulting generates high relief and exposes fresh rock to chemical weathering. Thus Edmond et al. emphasize the importance of tectonics in controlling global chemical weathering rates and, ultimately, atmospheric CO_2 levels. On the other hand, White and Blum (1995) found no correlation between chemical fluxes and relief, and argued that physical erosion rates do not have a critical influence of chemical weathering rates. However, the data they used came almost exclusively from North America and Europe, and thus did not include the tropical regions that were of primary interest to Edmond et al. (1995).

The $ke^{-E_A/RT}$ term in equation 13.26 expresses the usual Arrhenius exponential temperature dependence of reaction rates. We thus expect weathering rates to be higher in warm climates than in cold

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

ones, and this is indeed observed. The exact degree of temperature dependence will in turn depend on the activation energy. Various studies of the dependence of weathering rates on climate suggest the activation energy for chemical weathering is in the range of 40 to 80 kJ/mol (Lasaga et al., 1994; White and Blum, 1995). This is consistent with the average of activation energies of weathering reactions determined in laboratory studies (e.g., Table 5.4). This activation energy means that in increase in temperature of 8°C would result in a doubling of the weathering rate if all other factors remain constant.

Values of k and E_A vary from mineral to mineral (Table 5.4), hence weathering rates also depend on rock type. By examining stream composition in relation to rock type, Meybeck (1987) concluded that compared to granite, gneiss, and mica schist, gabbros and sandstones weathered 1.3 times as fast, volcanic rocks 1.5 times as fast, shales 2.5 times, serpentinites, marbles, and amphibolites 5 times, carbonates 12 times, gypsum 40 times, and rock salt 80 times as fast. These relative weathering rates are in at least qualitative agreement with rates predicted from laboratory studies of reaction rates (compare Table 5.3).

Numerous laboratory studies have found that the rates of weathering reactions depend on pH, but this dependence can be complex. In general, rates increase with decreasing pH in the acid range and increase with pH in the basic range. In some cases, rates appear to be independent of pH in the nuetral range. For example, Chow and Wollast (1985) found that the albite dissolution showed a dependence of the form a_{μ^+} below pH7 and a_{μ^+} above it (see Figure 5.33). Though the data on ferromagesian silicates (such as pyroxenes and amphiboles) is less clear, it appears that the pH dependence is grossly similar, at least in the acid range, to that of feldspars (Brantley and Chen, 1995). pH of weathering solutions will depend primarily on three factors: dissolved CO₂, the presence of organic acids, and the extent of weathering. Biological activity in the soil increases dissolved CO₂ and decreases pH. At the same time, it releases organic acids that also decrease pH. On the other hand, weathering reactions consume H⁺. Hence the pH of ground and soil water will progressively decrease as it reacts with rock. The presence of specific ions in solution may also affect reaction rates. For example, Chou and Wollast (1985) argued that dissolved Al inhibits feldspar dissolution, though this remains controverial. Organic acids may also have an effect on weathering rates beyond merely decreasing pH, because such acids can form surface complexes that directly promote weathering.

The final term in equation 13.26 expresses the expectation that reaction rates will decrease as equilibrium is approached. If the overall reaction rate is controlled by a single rate-limiting elementary reaction, we would expect this dependence to have the form predicted by transition state theory (see Chapter 5). However, Lasaga et al. (1994) point out that critical role played by crystal defects in weathering reaction leads to more a complex dependence of reaction rates on ΔG . The extent of disequilibrium will depend mainly on the rate at which water percolates through the soil. At high rates of flow, ΔG is large (i.e., equilibrium is not approached) and reaction rates are high. At lower flow rates, solution and rock more closely approach equilibrium and reaction rates slow. The dependence on ΔG helps explain the complex dependence of concentration on runoff observed by Bluth and Kump (1994).

The Effect of Plants on Weathering Rates

The biota clearly has an impact on chemical weathering, and for the most part, the effect of the biota is to increase weathering rates. Plants, and the soil microbes that survive on their detritus, increase soil P_{CO_2} and thereby decrease soil pH through respiration. They take up a variety of nutrients from the soil solution, and thus increase the disequilibrium between primary minerals and the soil solution. The decomposition products include organic acids, which also reduce soil pH. Plants also contribute to the physical disintigration of bedrock and by helping to form and retain a soil, keep rock in contact with water, which is essential for weathering reactions. Furthermore, transpiration returns water to the atmosphere, and hence increases rainfall. However, just how much terrestrial life has accelerated weathering is a matter of debate. Schwartzman and Volk (1989), for example, concluded that the existence of land plants has increased weathering rates by 2 to 3 orders of magnitude. Drever

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

(1994), however, argues that the direct chemical effect of land plants on weathering rates is probably no more than a factor of 2, and the overall effect is less than an order of magnitude.

Drever (1994) notes that the rates of many weathering reactions are independent of pH in the pH range of 4.5 to 8 (he acknowledges that weathering rates of many ferromagnesian silicates, such as pyroxenes and amphiboles, do increase with decreasing pH in this range). Most soil solutions have pH above 4.5, so pH decreases due to the biota have only a small effect on weathering rate. He also argues that while there is evidence that organic acids, such as oxalic acid, accelerate weathering reactions, the effect is typically only a factor of 2 at concentrations in the mM range. In nature, the organic acid concentrations are much lower and the overall effect of organic acids is likely to be much smaller. Although significant amounts of nutrients are stored in the biomass, mature ecosystems tend to be in steady state where the net uptake of nutrients is 0. Drever (1994) also points out that the effect of plants will be different depending on whether weathering is limited by rates of weathering reactions (weathering-limited) or by the tranport of weathering products. The Alps, where physical weathering and tranport of weathering products is rapid, are a good example of the former. Here, plants should cause an increase in weathering through increasing the contact time between water rock, hence their effect is to increase chemical weathering rates. The Amazon Basin, where weathering rates are extremely low, is an example of the latter. The combination of subdued topography and dense vegetation limits the rate at which weathering products can be transported, leading to the accumulation of extremely deep (100 m) soils. The thick soil cover isolates the bedrock from incoming precipitation. Thus the effect of plants in the tranport-limited environment is to reduce chemical weathering rates.

The Composition of Rivers

Rivers return rain to the oceans. Except in estuaries, where river and sea water mixes, we think of river water as "fresh water". However, as can be seen from Table 13.8, river water is by no means pure, having on average about 100 ppm dissolved substances. Most, but not all, dissolved substances in rivers are the products of weathering. Some were present in rain to begin with. These "cyclic salts" enter the atmosphere via water droplets produced by breaking waves in the ocean. Since this is a physical process, there is little associated chemical fractionation and hence these seawater-derived salts are present in rain in the same concentrations as in seawater. Rain can also dissolve aerosols from other sources, such as continent-derived dust, compounds transpired by trees, volcanic ash and

	Ave Upper	Ave	Ave	River	Percent	Ave River	River	Suspended/
	Crust	Rain	Dissolv	ed Load	Cyclic	Suspended	Dissolved	Dissolved
						Load	/Crust	
	mg/g	mg/kg	mg/kg	µM/kg		mg/kg	$x \ 10^4$	
Na	28.9	0.9	3.9	169.6	51%	6.6	1.35	0.5909
Mg	13.3	0.15	3.4	139.9	10%	11	2.56	0.3091
Al	80.4	-	0.04	1.5	_	97.4	0.00	0.0004
SiO_2	660	-	7.9	131.5	-	267	0.26	0.0296
Κ	28	0.23	1.45	37.2	35%	19.6	0.52	0.0740
Ca	30	1	14.5	362.5	15%	26	4.83	0.5577
Fe	35	_	0.05	0.9	_	50.7	0.01	0.0010
Cl	_	1.13	4.7	132.6	53%	_	_	_
SO_4	_	2.02	8.5	88.5	52%	_	_	_
HCO ₃	-	_	53.8	882.0	_	_	-	-
Total	_	5.43	98.24	209.6		478.3		

TABLE 13.8 AVERAGE COMPOSITION OF DISSOLVED AND SUSPENDED LOADS OF RIVERS

Average rain is U. S. average from Berner and Berner (1996). Average upper crust from Table 11.4, average river dissolved and suspended loads from Meybeck (1988). Polluted rivers were excluded in the average.

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

gases, pollutants, etc., so that total ionic concentrations in rain are not always in the same proportion as in seawater. In addition to cyclic salts and solutes derived from weathering, rivers also contain a variety of organic compounds derived from biological activity, as well as suspended solids, including both organic matter and mineral grains derived from erosion.

The composition of river water vary widely from the average given in Table 13.8. As we might expect, the major cations are the alkalis and alkaline earths, reflecting their relatively high solubility. Dissolved silica (usually reported as SiO_2 although the dominate dissolved species is H_4SiO_4) is also among the major dissolved species in rivers, though nevertheless depleted compared to it concentration in the crust and in the suspended load. This contrasts with seawater, where biological utilization results in very low SiO_2 concentrations. Concentrations of elements such as Fe and Al are extremely low, as we might expect. In most rivers, Ca^{2+} is the dominant cation and HCO_3 is the dominant anion. This contrasts with rain and seawater, where Na^+ and Cl^- are the dominant ions and reflects the importance of carbonate rock weathering in controlling river composition, as we discuss below. Table 13.9 lists the compositions of a number of rivers, selected to illustrate the range observed. River water composition depends on a number of factors, which we discuss below.

Those elements present in high dissolved concentrations tend to be depleted in the suspended load compared to average upper continental crust, and visa versa. This reflects the weathered nature of the suspended load carried by rivers: the most liable elements have been leached from them before the particles enter the river.

Gibbs (1970) used a plot of $Na^+/(Na^++Ca^{2+})$ versus total dissolved solids (TDS) to divide rivers into 3 classes: precipitation-dominated, rock-dominated, and evaporation/crystallization-dominated (Figure 13.12). He reasoned that rainfall has low TDS and high $Na^+/(Na^++Ca^{2+})$ and concluded that dissolved solids in rivers with these characteristics were derived mainly from precipitation. Reactions between water and rock (weathering) increase TDS and generally lower $Na^+/(Na^++Ca^{2+})$, thus rivers with these properties contained water that had interacted more extensively with rock. Evaporation increases TDS, and because calite and gypsum eventually crystallize from water when evaporative concentration occurs, the $Na^+/(Na^++Ca^{2+})$ ratio increases. Thus he reasoned that water with high TDS and $Na^+/(Na^++Ca^{2+})$ had undergone evaporative concentration and fractional crystallization of calcite and/or gypsum. Many rivers in this category, such as the Colorado and Rio Grande, come from arid regions, where evaporation rates are high.

Stallard and Edmond (1981, 1983), however, analyzed the composition of both precipitation and river water in the Amazon Basin and found that cyclic salts (i.e., those present in rain) contributed

	Negro	L. Amazon	Niger	Changjiang	Mississippi	Nile	Hwanghe	Colorado
				(Yangtze)			(Yellow)	
Na	0.4	1.5	3.5	7.6	11.0	17	55.6	95.0
Mg	0.1	1.0	2.6	7.4	8.9	7	17.7	24.0
SiO ₂	4.1	7.2	15.0	6.9	7.6	21	5.1	9.3
Κ	0.3	0.8	2.4	1.5	2.8	4	2.9	5.0
Ca	0.2	5.2	4.1	30.2	34.0	25	42.0	83.0
Cl	0.3	1.1	1.3	9.1	10.3	8	46.9	82.0
SO_4	0.2	1.7	1.0	11.5	25.5	9	71.7	270.0
HCO ₃	0.7	20.0	36.0	120.0	116.0	134	182.0	135.0
Total	6.3	38.5	65.9	194.2	216.1	224.7	423.9	703.3
Alk (µeq)	30.7	361.6	575.2	1991.8	2161.1	2263.2	3232.7	2449.3
ΣZ^+ (µeq)	43.3	428.0	632.7	2487.9	2982.6	2667.8	6048.5	10383.9
${ m SiO_2}/\Sigma Z^+$	1.58	0.28	0.39	0.05	0.04	0.13	0.01	0.01

Table 13.9. Representative Compositions of Major Rivers

Concentrations in mg/l. Negro and Lower Amazon data from Stallard (1980), Changjiang and Hwanghe data from Zhang et al. (1990), others from Meybeck (1979).

Geochemistry





Figure 13.12. Plot of total dissolved solids vs. $Na^+/(Na^+ + Ca^{2+})$ used by Gibbs (1970) to define "precipitation-dominated", "rock-dominated", and "evaporation-dominated" river compositions. Solid symbols are the rivers in Table 13.9.

only a few percent of the total dissolved solids in rivers plotting within Gibb's "precipitation-dominated" field. Furthermore, these rivers had high ratios of SiO_2 to total cation charge $(SiO_2/\Sigma Z^+)$, whereas rain has very low SiO₂/ ΣZ^+ . Stallard and Edmond (1983) proposed an alternative classification of rivers of the Amazon, based on ΣZ^+ , and argued that *the prin*cipal factors controlling water chemistry are rock type and the intensity of weathering. Berner and Berner (1996) showed that this classification can be extended to all rivers. The classification is as follows.

Transport-Limited Silicate Terranes. Stallard and Edmond's (1983) first category was rivers with $\Sigma Z^+ < 200 \ \mu eg/l$. They found that such rivers drain intensely weathered materials in a transport-limited regime (e.g., Rio Negro). Rivers in this category are enriched in SiO₂, Al, Fe, and organic anions and have a low pH. Perhaps the most significant feature of rivers in this category is their high $(SiO_2/\Sigma Z^+)$. The $SiO_2/\Sigma Z^+$ of the average upper crust is 2.4; silicate sedimentary rocks, derived from rocks that have been partially weathered have a somewhat higher ratio (for example, the ratio in riverine suspended matter is 3.2). Taken to its extreme, weathering leaves a residue of Fe and Al oxides, quantitatively stripping the alkali and alkaline earth cations as well as silica. Thus the most intense weathering would produce wa-

ter with SiO₂/ ΣZ^+ of 2.4 or above, depending on the nature of the bedrock. Where weathering is somewhat less intense, leaving a kaolinite residue for example, the expected ratio would be closer to 1. Whereas in most rivers, the bicarbonate concentration approximately equals the total alkalinity, indicating carbonate species are the principal non-conservative ions, bicarbonate concentrations are much lower than alkalinity in rivers in this category. This is a result of high concentrations of organic anions, which account for much of the alkalinity total anion concentration.

Weathering-Limited Silicate Terranes. Stallard and Edmond (1983) proposed that rivers with ΣZ^+ between 200 and 450 µeq/l drained "weathering-limited siliceous terranes" (e.g., Congo, Amazon). In such regions the rate of erosion exceeds the rate of chemical weathering, and cations are leached from minerals in preference to SiO₂, leading to lower SiO₂/ ΣZ^+ ratios of the water, while the availability of fresh rock results in higher ΣZ^+ and TDS than in transport-limited regimes. Rivers in this category have (SiO₂/ ΣZ^+) of between 0.1 and 0.5.

Carbonate Terranes. Stallard and Edmond's third category was those rivers with ΣZ^+ between 450 and 3000 μ eq/l. Such rivers have low Na⁺/(Na⁺+Ca²⁺) and high Ca, Mg, alkalinity, SO₄ (from oxidation of pyrite in reduced shales). Also, such rivers tend to have 1:1 ratios of Na to Cl and (Mg+Ca) to (HCO₃ +SO₄). These features indicate that these ions are derived from weathering of carbonates and evaporite minerals such as halite and gyspum. Rivers with these properties tend to drain areas

Geochemistry

CHAPTER 13: WEATHERING., SOILS., AND STREAM CHEMISTRY

underlain by marine sediments containing carbonates, reduced shales, and minor evaporites. Most of the worlds major rivers fall in this category (Berner and Berner, 1996).

Evaporite Terranes. The fourth category of Stallard and Edmond was those rivers with $\Sigma Z^+ > 3000$

 μ eq/l. These rivers also tend to have 1:1 ratios of Na to Cl and (Mg+Ca) to (HCO₃ +SO₄). Such high ionic strength rivers drain terranes with abundant evaporites. Rivers in this and the third category have SiO₂/ Σ Z⁺ less than 0.1.

Stallard and Edmond (1983) used a ternary plot of carbonate alkalinity (HCO₃), SiO₂, and Cl + SO₄ (all in μ eq/l) to illustrate the differences between these categories of rivers. A similar plot is shown in Figure 13.13, showing data from river compositions listed in Table 13.9.

Maybeck (1987) also emphasized the importance of rock type in controlling river dissolved load composition. He compared data from French rivers draining monolithic watersheds to data for global He found rivers. that "crystalline" (i.e., igneous and metamorphic) silicate rocks have only a minor influence on dissolved loads. Crystalline rocks account for 34% of the global surface rock outcrop, but only 12% of the global riverine dissolved load. In contrast, evaporites constitute only 1.25% of the outcrop, but 17% of the dissolved load. Carbonate rocks, which constitute 16% of the outcrop, account for half of the global riverine dissolved load. This is consistent with the observation that most major rivers fall in Stallard and Edmond's carbonate category, even through carbonate rocks are rarer than silicate ones. At a more fundamental level, it is the weathering rates of minerals the are important in controlling river chemistry: silicates release dissolved components slowly; carbonates and evaporites dissolve rapidly.

Although it is clear that rock type and weathering regime (transport-limted vs. weathering limited) are the most important factors controlling river chemistry, we should emphasize that both precipitation and evapotranspiration do play some role, albeit a lesser one. The concentration of cyclic salts in rain decrease with distance from the ocean. For exam-



Figure 13.13. Ternary plot of SiO₂, Cl + SO₄, and HCO₃ (equivalent to carbonate alkalinity) used by Stallard and Edmond (1993) to illustrate how rock type and weathering intensity control river composition. (a) shows representative major rivers. Symbols are: diamonds: $\Sigma Z^+ < 200 \ \mu eq$, circles: $200 < \Sigma Z^+ < 450 \ \mu eq$, triangles: $450 < \Sigma Z^+ < 3000 \ \mu eq$, squares: $\Sigma Z^+ > 3000 \ \mu eq$. Filled symbols are rivers listed in Table 13.9. (b) shows where Stalland and Edmond's (1983) 4 categories plot.

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

ple, concentrations of Cl⁻ in rain in the Amazon Basin measured by Stallard and Edmond (1981) were as high as 200 mg/l near the mouth of the Amazon and as low as 2 mg/l at a station on the western edge of the basin 2800 km from the Atlantic coast. For this reason, coastal rivers in areas of high rainfall, such as the southeastern U.S., are the ones most likely to be truely precipitation-dominated (Berner and Berner, 1996). Evaporative concentration, and precipitation of calcite in soils, does indeed influence the composition of rivers draining arid regions such as the Colorado and Rio Grande. Agricultural use of water increases evaporation, as 2/3 or more of the water used in irrigation evaporates. Thus this activity serves to further increase the level of dissolved solids in such rivers.

CONTINENTAL SALINE WATERS

Saline waters result from evaporative concentration of fresh water. Table 12.14 lists the principal components in a number of brines from North America. As we might expect, there is a fair variation in concentrations, but a close inspection reveals something we might not expect: the relative concentrations of these elements also vary greatly. Some are carbonate brines, some are chlorine-rich brines. Some have high sulfate concentrations and some do not. Na⁺ is always a major cation, but relative proportions of Ca²⁺ and Mg²⁺ vary greatly. What leads to this diversity in composition?

As is the case for rivers, the nature of the rock with which the source waters of saline lakes equilibrates is important in controlling concentrations. However, a number of other factors also play a role. Most important, perhaps, is the role of crystallization in magnifying relatively small differences in the composition of source waters. This is very much similar to the role played by crystallization in producing compositional diversity in igneous rocks. Let's consider what happens when dissolved solids in water are evaporatively concentrated.



Figure 13.14. Chemical divides and evolutionary paths of evaporating natural waters. After Drever (1988).

In almost all natural waters, the first mineral to precipitate will be calcite. If the molar concentrations of Ca^{2+} and CO_3^{2-} are equal, the precipitation of calcite does not change the relative concentrations of these two ions. If, however, the concentration of CO_3 exceeds that of Ca, even by a small amount, then crystallization of calcite leads to an increase in the relative concentration of CO_3 and a decrease in the concentration of Ca. As long as evaporation continues, calcite will continue to crystallize and Ca concentrations will continue to decrease, and those of CO_3 to increase. This process leads to a Ca-poor, CO_3 -rich brine. If the opposite is true, namely $[Ca]>[CO_3]$, then Ca increases and CO_3 decreases, leading to a CO_3 poor brine. This illustrates the concept of *chemical divide*, which is shown in Figure 13.14. Depending on initial composition, an evaporating solution will come to forks in the compositional evolution paths that lead to very different compositions.

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

	Kamloops Lake No.	Hot Lake,	Lenore Lake,	Harney Lake,	Alkali Valley,	Abert Lake,	Surprise Valley,	Great Sal Lake,	t Honey Lake,	Pyramid Lake,
	7, B.C.	Wash.	Wash.	Ore.	Ore.	Ore.	Calif.	Utah	Calif.	Nev.
SiO ₂			22	31	542	645	36	48	55	1.4
Ca	tr	640	3	7	-	-	11	241	-	10
Mg	34,900	22,838	20	tr	-	-	31	7,200	-	113
Na	10,900	7,337	5,360	8,826	117,000	119,000	4,090	83,600	18,300	1,630
Κ		891	с	336	8,850	3,890	11	4,070	1,630	134
HCO ₃	2,400	6,296	6,090	4,425	2,510	-	1,410	251	5,490	1,390
CO ₃	-	-	3,020	d	91,400	60,300	644	-	8,020	-
SO_4	160,800	103,680	2,180	1,929	46,300	9,230	900	16,400	12,100	264
CI Total pH	200 209,000	1,688 143,000	$1,360 \\ 18,000$	6,804 22,383	45,700 314,000 10.1	115,000 309,000 9.8	4,110 10,600 9.2	$140,000 \\ 254,000 \\ 7.4$	9,680 52,900 9.7	1,960 5,510
	Carson Sink,	Rhodes Marsh,	Mono Lake,	Saline Valley,	Owens Lake,	Death Valley	Searles Lake,	Soda Lake,	Danby Lake,	Salton Sea,
	Nev.	Nev.	Calif.	Calif.	Calif.	Calif.	Calif.	Calif.	Calif.	Calif.
SiO_2	19	142	14	36	299	-		-	-	20.8
Ca	261	17	4.5	3.1	43	-	16	-	325	505
Mg	129	0.5	34	552	21	150		-	108	581
Na	56,800	3,680	21,500	103,000	81,398	109,318	110,000	114,213	137,580	6,249
Κ	3,240	102	1,170	4,830	3,462	4,043	26,000	tr	-	112
HCO ₃	322	23	5,410	614	52,463	-	-	-	tr	232
CO ₃	-	648	10,300	-	d		27,100	12,053	-	-
SO_4	786	2,590	7,380	22,900	21,220	44,356	46,000	52,026	13,397	4,139
CI	88,900	3,070	13,500	150,000	53,040	140,196	121,000	124,618	119,789	9,033
Total	152,000	10,400	56,600	282,360	213,700	299,500	336,000	305,137	271,200	20,900

Table 13.10. North American Saline Lake Brines

Notes tr: trace. c: Reported Na represents Na + K. d: Reported HCO_3 represents $HCO_3 + CO_3$. From Eugster and Hardie (1978).

The first divide is, as we have seen, calcite. Depending on the path taken at that divide, the next divide is either precipitation of gypsum or precipitation of a Mg mineral, either dolomite or sepio-lite $(Mg_4Si_6O_{15}(OH)_2GH_2O)$, a mixed-layer, 2:1 clay). Gypsum precipitation leads to either a sulfate-depleted or calcium-depleted brine. Although dolomite is the Mg mineral predicted by thermody-namics to precipitate in Mg, CO₃ rich waters, dolomite reactions are so sluggish that sepiolite is actually more likely to precipitate. Both dolomite and sepiolite lead to either magnesium or carbonate depleted brines, because both reactions consume carbonate:

$$Mg^{2+} + 2HCO_3^{-} + 3H_4SiO_4 \rightleftharpoons MgSi_3O_6(OH)_2 + 2CO_2 + 6H_2O$$
 13.27

Note that most of the chloride, and often much of the sulfate and sodium in saline lakes is derived from rain. Sodium and chloride minerals precipitate only at extreme concentrations.

Several other factors play a role in determining the composition of saline waters. One is whether sulfate reduction occurs. The solubility of oxygen decreases with increasing salinity, so reducing conditions are more likely in saline than in fresh waters. Sulfate reduction obviously depletes sulfate (by converting it to insoluble sulfides), but it also increases carbonate alkalinity by production of carbonate by oxidation of organic matter. Another factor is ion exchange and absorption. Many saline lakes are fed by subsurface flow, providing the opportunity for ion exchange with clays and other minerals. This accounts for the low K concentrations of most saline waters. Finally, cyclic wetting and drying can lead to some interesting effects resulting from kinetics. In dry periods, evaporite minerals will precipitate subsurface. Rains in many dry areas come as occasional or rare cloud bursts and

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

may wet the soil long enough to dissolve highly soluble salts such as sodium chloride, but not long enough to achieve equilibrium with slightly soluble salts such as gypsum. This can lead to sulfate concentrations lower than that expected from evaporative concentration of rain water.

References and Suggestions for Further Reading

- Berner, R. A., A. C. Lasaga, and R. M. Garrells. 1983. The carbonate-silicate geochemical cycle and its effect on atmospheric carbon dioxide over the past 100 million years. *Amer. J. Sci.* 283: 641-683.
- Berner, E. K. and R. A. Berner, 1996. Global Environment: Water, Air, and Geochemical Cycles, Upper Saddle River (NJ): Prentice Hall.
- Berner, R. A. 1991. Atmospheric CO₂ levels over Phanerozoic time. Science. 249: 1382-1386.
- Bluth, G. J. S. and L. R. Kump. 1994. Lithologic and climatological controls of river chemistry. *Geochim. Cosmochim. Acta.* 58: 2341-2360.
- Brantley, S. L. and Y. Chen.1995. Chemical weathering rates of pyroxenes and amphiboles. in *Chemical Weathering Rates of Silicate Minerals, Reviews in Mineralogy 31*, ed. A. F. White and S. L. Brantley. 119-172. Washington: Min. Soc. Amer.
- Brimhall, G. H. and W. E. Dietrich. 1987. Constitutive mass balance relations between chemical composition, volume, density, porposity, and strain in metasomatic hydrochemical systems: results on weathering and pedogenesis. *Geochim. Cosmochim. Acta*. 51: 567-588.
- Brooks, R. R. 1972. *Geobotany and Biogeochemistry in Mineral Exploration*. New York: Harper and Row.
- Chou, L. and R. Wollast. 1985. Steady-state kinetics and dissolution mechanisms of albite. *Amer. J. Sci.* 285: 965-993.
- Drever, J. I. (ed.), 1985, The Chemistry of Weathering, Dordrecht: D. Reidel Publ. Co.
- Drever, J. I., 1988. The Geochemistry of Natural Waters, Prentice Hall, Englewood Cliffs, 437 p.
- Drever, J. I. 1994. The effect of land plants on weathering rates of silicate minerals. *Geochim. Cosmo-chim. Acta.* 58: 2325-2332.
- Drever, J. I. and D. W. Clow. 1995. Weathering rates in catchments. in *Chemical Weathering Rates* of *Silicate Minerals, Reviews in Mineralogy 31*, ed. A. F. White and S. L. Brantley. 463-483. Washington: Min. Soc. Am.
- Edmond, J. M., M. R. Palmer, C. I. Measures, B. Grant and R. F. Stalland. 1995. The fluvial geochemistry and denudation rate of the Guayana Shield in Venezuela, Columbia, and Brazil. *Geochim. Cosmochim. Acta*. 59: 3301-3326.
- Eugster, H. P. and L. A. Hardie, 1978, Saline Lakes, in *Lakes-Chemistry*, *Geology*, *Physics*, A. Lerman (ed.), Springer-Verlag, New York, pp. 237-293.
- Feth, J. H., C. E. Roberson and W. L. Polzer. 1964. Sources of mineral constituents in water from granitic rocks, Sierra Nevada, California and Nevada: U. S. Geol. Surv. Water-Supply Pap. 1535
- Garrels, R. M. 1967. Genesis of some ground waters from igneous rocks. in *Researches in Geochemistry*, 2 ed. P. H. Abelson. 405-420. New York: Wiley.
- Garrels, R. M. and C. L. Christ. 1965. Solutions, Minerals and Equilibria. New York: Harper and Row.
- Garrels, R. M. and R. T. Mackenzie. 1967. Origin of the chemical composition of some springs and lakes. *Equilibrium Concepts in Natural Water Systems, Am Chem. Soc. Adv. Chem. Ser.* 67: 222-242.
- Gibbs, R. J. 1970. Mechanisms controlling world water chemistry. Science. 170: 1088.
- Lasaga, A. C., J. M. Soler, G. J., T. E. Burch and K. L. Nagy. 1994. Chemical weathering rate laws and global geochemical cycles. *Geochim. Cosmochim. Acta*. 58: 2361-2386.
- Martin, J.-M. and M. Meybeck. 1979. Elemental mass balance of material carried by major world rivers. *Mar. Chem.* 7: 173-206.
- Meybeck, M. 1979. Concentration des eaux fluviales en éléments majeurs et apports en solution aux océans. *Rev. Géo. Dyn. Géogr. Phys.* 21: 215-246.
- Meybeck, M. 1987. Global chemical weathering of surficial rocks estimated from river dissolved loads. *Am. J. Sci.* 287: 400-428.

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

- Meybeck, M.1988. How to establish and use world budgets of riverine materials. in *Physical and Chemical Weathering in the Global Environment*, ed. A. Lerman and M. Meybeck. 247-272. Dor-drecht: D. Riedel Publ.
- Morel, F. M. M. and J. G. Hering. 1993. *Principles and Applications of Aquatic Chemistry*. New York: John Wiley and Sons.
- Probst, A., E. Dambrine, D. Viville and B. Fritz. 1990. Influence of acid atmospheric inputs on surface water chemistry and mineral fluxes in a declining spruce stand within a small granitic catchment (Vosage Massif, France). *J. Hydrol.* 116: 101-124.
- Richardson, S. M. and H. Y. McSween. 1988. *Geochemistry: Pathways and Processes*, New York: Prentice Hall.
- Robbins, J. A. and E. Callender. 1965. Diagenesis of manganese in Lake Michigan sediments. *Am. J. Sci.* 275: 512-533.
- Schlesinger, W. H. 1991. Biogeochemistry. San Diego: Academic Press.
- Schwartzman, D. W. and T. Volk. 1989. Biotic enhancement of weathering and the habitability of the Earth. *Nature*. 340: 457-460.
- Sposito, G. 1989. The Chemistry of Soils. New York: Oxford University Press.
- Stallard, R. F. 1980. Major Element Geochemistry of the Amazon River System. PhD Dissert. Cambridge: M. I. T.
- Stallard, R. F. 1985. River chemistry, geology, geomorphology, and soils in the Amazon and Oronoco Basins. in *The Chemistry of Weathering*, ed. J. I. Drever. 293. Dordrecht: D. Reidel Publ.
- Stallard, R. F. and J. M. Edmond. 1981. Geochemistry of the Amazon 1. Precipitation chemistry and the marine contribution to the dissolved load at the time of peak discharge. *J. Geophys. Res.* 86: 9844-9858.
- Stallard, R. F. and J. M. Edmond. 1983. Geochemistry of the Amazon 2. The influence of geology and weathering environment on the dissolved load. *J. Geophys. Res.* 88: 9671-9688.
- Stallard, R. F. and J. M. Edmond. 1987. Geochemistry of the Amazon 3. Weathering chemistry and the limits to dissolved inputs. *J. Geophys. Res.* 92: 8293-8302.
- Stumm, W. and J. J. Morgan. 1995. Aquatic Chemistry, New York: Wiley and Sons.
- Velbel, M. A. 1985a. Geochemical mass balances and weathering rates in forested watersheds of the Southern Blue Ridge. *Am. J. Sci.* 285: 904-930.
- Velbel, M. A. 1985b. Hydrogeochemical constraints on mass balances in forested watersheds of the Southern Appalachians. in *The Chemistry of Weathering*, ed. J. I. Drever. 231-247. Dordrecht: D. Reidel Publ. Co.
- Velbel, M. A. 1993. Constancy of silicate-mineral weathering-rate ratios between natural and experimental weathering: implications for hydrologic control of differences in absolute rates. *Chem. Geol.* 105: 89-99.
- White, A. F.1995. Chemical Weathering rates of silicate minerals in soils. in *Chemical Weathering Rates of Silicate Minerals, Reviews in Mineralogy* 31, ed. W. A. F. and S. L. Brantley. 407-461. Washington: Min. Soc. Am.
- White, A. F. and A. E. Blum. 1995. Effects of climate on chemical weathering in watersheds. *Geochim. Cosmochim. Acta*. 59: 1729-1747.
- White, A. F. and S. L. Brantley (ed.), 1995, *Chemical Weathering Rates in Silicate Minerals*, Washington: Mineral. Soc. Am.
- Zhang, J. W., W. Huang, M. G. Lin and A. Zhon. 1990. Drainage basin weathering and majro element transport of two large Chinese rivers (Hwanghe and Changjiang). J. Geophys. Res. 95: 13277-13288.

Problems

1. The dissolution of *talc* may be described by the reaction:

 $Mg_3Si_4O_{10}^{(OH)} + 6H^+ + 4H_2O \rightleftharpoons 3Mg^{2+} + 4H_4SiO_4$

that of *brucite* as:

(1)

Geochemistry

CHAPTER 17: WEATHERING., SOILS., AND STREAM CHEMISTRY

$$Mg(OH)_2 + 2H^+ + 4H_2O \rightleftharpoons Mg^{2+} + 2H_2O$$
(2)

and that of *serpentine* as:

 $Mg_{3}Si_{2}O_{5}(OH)_{4} + 6H^{+} \rightleftharpoons 3Mg^{2+} + 2H_{4}SiO_{4} + H_{2}O$ (3)

The solubility of quartz is 10^{-4} . The ΔG for these reactions (at 25° C) have been estimated as -114.09 kJ, -96.98 kJ, and -193.96 kJ respectively. On a plot of log ($[Mg^{2+}]/[H^+]^2$) vs. log $[H_4SiO_4]$ show the stability fields for talc, brucite, and serpentine.

2. For this question, assume that the dissolution rate of brucite shows the dependence on ΔG given in equation 6.48, that the value of k_i is 10⁻⁶ moles/sec, and a temperature of 25°C. If all other factors remain constant, how will the dissolution rate change as equilibrium between brucite and solution is approached? Make a plot showing the *relative* reaction rate as a function of log ($[Mg^{2+}]/[H^+]^2$).

3. The following is an analysis of the Congo River in Africa (units are mg/l).

pН	Ca ²⁺	Mg^{2+}	Na^+	K^+	Cl-	SO_4	HCO ₃	SiO_2	TDS
6.87	2.37	1.38	1.99	1.40	1.40	1.17	13.43	10.36	54

a. Calculate the *alkalinity* of this water. How does alkalinity compare to the bicarbonate concentration (*HINT:* be sure to use molar units in your comparison).

b. Based on the $Na^+/(Na^+ + Ca^{2+})$ ratio (mg/l) of this analysis, into which of Gibb's 3 categories would you play the Congo River?

c. Calculate $\hat{\Sigma}Z^+$ (μ eq/l). Into which of Stallard and Edmond's categories would this river fall?

d. Calculate the ΣZ^+ /SiO₂ ($\mu eq/\mu M$) ratio for this river. Do all calculations at 25°C.

4. Referring to Figure 6.22, would you expect the Congo River water of Problem 3 to be in equilibrium with gibbsite, kaolinite, muscovite, pyrophyllite, or K-feldspar?

5. Assuming a dissolved aluminum concentration of 0.05 mg/l in Congo River water, calculate the concentration of the various aluminum species at the pH given in problem 2 and the equilibrium constants in equations 6.68 through 6.71.

6. Use the analysis of Congo River water in Problem 3 for these questions.

- a. What is the ionic strength of this solution?
- b. Calculate the Debye-Hückel activity coefficients for this solution.
- c. The bicarbonate concentration reported is actually total carbonate (Σ CO₂). Calculate the actual concentrations of each of the 3 carbonate species.
- d. Calculate the buffer capacity of this water at the given pH and assuming all ions other than the carbonate species remain completely dissociated.

7. Calculate the *speciation* of the ions listed in the Congo River water analysis in Problem 3. Use the stability constants given in Example 6.7.

CHAPTER 14: ORGANIC GEOCHEMISTRY

Chapter 14: Organic Geochemistry

INTRODUCTION

rganic compounds are ubiquitous on the surface of the Earth: some form living organisms, some are a result of pollution, some have been leaked or excreted from living cells, but most are break ^e down products of dead organisms. Organic substances may be either water insoluble or water soluble. Rain typically contains about $1 \mu g/l$ dissolved organic carbon (DOC), the ocean around 0.5 mg/l DOC, and soils up to 250 mg/l. In general, insoluble compounds in soil are more abundant than water soluble ones. In addition to their value as fuel (as peat, coal, petroleum, and natural gas), organic substances are also important in controlling the properties of soil, as weathering agents, and as a significant fraction of surficial carbon, which cycles between the ocean, the atmosphere, the biosphere, soils, and rocks as both inorganic carbon in carbonates and organic carbon. Organic carbon reservoirs (soils and coal, petroleum, etc.) exchange with the atmosphere, where carbon (as CO_2) plays an important role in regulating the Earth's surface temperature. Hence, organic carbon plays an indirect role in climate regulation. All of these provide very good reasons for acquiring a better understanding of organic substances and the role they play in geological processes.

In this chapter, we begin with a brief review of some basic biology and organic chemistry. We then examine how organic compounds are produced and used by organisms. We continue on to survey the distribution of organic compounds in water and soil. We then examine their geochemical properties and their roles as complexing agents and adsorbents. In the final section of this chapter we examine the processes by which sedimentary organic matter evolves into important energy and chemical resources such as coal, gas, and oil.

A Brief Biological Background

Living organisms are the ultimate source of most, but not all^{*}, organic compounds in the environment. While life forms are extremely varied, the basic principles of cell operation are the same in all organisms. For example, all derive the energy they require through oxidation of organic compounds, all contain DNA (deoxyribonucleic acid), which is the blueprint for synthesis of proteins that carry out various intracellular functions, and all use ATP (adnosine triphosphate) in intracellular energy transactions. This commonality suggests all organisms have evolved from a common ancestor.

Over the years there have been a variety of schemes for classifying organisms, and one may all other organisms are eukaryotes.



Figure 14.1. Phylogenetic relationships among organisms. The eubacteria and the archeobacteria constitute the prokaryotes;

^{*} For example, abiogenic methane is found in hydrothermal fluids on mid-ocean ridges.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

still find any one of several in biological textbooks. Figure 14.1 presents a simplified version of the Woese system, based on ribosomal RNA sequencing, which has now achieved wide acceptance. It divides organisms into three kingdoms: *archeobacteria*, *eubacteria*, and *eukaryotes*. Two of these kingdoms consist only of bacteria; all other organisms, including all multicellular organisms, are eukaryotes. Eubacteria and archeobacteria are collectively called *prokaryotes*. They are distinguished from eukaryotes in that their DNA is not contained within a membrane-bound nucleus as is the case with eukaryotes; rather the DNA of prokaryotes is simply concentrated within one or more regions, called nucleoids, of intracellular fluid, or cytoplasm (a protein solution). In contrast, the DNA of eukaryotes is contained within the cell nucleus, usually organized into chromosomes. In both eukaryotes and prokaryotes, messenger RNA (m-RNA) carries information from DNA, in the form of a complement of a portion of a DNA strand, to *ribosomes*, which consist of RNA and proteins, where proteins are synthesized by transcription RNA (t-RNA) from amino acids.

An additional contrast between prokaryotes and eukaryotes is that eukaryotes contain a variety of specialized intracellular structures whereas, except for ribosomes, prokaryotes do not. These include *mitochondria*, where energy is generated by oxidation of carbohydrates, *chloroplasts* (in plants), where *photosynthesis* takes place, the Golgi apparatus, which is involved in modifying proteins, and networks of structural proteins that, among other things, participate in changing the shape of the cell so that it can move. Some of these *organelles*, such as the mitochondria and chloroplasts, have their own DNA with affinities to that of prokaryotes. This and other evidence suggests these organelles may have evolved from prokaryote cells living symbiotically within eukaryote hosts. Eukaryotic cells typically have dimensions of about 10μ , and are therefore much larger than prokaryotes, which generally have largest dimensions of 1μ or less. The eukaryotes may be further divided into single-celled organisms (such as algae and foraminifera) called *protists* and multicellular organisms or *metazoans* (such as us). Prokaryotes, also called *monera*, are divided into Eubacteria and Archeobacteria. These groups differ in the nature of their cell walls and lipid chemistry.

Organisms may also be divided between *autotrophs*, which supply their energy needs through *photosynthesis* or *chemosynthesis*, and *heterotrophs*, whose energy needs are supplied by assimilating organic compounds previously produced by autotrophs. Thus all plants are autotrophs, and all animals are heterotrophs. The bacteria include both autotrophs, such as photosynthetic cyanobacteria (previously called blue-green algae), and heterotrophs.

Perhaps because we are most keenly aware of the presence of bacteria when they infect us, we often think of bacteria as "bad". However, bacteria are ubiquitous and play essential roles in every ecosystem; only a small fraction are pathogenic. All chemosynthetic organisms are bacteria, as are most organisms capable of living without free oxygen (anaerobes). Archeobacteria may be found in extremely hostile environments such as saline lakes and fumaroles. Bacteria play the most important roles in converting the chemical products of life to organic substances found in sediments ("geopolymers") and soils; thus they are of particular interest in geochemistry. Some bacteria reduce sulfate, others oxidize sulfide, some produce methane, others consume it. Some bacteria reduce nitrogen to ammonia, a process called nitrogen fixation, others oxidize ammonia to nitrate (nitrification), yet others convert nitrate to nitrogen (denitrification). Symbiotic bacteria are essential to the digestive systems of higher animals. E. Coli in the human gut, for example, synthesize a number of essential nutrients that are then assimilated through intestinal walls. On the whole, bacteria, and more generally the unicellular organisms, or *microbes*, play a much more important role in biogeochemical cycling and geochemical processes than do the metazoans.

Organic Compounds and Their Nomenclature

Organic chemistry can an intimidating subject involving a bewildering array of compounds, and names for each, whose properties depend as much on the details of their structures as on their composition. A complex nomenclature of organic chemistry has evolved because simply referring to an organic compound by its chemical formula is not sufficient to identify the compound. Compositionally identical compounds can have different structures and different properties. This is, of course, true of

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

inorganic compounds as well, for example, quartz and tridymite or graphite and diamond, but isomers, i.e., compositionally identical but structur-







2,4-dimethylhexane

Figure 14.3. Some simple and branched hydrocarbon chains.





ally distinct compounds, are particularly common among organic compounds. Organic nomenclature is in some sense a language unto itself. Like any language, organic nomenclature has a "grammar". Once this grammar is mastered, the exact composition and structure of a compound can be communicated through its name alone. We will need to concern ourselves with only a part of that language.

We will make one simplification in the brief review of organic chemistry that follows: we will ignore the phenomenon known as *stereoisomerism*. Organic compounds that are otherwise structurally identical but are mirror images of one and other are said to be *stereoisomers*. The difference is physical properties of stereoisomers can be quite small, no more than rotation of the plane of polarized light in an opposite sense, but the difference is often biochemically important. Stereoisomerism is usually less important from a geochemical perspective (but see Example 5.3), so in the interest of clarity, we will not discuss it further.

Hydrocarbons

Hydrocarbons are the simplest group of organic compounds in that they consist only of carbon and hydrogen. We can think of hydrocarbons, and indeed all organic compounds, as consisting of a basic skeleton of carbon atoms to which various functional groups can be attached to form other sorts of organic compounds. The simplest skeleton is that of the *acyclic* or *aliphatic* hydrocarbons, which consist of a straight or branched chain of

M. White

CHAPTER 14: ORGANIC GEOCHEMISTRY

carbon and hydrogen atoms. These compounds are also called *paraf*fins, or alkanes. If they are unbranched, they called normal or n-alkanes. Individual compounds are given names with a prefix indicating the number of CH_x (x = 2, 3) groups present and an ending "-ane". The simplest such molecule is *methane*, CH₄. Ethane consists of 2 CH groups, propane three, and butane four. Beyond that, the root of the name is based on the Greek word for the number of carbon atoms in the chain (i.e., pentane for 5 carbons, hexane for 6, heptane for 7, etc.). Radials formed by removing a hydrogen from a terminal carbon are named by replacing the "-ane" suffix with "-yl", e.g., methyl, butyl. As a group, the radicals formed from alkanes in this way are called alkyls (and thus the group name also conforms to the naming convention). Some simple alkanes are illustrated in Figure 14.2.

Branched chains are formed by replacing one of the hydrogens by Figure 14.4. a carbon atom, to which additional hydrogens are attached. These are named by prefixing the designations of the side chains to the hydrocarbons. name of the longest chain in the formula. A number is prefixed indi-

cating the carbon, counting from the nearest end, to which the secondary chain is attached. An example is 3-methylpentane (Figure 14.3). If more than one secondary group is attached, the groups are listed in alphabetical order, for example, 3-ethyl-4methylhexane shown in Figure 14.3. When several of the same group are attached, a multiplier corresponding to that number precedes the name of the group. An example is 2,4-dimethylhexane, shown in Figure 14.3.

Compounds where all carbon atoms have single bonds to 4 other atoms are said to be saturated (the term saturated arises from carbon being bonded to the maximum possible number of hydrogens; i.e., the carbon is hydrogen-saturated). Carbon atoms that are double bonded are termed olefinic units. Compounds containing one or more pairs of doubly bonded carbons are said to be unsaturated. Unsaturated, unbranched acyclic hydrocarbons having one double bond are named by replacing



Some simple doubly and triply-bonded



Figure 14.5. Two illustrations of the benzene ring, the foundation of aromatic hydrocarbons. In (b) the C and H atoms have been removed for clarity. Since all carboncarbon bonds in the benzene ring are equivalent, (b) is actually a better representation than (a).

the suffix "-ane" by "-ene", e.g., ethylene: $CH_2=CH_2$. A number is used to specify the location (the carbon atom) of the double bond, for example, 2-butene, shown in Figure 14.4. If there are more than 2 double bonds the ending becomes "-adiene", "-atriene", etc. Generic names are alkene, alkadiene, etc. Triple carbon bonds are also possible, in which case the suffix becomes "yne" for a single triple bond. Acetylene^{*} (Figure 14.4) in an example of a compound containing triply bonded carbon. Particularly stable compounds result when carbon bonds are *conjugated*, that is, alternately singly and doubly bonded, e.g., -C=C-C=C-C=C-.

Instead of forming chains, the C atoms may form rings; the resulting compounds are called cyclic hydrocarbons. Naming conventions for the simple groups are similar to those for chains, with the prefix "cyclo-" used to indicate the cyclic nature, e.g., cyclopropane.

A particularly important cyclic structure is the benzene ring, which consists of 6 conjugately bonded carbon atoms lying in a single plane (Figure 14.5). Compounds based on this structure are particularly

As is the case in most languages, there are words in the lexicon of organic geochemistry that do not conform to the standard grammar. Acetylene, named before the names convections were developed, is an example of such an irregular term.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

referred stable and are to as aromatic[†]. Representation of this structure as alternating single and double bonds is not entirely accurate. The carboncarbon bond in a saturated alkane such as ethane is 1.54 Å in length; the double bond in ethylene is 1.33Å in length. All carbon-carbon bonds in the benzene ring are found to be intermediate in length (1.40Å). Thus bonding is delocalized, that is, all carbon-carbon bonds in the ring are of approximately equal strength and the double bonds appear to be shared among all carbon atoms in the ring. This delocalization is responsible for the particularly stability of this structure. Several rings may be joined to form *polyaromatic units*.

FUNCTIONAL GROUPS

From these basic hydrocarbon structures, a great variety of other organic compounds may be formed by attachment of atoms or functional groups by replacement of hydrogen or carbon atoms. The basic naming conventions discussed above for hydrocarbons also apply to these other organic molecules. Additional names, prefixes, and suffixes are used to indicate the presence of Figure 14.6. Important functional groups attached groups replacing hydrogen or other atoms re- found in organic compounds. placing carbon in the basic structure.

Group	Name	Resulting Compound
_ OH	Hydroxyl	Phenols (Aromatics) Alcohols (Aliphatics)
_ 0 _	Охо	Esters
R-C=0	Carbonyl	Aldehydes Ketones, Quínones
_C=0 ∣ OH	Carboxyl	Carboxaylic Acid
$-NH_2$	Amino	Amines
-C=0 H_2	Amido	Amides
— SH	Thio	Thiols

The most important functional groups are summarized in Figure 14.6. The hydroxyl (OH) unit may be attached to an aromatic ring to form *phenols* or to aliphatic units to form *alcohols*. The biologically important carbohydrates, e.g., sugars and starches, are compounds with the general formula of CH₂O. They are either aliphatic or cyclic hydrocarbons in which a number of the hydrogens has been replaced by OH. The carbonyl group (C=O) forms aldehydes when the "R" in Figure 14.6 is H, ketones when "R" is either an aliphatic or aromatic group, and quinones when the carbon in the group is incorporated into an aromatic ring.

A particularly important functional group from a geologic perspective is the *carboxylic acid* group: COOH. The suffix "-oic acid" is used to designate compounds formed by carboxylic acid groups: for example, menthanoic acid is a carboxyl group with an attached H, ethanoic acid is a carboxyl group with an attached CH₃, benzoic acid is a carboxyl group with an attached benzene ring. Many of the carboxylic acids also have more familiar names, for example, ethanoic acid is more commonly called acetic acid, methanoic acid is also called formic acid. Two carboxyl units bound together form oxalic acid. A compound formed by replacement of the dissociable hydrogen in the carboxyl group with some other group is designated by the suffix "-ate".

Carboxyl groups attached to hydrocarbon chains form *fatty acids*, which are important components of lipids. If one or more of the hydrogens attached to the carbon chain or ring is substituted by a hydroxyl group, the compound is known as an hydroxy acid (for example, salicylic acid illustrated in Figure 14.7). If a doubly bonded oxygen is substituted for two hydrogens attached to the chain (forming a *carbonyl group*), the compound is known as a *keto acid*. Hydroxy acids such a lactic acid and keto acids such as pyruvic acid are important in both the Calvin Cycle, by which autotrophs synthesize organic compounds, and the Krebs Cycle, by which organisms oxidize organic compounds to release energy.

[†] Some aromatic compounds, such as benzene, toluene, and a variety of chorinated phenols, are highly toxic. Because of the stability imparted by the conjugate bonding, they are particularly environmentally hazardous.

CHAPTER 14: ORGANIC GEOCHEMISTRY

Esters form by combining a carboxylic acid and an alcohol. In the reaction between these two, the OH is removed from the carboxyl group and the H removed from the alcohol, leaving the two units bound by an oxygen atom[†]. An example of such a reaction is the formation of pentyl acetate (which gives bananas their familiar odor) from acetic acid and pentyl alcohol:



Another geologically and biologically important functional group is the *amino group*, NH_2 (the name being derived from ammonia, NH_3). *Amino acids*, the building blocks of proteins, consist of molecules containing both amino and carboxylic groups. Other important functional groups include the carbonyl group, consisting of a carbon that is double bonded to oxygen, and thiol groups, where S replaces O in



Figure 14.7. Some examples of compounds formed by substituting functional groups for hydrogen in basic hydrocarbon structures.

the OH group (i.e., SH). Figure 14.7 illustrates a few of the important compounds formed by replacement of hydrogen in basic hydrocarbons by one or more of these functional groups.

Organic compounds can contribute significantly to the pH balance of natural waters. The acid-base properties of organic compounds depend on the attached functional groups rather than the basic structure. Functional groups may be classified as acidic or basic depending on their tendency to give up (acidic) or accept (basic) hydrogen ions. Carboxyl groups tend to be strongly acidic, phenols and quinones tend to be mildly acidic. Alcohols, ethers, ketones, and aldehydes are generally classified as neutral. The nitrogen in amines and amides binds free hydrogen ions fairly easily, hence these groups are basic.

The acid-base properties of organic substances are also directly related to their solubility in water. Because water can more easily dissolve ionic substances than neutral ones, strongly basic or strongly acidic compounds (those that have given up or accepted a proton) tend to be more soluble than neutral compounds. Thus carboxylic acids are very water soluble, amines somewhat less soluble, and neutral compounds such as esters and ether least soluble.

It is also possible to substitute other elements for carbon in the basic hydrocarbon structure; such atoms are known as *heteroatoms*. Examples are illustrated in Figure 14.8. The pyranyl group is a particularly important one because it forms the basis of many cyclic carbohydrates, the pyridinyl group is an important component of nucleic acids.

⁺ Such a reaction, in which two molecules combine to form a larger molecule with the elimination of a small molecule (H_2O in this case), is called a condensation reaction or simply *condensation*. The reverse of this reaction is termed *hydrolysis*.

CHAPTER 14: ORGANIC GEOCHEMISTRY

Short-Hand Notations of Organic Molecules

For both aliphatic and cyclic molecules, the number of carbons in the ring or chain is commonly denoted with C followed by a subscript corresponding to the number of carbons, e.g., C₆ for hexane. It is often convenient to use this notation in condensed structural formulae of long-chained aliphatic compounds. The basic repeating unit of such chains is CH₂. The number of repetitions can be expressed by enclosing the CH₂ in parentheses followed by a subscript indicating the number or for example repetitions, $(CH_2)_{6}$. Groups placed on the ends then precede and follow. Thus for example stearic acid, a fatty acid consisting of a 17 carbon chain with a hydrogen on one end and a carboxyl group on the other, can be denoted as $CH_3(CH_2)_{16}C(O)OH$. If a double carbon bond occurs, this is designated by CH=CH inserted at the appropriate place. Thus palmitoleic acid, a common unsaturated fatty acid with the formula CH₃(CH₂)₅CH=CH (CH₂)₇C(O)OH is a 15 carbon aliphatic compound with a olefinic unit between the 7th and 8th carbons, an H on one end, and a carboxyl acid group on the other. If an attached group occurs in the middle part of the chain, then the formula for the group is inserted in parentheses after the appropriate number of repetitions of the CH_2 unit, and a second $(CH_2)_n$ follows, n indicating the number of subsequent repetitions. Thus the formula for 3-methylhexane would be $CH_3(CH_2)_2(CH_3)(CH_2)_2CH_3$.

There are also several conventions for illustrating the structure of organic molecules. We have already seen one: the hexagon with an enclosed circle to denote the benzene ring (Figure 14.5). As in this illustration, carbons and hydrogens bonded to them are often omitted from representations of organic



Geochemistry

(a 3-n-hexene



(b Retinol (Vitamin A)

Figure 14.9. Short-hand structural representation of 3-*n*-hexene and retinol. Carbon and hydrogen atoms are not shown in the short-hand representations. Carbons occur a t each joint in the lines as well as at ends of lines. Number of hydrogens bonded to each carbon is equal to 4 less the number of other bonds shown.



Figure 14.8. Examples of functional groups formed replacing one carbon atom in the cyclic skeleton with atoms of O, N and S. The names of the functional groups are given in the figure; names of the resulting compounds are (a) pyran, (b) pyridine, and (c) thiophene.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

molecules. We infer a carbon at each bend of the line as well as hydrogens bonded to it. A similar short-hand may be used for aliphatic molecules as well, as illustrated in Figure 14.9. We may summarize these abbreviated illustrations as follows. A carbon is inferred at each change in angle of the line as well as at the ends. Double bonds are indicated by double lines. Each carbon may have one or more hydrogens bonded to it. Since carbon always forms 4 bonds, the number hydrogens is easily deduced as 4 less the number of other bonds shown.

Biologically Important Organic Compounds

Obviously, life is based upon organic compounds. A remarkable variety of organic compounds can be found in even the simplest cells. Many of these compounds are incredibly complex, commonly having molecular weights exceeding 10,000 daltons. The most important of these compounds can be divided into a few fundamental classes: carbohydrates, proteins, lipids, nucleotides, and nucleic acids. Essentially all naturally occurring organic compounds originate from these classes, and most from the first three. Here we briefly review the chemistry of these biologically important compounds.

Carbohydrates

Carbohydrates are the most fundamental of the biologically important compounds in the sense that they are the direct products of photosynthesis. Virtually all life is ultimately dependent on photosynthesis, and virtually all other compounds necessary for life are synthesized in cells from carbohydrates. Some organisms, such as those of hydrothermal vent communities, depend on chemo-



Figure 14.10. Some simple sugars. (a) linear glucose, (b) cyclic glucose, (c) fructose, and (d) sucrose, a disaccharide formed by condensation of glucose and fructose.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

synthesis rather than photosynthesis, but carbohydrate is also the immediate product of chemosynthesis. Thus carbohydrates may be regarded the fundamental substance of life.

Carbohydrates, as we mentioned earlier, are related to hydrocarbons by substitution of hydroxyl groups for hydrogen atoms. Two of the simplest carbohydrates are the sugars glucose and fructose, both of which have the composition $C_6H_{12}O_6$. Both can exist as straight chains or cyclic structures (Figure 14.10), though the cyclic structures predominate. Glucose and fructose are examples of *monosaccharides*, the mono-prefix indicating they consist of single chains or rings. General names for these compounds are formed from the Greek prefix corresponding to the number of carbons and the suffix *-ose*. Thus fructose and glucose are pentose sugars and ribulose (a building block of nucleic acids) is a hexose sugar. Two monosaccharide units may be linked together by elimination of H_2O to form a disaccharide (another example of a *condensation* reaction). Sucrose, or common table sugar, is the condensation product of glucose and fructose.

Molecules consisting of 10 or more monosaccharide units are called *polysaccharides*. Among the biologically most important polysaccharides are *cellulose* and *starch*. Cellulose, the basic structural material of plants, has the general formula of $(C_6H_{10}O_5)_n$ and consists of long (i.e., $n \ge 10,000$) chains of glucose units. The chains are cross-linked to each other by hydrogen bonds, building up fibrils. In cell walls of vascular plants these fibrils are interwoven or cemented by other polysaccharides called hemicellulose. Cellulose is an example of a *homogeneous polysaccharide*, i.e., one that is formed by linkage of a single kind of monosaccharide. *Chitin*, the material forming hard structures in arthropods, molluscs, and some fungi and algae, is also a homogeneous polysaccharide. It is related to cellulose by replacing one of the hydroxyl groups with a amido group. Starch, which serves to store energy in plants, is also a $(C_6H_{10}O_5)_n$ polysaccharide in its simplest form. Amuylose, for example, which

represents about 25% of starch, consists of long, unbranched chains of (cyclo-) glucose. Starches, however, also include heterogeneous polysaccharides, i.e., polysaccharides containing more than one kind of monosaccharide unit. Water soluble starches consist of relatively short chains ($n \approx 25$); insoluble starches are typically longer, up to 500,000 daltons. Glycogen, a polysaccharide used for energy storage in animals, is closely related to starch. From a geochemical perspective, an important difference between cellulose and starch is that the former is much more stable and less readily metabolized. Although organisms generally store energy in the form of complex carbohydrates and lipids, these are always first converted to back to glucose before oxidation releases this energy.



Figure 14.11. (a) Three of the twenty common amino acids that can combine to form proteins. (b) Peptide segment consisting of glycine, serine and tyrosine bound by peptide linkages (dashed boxes). Peptide linkage occurs between amine and carboxyl groups with the elimination of H_2O .

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

Nitrogen-Bearing Organic Compounds: Proteins, Nucleotides, and Nucleic Acids

Amino acids are the basic building blocks of proteins. There are 20 common amino acids, the essential characteristic of which is the presence of both an amine and a carboxylic group (Figure 14.7). The simplest amino acids are *glycine*, which consists of the amino acid group with a hydrogen at the free position (Figure 14.11), and alanine, which has CH_3 at the free position. Amino acids may be characterized as either neutral, acidic, or basic. *Acidic* amino acids have an additional carboxylic group, which acts as a proton donor. *Basic* amino acids, such as lysine (COOHCH((CH_2AH_2) NH₂), have an additional amine group, which can act as a proton acceptor. Neutral amino acids, such as glycine and alanine have equal numbers of carboxylic and amine groups.

Proteins are formed by condensation of many amino acid units into polymers called peptides. The simplest proteins consist of 40 amino acid units, the most complex ones consist of more than 8000 units. With 20 basic building blocks, the possible combinations are virtually limitless, making the diversity of life possible. The condensation reaction forming peptides consists of linking the carboxylic group of one amino acid to the amine group of another with the elimination of water, as illustrated in



Nicotinamide adenine dinucleotide (R=H)

(c)

Nicotinamide adenine dinucleotide phosphate $(R = PO_3^{2-})$

Figure 14.12. (a) The structure of pyrimidine and purine groups, essential components of nucleotides. (b) Structure of the nucleotide adenosine triphosphate (ATP). (c) Structure of NAD or nicotinamide adenine dinuclotide (when the radical labelled R is H) and NADP or nicotinamide adenine dinucleotide phosphate (when R is phosphate).

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

Figure 14.11. This bond is referred to as a peptide linkage. The biosynthesis of proteins is performed in ribesomes by RNA molecules. The genetic information contained in DNA is essentially a set of blueprints for protein synthesis.

There is an immense variety of proteins, and they play a wide variety of roles in life. Proteins such as *collagen* (bone) and *keratin* (hair, claws) are the essential structural and connective (e.g., tendons) materials of higher animals. It is contraction of proteins in muscles that provide movement. Enzymes, which are cell's catalysts, are often proteins, as are antibodies, which play an essential role in the immune system. Proteins also act to store and transport various elements and compounds; hemoglobin is a good example of such a protein. Hormone proteins serve as messengers and regulators.

This variety of function results from primary, secondary, and tertiary structures. The primary structure of proteins depends both on the kind of amino acid units composing them and on the order in which these units occur. These primary structures may then be folded. The folds are locked in by hydrogen bonds between adjacent parts of the chain (secondary structures). Other proteins are twisted into α -helix structures. Folding of the α -helix results in tertiary structures. All these structures contribute to the biological function of the protein.

Nucleotides are based on pyrimidine or purine groups (Figure 14.12). The nucleotides adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide phosphate (NADP), illustrated in Figure 14.11, play key roles in both the creation and storage of chemical energy (photosynthesis) and its transfer and release (respiration) in organisms. Another nucleotide, nicotinamide adenine dinucleotide (NAD) plays an important role in respiration. ATP can be formed from ADP (adenosine diphosphate) by the addition of an inorganic phosphate ion. This process, called *phosphorylation* involves a free energy change (Δ G) of about +40 kJ/mol. That energy is readily liberated on demand by the reverse reaction. Thus ATP serves as a general carrier of free energy within cells. NAD and NADP and their reduced equivalents (NADH and NADPH) serve as redox couples and as transport agents of reduced hydrogen.

Nucleic acids are related to proteins in the sense that are nitrogen-containing polymers built from a variety of fundamental groups. The amine, phosphate (PO $_4^3$), and pyridinyl (Figure 14.7b) groups are among the essential ingredients of nucleic acids. Unlike proteins, the carboxyl group is generally not present. The nucleic acids DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) contain the genetic code and control protein synthesis within the cell. DNA consists of two backbone strands of a polymer made up of phosphate and the pentose sugar β -D-ribfuanose connected to each other by pairs of the 4 nucleotides, cytosine, thymine, adenine, and guanine. The genetic information is encoded in the sequence of pairs.

Lipids

Unlike carbohydrates and proteins, *lipids* are defined not by their composition and structure, but by their behavior: lipids are those substances that are water insoluble but are soluble in organic solvents such as chloroform, toluene, acetone, and ether. Lipids include fats, oils, waxes, steroids, and other compounds. Fats and oils are generally *triglycerides*, which are esters of 3 *fatty acids* and *glycerol*, an alcohol (Figure 14.13). Fatty acids are straight-chained (aliphatic) carboxylic acids (i.e., an alkane with a carboxyl group at one end). They typically range in length from C_{12} to C_{36} . Because they are generally formed by successive additions of aceytl (C_2) units, fatty acids have predominantly even numbers of carbon atoms. Unsaturated fats, such as *oleic acid* ($CH_3(CH_2)_7CH=CH(CH_2)_7$ CH(O)OH), predominate in plants, whereas saturated fats, such as *stearic acid* ($CH_3(CH_2)_{16}$ CH(O)OH), predominate in animals. In *phospholipids*, one of the fatty acids in the triglyceride is replace by a phosphate unit, which is in turn often linked to a nitrogen base. Many *glycolipids*, which are combinations of lipids and carbohydrates, are triglycerides in which one of the fatty acids is replaced by a sugar. Ether lipids are glycerides formed from straight-chained alcohols, called *nalkanols* or fatty alcohols, rather than fatty acids.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

$$\begin{array}{c} H & O \\ & H & -C & -O & -C & -CH_3(CH_2)_{16} \\ & H & -C & -O & -C & -H \\ & H & -C & -O & -C & -H \\ & H & -C & -O & -C & -CH_3(CH_2)_{16} \\ & H & O \end{array}$$

Figure 14.13. A triglyceride fat formed from the alcohol glycerol and three molecules of stearic acid.

cause they are formed by *decarboxylation*^{*} of fatty acids. *Cutin*, which forms protective coatings on plants, is a polymerized hydroxy fatty acid (commonly C_{16} or C_{18}).

Another important class of lipids is the terpenoids. Terpenoids display a great diversity of structure, but the basic unit of all terpenoids is the isoprene unit, a branched, 5-carbon chain with a

methyl group attached to the second carbon atom (Figure 14.14). Terpenoids are named on the basis of the number of isoprene units present: monoterpenoids have 2, sesquiterpenoids have 3, diterpenoids have 4, triterpenoids have 6, tetraterpenoids have 8. Molecules consisting of more than 8 isoprene units are termed polyterpenoids. Terpenoids may be cyclic or acyclic, saturated or unsaturated. Among other things, terpenoids serve as pheromones (scents, attractants), hormones, antibiotics, resins, and vegetable and animal oils. The noncyclic terpenoids, often referred to as acyclic *isoprenoids*, are important from a geochemical perspective because they are common component of sedimentary organic matter. Furthermore, they are also found in petroleum, apparently having survived diagenesis. Many isoprenoids are derivatives of phytol, a diterpenoid that forms part of chlolophyll-a, a key compound in photosynthesis. During diagenesis, phytol is converted to phytane (Figure 14.14c and 14.14d) through the loss of



Cholesterol

e

Figure 14.14. Terpenoids. a. Isoprene, the building block of all terpenoids. b. Menthol, a simple cyclic monoterpenoid. c. Phytol, an acyclic diterpenene. d. Phytane, an acyclic diterpenane derived from phytol. e., Cholesterol, illustrating the ring system shared by all steroids. The carbon atoms at the apices of the phenols and the hydrogens bound to them are not shown.

Waxes are a mixture of many constituents. Among the most important are wax esters, which are esters of straightchained fatty acids and fatty alcohols, both of which have chain lengths generally in the range of C_{24} to C_{28} . The fatty alcohols also have predominantly even number of carbon atoms because they are synthesized from fatty acids. Plant waxes also contain long, straight-chained hydrocarbons (C_{23} to C_{33}). These hydrocarbons typically have odd numbers of carbons be-

^{*} Decarboxylation is a process whereby a CO_2 molecule is lost from an organic compound. Carboxylation is the addition of a CO₂ molecule to an organic molecule.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

the OH functional group and hydrogenization (loss of the C-double bond). Another important isoprenoid is squalene, a triterpenoid ($C_{30}H_{50}$). Squalene is abundant in both plants and animals. Among other things, it is the precursor of a biologically important class of compounds called *steroids*. *Cholesterol* is a triterpenoids that occurs in both plants and animals. Most cholesterol is found in cell membranes and in lipoproteins. Cholesterol also serves as the precursor of other animal steroids. A few terpenoids are illustrated in Figure 14.14.

Like proteins, lipids play a variety of roles in life. Phospholipids are the primary constituents of cell membranes. Fats and oils serve to store energy. Lipids include pigments that are essential in photosynthesis in plants (e.g., chlorophyll) and vitamin A production in mammals (e.g., carotenoids). Other lipids form protective barriers (e.g., cutin), act as sex pheromones (i.e., providing scent) or homones, or assist in digestion. From a geochemical perspective, lipids are important because they are thought to play a dominant role in petroleum formation. Furthermore, many oils and pigments are unique to groups of organisms. Some of these, isoprenoids in particular, can survive diagenesis, and can be used in reconstructing the origin of hydrocarbon deposits. Such compounds are called *biomarkers*.

Lignin and Tannins

Lignin is another important structural material in higher plants. It is a rigid, high molecular weight polyphenol. As such, it is quite stable and resistant to bacterial decomposition. As a result, it is a very important contributor to soil and terrestrial sedimentary organic carbon. Because marine plants are almost exclusively algae, which do not produce lignin, lignin is a much less important contributor to marine organic carbon.

Tannins are another class of compounds found only in higher plants. They occur predominantly in bark and leaves and function to make the plant less palatable to herbivores. Like lignin, they are high molecular weight (500-3000) polyphenol compounds. Functional groups include carboxyl as well as OH.

THE CHEMISTRY OF LIFE: IMPORTANT BIOCHEMICAL PROCESSES

Our main concern in this chapter is the role played by organic compounds outside living tissue. Nevertheless, given the remarkable impact of life upon the surface of the Earth and its chemistry, and the interplay and constant exchange between living and non-living organic matter, a very brief survey of the more important biochemical processes is appropriate at this point.

As was mentioned earlier, autotrophs produce the energy they need by synthesis of organic compounds from inorganic ones. Most autotrophs are phototrophic: they use light energy to synthesize organic compounds, a process known as photosynthesis. Some bacteria use chemical energy, derived from the oxidation of H_2S or other reduced species, to synthesize organic compounds in a process called chemosynthesis.

Photosynthesis

Like most biochemical processes, photosynthesis is a complex one, involving many steps catalyzed by a variety of compounds. The details of the photosynthetic processes also vary somewhat between photosynthetic bacteria and true plants. In true plants, photosynthesis takes place within specialized intracellular organelles called chloroplasts. In prokaryotic bacteria such specialized intracellular units are absent. In these organisms, the site of photosynthesis may internal membranes within the cell protoplasm, as in the case of the cyanobacteria, or on the cell's plasma membrane, as in the case of the halophiles.

We can describe photosynthesis with the following reaction:

$$CO_2 + 2H_2A + light \rightarrow CH_2O + H_2O + 2A$$
 14.1

In photosynthetic eukaryotes and cyanobacteria, *A* is oxygen, and hence the hydrogen donor is water. However, in anaerobic photosynthesis, carried out by some photosynthetic bacteria (for example the

CHAPTER 14: ORGANIC GEOCHEMISTRY

purple sulfur bacteria), *A* may be sulphur, so that H_2S is the hydrogen donor; alternatively, hydrogen may be taken up directly. In oxygenic photosynthesis, carried out by plants and some photosynthetic bacteria, molecular oxygen is a by-product of photosynthesis and it is this process that is to a large degree responsible for free oxygen in the atmosphere. Some bacteria can fix CO_2 without liberating oxygen, a process called non-oxygenic photosynthesis. Our brief description will focus on oxygenic photosynthesis.

Oxygenic photosynthesis can be divided into a light and a dark stage. The former involves two separate photoreactions. The first of these, governed by Photosystem II (or PS II) is the photodissociation of water. This reaction can be described as:

 $2H_2O \rightarrow 4H^+ + 4e^- + O_2$

Among the enzymes mediating this process is a Mn-bearing protein, which takes up the liberated electrons by oxidation of Mn. The hydrogen ions and electrons produced by PS II travel along distinct chemical pathways, which in the case of the latter include Fe- and Cu-bearing proteins, to the site of the dark reactions. The energy from the electrons liberated in PS II is used for phosphorylation of ADP to ATP. In Photosystem I (PS I), NADP⁺ is reduced to NADPH. In the subsequent dark reactions, this NADPH acts as an electron donor in the reduction of CO_2 , and the energy for this reaction is supplied by ATP.

Energy to drive both PS I and PS II is captured by chlorophyll (Fig. 14.15), a member of a class of lipids called *porphyrins*, or some other light absorbing pigment. There are several varieties of chlorophylls, several of which are generally present within a given plant. Chlorophyll-a is the principal photosynthetic pigment in plants and cyanobacteria. Some other bacteria, such as the green sulfur bacteria, utilize a closely related substance called bacteriochlorophyll. The halophiles, members of the archeobacteria, use retinol (the same light-sensitive pigment in the human retina) rather than chlorophyll to capture light energy. All chlorophylls are strongly absorb light in the red and blue parts of the visible spectrum (the green color of plants results from a lack of absorption of green light).

It is in the dark stage of photosynthesis that carbohydrate is actually synthesized. At this point, there is a divergence in the chemical pathways. In C_3 plants the initial carbohydrate produced (3-

phosphoglycerate) is a 3-carbon chain. This process, known as the Calvin Cycle (illustrated in Figure 14.16) is used by all marine plants and about 90% of terrestrial plants. The first step is the phosphorylation, or the addition of a phosphate group. In this reaction, ribulose 5-phosphate, a C₅ sugar containing 1 phosphate, is converted to ribulose 1,5-bisphosphate, with the additional phosphate coming from the ATP generated during the light stage. In the next step, an enzyme called *ribulose* bisphosphate carboxylase oxygenase (RUBISCO) catalyzes a reaction in which ribulose 1,5-bisphosphate reacts with one molecule of CO_2 to produce 3 molecules of 3-phosphoglyceric acid. Then an additional phosphate is added to each of these



Figure 14.15. Structure of chlorophyll-a. The linear chain is a phytol unit.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY



Figure 14.16. The Calvin cycle of dark reactions for the production of glucose in C₃ plants.

molecules to form 1,3-bisphosphoglycerate. This phosphate is then replaced by hydrogen supplied by NADPH to form glyceraldehyde 3-phosphate. It is in this step that the new CO_2 in the phosphoglycerate is reduced. The ΔG for reduction of CO_2 to CH_2O is about +480 kJ/mole. Most of this energy is supplied by the oxidation of two molecules of NADPH to NADP⁺. Most of the resulting triose phosphate is converted back to ribulose 5-phosphate for further synthesis, but some is used for synthesis of fructose or glucose. Monosarccharides produced in this way, if not immediately metabolized, are eventually converted to more complex carbohydrates or other essential compounds.

The other photosynthetic pathway is the Hatch-Slack Cycle, used by the C_4 plants, which include hot-region grasses and related crops such as maize and sugarcane. These plants use *phosphoenol pyruvate carboxylase* (PEP) to initially fix the carbon and form oxaloacetate, a compound that contains 4 carbons. CO_2 is fixed in outer mesophyll cells as oxaloacetate and is then transported, either as malate or asparatate, to inner "bundle sheath" cells, where it is decarboxylated and refixed in the Calvin Cycle (Fig. 14.17). C_4 photosynthesis appears to have evolved only recently, the oldest known C_4 plants are of late Miocene age (though there has been speculation they may have first evolved in the Cretaceous). Because the C_4 pathway is somewhat more efficient at low ambient con-

CHAPTER 14: ORGANIC GEOCHEMISTRY

centrations of CO_{2} there has been speculation their that appearance may reflect an evolutionary response to low atmospheric CO₂ concentrations of the late Cenozoic.

From a geochemical perspective, the most important aspect of the difference between C_3 and C_4 plants is the difference in carbon



isotope fractiona- Figure 14.17. The Hatch-Slack photosynthetic cycle utilized by C_4 plants.

tion during photosynthesis. The C₃ pathway produces a δ^{13} C fractionation of from -20 to -30‰, where as the C₄ pathway produces a much smaller fractionation of about -13‰. Further aspects of this are discussed in Chapter 9.

Some bacteria, the chemoautotrophs, obtain the energy necessary to reduce CO_2 to carbohydrate from oxidation of sulfide to sulfate or of ferrous iron to ferric iron. As in the light stage of photosynthesis, the oxidation reactions, catalyzed by an enzyme called dehydrogenase, provide both protons (for reduction of CO_2) and energy in the form of electrons. Though they require reduced species for energy, chemosynthetic bacteria are aerobic. Thus they live at boundaries between oxic and anoxic conditions.

RESPIRATION

Respiration is the process by which organic carbon is oxidized to H_2O and CO_2 , liberating energy in the process. Polysaccharides serve as the primary energy storage in both plants and animals. In plants, these generally take the form of starches, while in animals glycogen serves as the primary energy store. In both cases, single glucose molecules are first liberated from these chains before being oxidized. This is accomplished through the catalytic action of an enzyme[‡]. There are a very large number of enzymes, many designed to catalyze the breaking (or formation) of a single specific bond. Glucose molecules liberated from complex carbohydrates in this fashion are then converted to two molecules of pyruvic acid in a multi-step process called *glycolysis*. The first steps involve phosphorylation and conversion of glucose to fructose diphosphate. The fructose diphosphate is then cleaved into two triose phosphates and converted to pyruvic acid (CH₃COCOOH) upon removal of the phosphates. This liberates 2 hydrogen atoms, which are consumed by the reduction of NAD⁺ to NADH, per glucose molecule. The reduction of each NAD⁺ consumes about 219 kJ/mol. This energy can be later used to produce ATP. Although ATP is consumed in the initial phosphorylation steps, it is produced in subsequent steps and the entire process results in a net production of 2 ATP molecules per glucose molecule (and an additional energy gain of 80kJ/mole).

[‡] Enzymes are denoted by the ending "-ase". Classes of enzymes that catalyze hydrolysis reactions are named by combining the name of the class of compounds whose hydrolysis they catalyze with the suffix "-ase". Thus lipases catalyze the hydrolysis of lipids, amylases catalyze the hydrolysis of starches, etc.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

Glycolysis does not release CO₂. Energy released in the process comes primarily from oxidation of hydrogen in the sugar to hydroxyl and resulting conversion of a sugar (glucose) to a carboxylic acid (pyruvic acid). Thus glycolysis releases only a small fraction of the energy stored in the glucose molecule. Under aerobic conditions, further energy may be obtained by oxidation of the pyruvic acid molecule. This oxidation of pyruvic acid occurs within the mitochondria of eukaryote cells in a process called the *Krebs Cycle* (illustrated in Figure 14.18), or alternatively the *Citric Acid Cycle* (in con-



Figure 14.18. Illustration of the Krebs, or Citric Acid, Cycle, in which organic matter is oxidized by organisms to produce energy in the form of ATP. New pyruvate enters the cycle at the top, where one carbon removed from the structure. The remaining two carbons join with a four carbon oxaloacetate chain from a previous cycle to form citric acid, from which 2 additional carbons are successively stripped.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

trast, glycolysis occurs within the general cell cytoplasm). In the Krebs cycle, pyruvic acid is first oxidized to acetyl coenzyme A (a reactive thioester) plus CO_2 . This acetate is then combined with an oxaloacetate produced in an earlier Krebs cycle to form citrate. Two additional CO_2 molecules as well as additional hydrogens are then successively stripped, ultimately producing an oxaloacetate. Thus one pass through the Krebs cycle completely oxidizes one pyruvic acid molecule, producing 4 molecules of NADH, one of FADH₂ (FAD, flavin adenosine dinucleotide, is an oxidizing agent similar to NAD), and one of ATP. The hydrogen stored in FADH₂ and NADH are passed along through a series of protein carriers until it is finally combined with free oxygen to form H₂O. The last step generates 3 molecules of ATP per water molecule produced. In all, 36 molecules of ATP can be produced per molecule of glucose, 34 of these being produced in the Krebs cycle.

Fats serve as secondary energy storage in both plants and animals. Because they contain less oxygen, they store approximately twice as much energy per unit weight as carbohydrates. Respiration of fats involves breaking them into their component fatty acid and glycerol units. The glycerol is converted to triose phosphate, much as glucose, and then enters the Krebs Cycle as pyruvate. The fatty acids undergo successive loss of C_2 units that subsequently enter the Krebs Cycle as acetyl coenzyme A.

When oxygen is absent in cells, hydrogen carried by NADH cannot be oxidized to H_2O and the Krebs Cycle cannot operate. Some organisms that normally live under aerobic conditions can switch to an alternative metabolism, called *fermentation*, which does not require oxygen. Such organisms are called *facultative anaerobes*. When such cells are deprived of oxygen, NAD⁺ is regenerated from NADH by reduction of pyruvic acid. Lactic acid^{*} (CH₃CH(OH)COOH) is one of several possible products of this process. In alcoholic fermentation, carried out by yeast, CO₂ is removed from pyruvate leaving acetalaldehyde (CH₃CHO), which is then reduced to ethanol (CH₃CH₂OH). Many bacteria can live only by fermentation and are poisoned by free oxygen. Such organisms are called *obligate anaerobes*. These bacteria, of which there is a great variety, carry out cellular respiration by using a compound other than oxygen as an electron acceptor. The most common such receptors are nitrate, nitrite, and sulfate, but Fe³⁺ and Mn⁴⁺ can be reduced by some bacteria.

ORGANIC MATTER IN NATURAL WATERS AND SOILS

Dissolved Organic Substances

Virtually all water at the surface of the Earth contains organic substances. These can be divided into *particulate organic matter* (POM) and *dissolved organic matter* (DOM). In a sense, though, there is a continuum, with colloids being intermediate between dissolved and particulate forms. By definition, dissolved substances are those that pass through a 0.45 μ filter; particulate organic matter is anything coarser than this. Together, dissolved and particulate organic carbon are called *total organic matter* (TOM). The carbon fraction of this organic matter is called *total organic carbon* (TOC) and can likewise be divided into *particulate organic carbon* (POC), which is also called *suspended organic carbon* (SOC), and *dissolved organic carbon* (DOC). In some instances, several other distinctions are made. Those organic substances that are volatile at ambient temperature constitute volatile organic carbon (VOC). Colloids have sizes in range of 1 nanometer to 0.45 μ , corresponding to molecular weights greater than about 10,000. They may be specifically referred to as colloidal organic matter (COM). Figure 14.19 illustrates the range in concentrations of dissolved and particulate organic carbon concentrations, while wet-

^{*}Animal muscle cells have the ability to switch to fermentation when oxygen cannot be supplied fast enough for NAD⁺ to be regenerated. Under these circumstances, lactic acid, the formal name for which is 2-hydroxy-propanoic acid, builds up in muscle tissue. The presence of lactic acid is what produces the sore muscle feeling that often accompanies vigorous excercise. "Conditioning" involves a number of processes, including increasing efficiency of oxygen delivery and a build-up of glycogen stores in muscle tissue, but also increasing tolerance to lactic acid. Nerve and brain cells, however, cannot respire anaerobically and quickly die when deprived of oxygen.

Geochemistry

W. M. White

CHAPTER 14: ORGANIC GEOCHEMISTRY

land waters (bogs, swamps, marshes) have the highest. The interstitial waters of the upper layers of soils often have DOC concentrations in the range of 20-30 mg/l. Organic carbon concentrations in natural waters generally show strong seasonal variations, as factors such as rainfall, biological productivity and mircobial activity all may vary seasonally. One might also note that even rain and snow typically contain some DOC; though the concentration is generally less 1 mg/l or less, much higher concentrations have been measured.

Organic substances dissolved in natural waters can be divided into simple compounds and humic substances. Simple compounds include *biomolecules*, which we will define as molecules synthesized by organisms to directly support life, and breakdown products of biomolecules, as well as excreted waste products and metabolites and substances released, or exuded, by cells by organisms to assist in breaking down and assimilating nutrients. Simple compounds include carboxylic acids (including fatty acids), phenols, carbohydrates, amino acids, nucleic acids, and hydrocarbons. As we shall see, humic substances are complex and poorly understood substances found in water and soil. Their exact origin is not well understood and is still somewhat controversial, but they probably include both residues of complex biomolecules and polymers that form from condensation of simpler biomolecules. Humic substances are further



Figure 14.19. Average concentration of dissolved, particulate, and total organic carbon in various natural waters. From Thurman (1985).



Figure 14.20. Components of dissolved organic carbon in typical river water. After Thurman (1985).

divided into fulvic and humic acids based on their solubility. *Hydrophilic acids* are a third set of compounds closely related to humic substances. Hydrophilic acids are, however, more soluble (as their name implies) and generally simpler than humic and fulvic acids. Humic, fulvic, and hydrophilic acids are the most abundant organic substances in most natural waters, as Figure 14.20 illustrates.

Simple Compounds

Simple compounds typically constitute 10 to 60% of dissolved organic carbon in natural waters, with an average of about 25%. Simple compounds are derived directly from cells or the breakdown of dead cells. Most of them are readily metabolized by bacteria, hence their residence time in water is usually quite short, probably a matter of minutes to days. Concentrations are maintained by continuous production by the biota. We might expect under these circumstances that the dissolved concentrations of these substances would be quite variable, and this is indeed the case. A number of instances of

CHAPTER 14: ORGANIC GEOCHEMISTRY

regular diurnal or seasonal variations in concentrations of carbohydrates and other simple compounds have been documented (reviewed in Thurman, 1985).

Fatty acids are among the most abundance of the dissolved simple organic compounds. They are also among the most important for several reasons. First, the carboxyl group acts as an acid, and hence these substances can significantly affect pH. Second, carboxylate can complex metal ions and will thus affect the concentration and activity of metal cations in solution. Third, they apparently form the basic building blocks of marine and aquatic humic substances.

The short chained carboxylic acids (C_1 to C_5), which include formic acid, acetic acid, propionic acid (CH_3CH_2COOH), butyric acid ($CH_3(CH_2)_2COOH$), and valeric acid ($CH_3(CH_2)_3COOH$) have relatively high vapor pressures and hence are known as volatile fatty acids. Acetic acid is the most abundant of these. Carboxylic acids are produced by microbial fermentation of more complex compounds and can accumulate in anaerobic environments. They are typically present in concentrations of 10 to 100 $\mu g/l$ in natural waters, but can reach higher concentrations in petroleum-related ground waters.

The longer chained, or non-volatile, fatty acids are produced by hydrolysis of triglycerides. Though these may range from C_{12} to C_{36} , C_{16} and C_{18} fatty acids (e.g., palmitoleic acid, its saturated equivalent palmitic acid, stearic acid and its saturated equivalent linoleic acid) are most common because they are most common constituents of lipids in marine and aquatic organisms. The concentrations of non-volatile fatty acids is typically 5 to 500 g/l and is generally somewhat higher than that of volatile fatty acids (Thurman, 1985).

Hydroxy and keto acids, particularly lactic, glycolic (also called hydroxy-acetic acid, OHCH₂COOH), and pyruvic acid, also occur in solution in natural waters in concentrations in the range of 10 to 250 μ g/l. They can originate from breakdown of plant matter, but most are probably directly excreted by algae and other organisms; hydroxy acids are produced in both the Krebs and Calvin cycles, as we have seen. Other related compounds that have been identified in solution in natural waters include dicarboxylic acids such as oxalic acid and aromatic carboxylic acids such as salicylic and benzoic acid. These generally do not exceed concentrations of a few μ g/l.

Phenols and other monomeric or dimeric aromatic compounds are generally present in concentrations of a few μ g/l or less in uncontaminated waters. Natural phenols result primarily from breakdown of lignins and tannins. Anthropogenic activity (i.e., pollution), however, can result in higher concentrations of phenols. An interesting example of "natural pollution" was the occurrence of high concentrations of phenol and cresol (a phenol unit with a methyl group attached to the ring) in South Fork Castle Lake in Washington, USA. This lake was created when debris flows from the 1980 catastrophic eruption of Mt. St. Helens blocked the South Fork of Castle Creek. Concentrations of phenol approached 1000 μ g/l and that of cresol approached 100 μ g/l (McKnight et al., 1982). These compounds were presumably produced by decomposition of the many dead trees floating in the lake.

Carbohydrates constitute about 10% of dissolved organic matter. Carbohydrates are found both as monosaccharides and polysacharides, as well as a variety of other forms, such as amino sugars, and sugar alcohols. In addition, carbohydrates are present as sugars bound to humic substances. Table 14.1

TAble	14.1.	Total	Dissolved	Carbohydrates	in
NATUR	al Wate	RS			

	Concentration $\mu g/l$	% DOC	
Groundwater	65-125	1-4	
Seawater	100-1000	5-10	
River Water	100-2000	5-10	
Lake Water	100-3000	8-12	

From Thurman (1985).

summarizes typical carbohydrate concentrations in various natural waters. In fresh waters, most carbohydrates are probably derived from leaching of terrestrial plant matter. In seawater, phytoplankton are thought to be the principle source of dissolved carbohydrates. Figure 14.21 summarizes the concentrations of carbohydrates in various natural waters. As this Figure shows, polysaccharides are the most abundant carbohydrates, fol-

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

lowed by saccharides associated with humic substances. Monosaccharides, sugar acids, amino sugars, methylated sugars and sugar alcohols are less abundant. Of the monosaccharides, glucose is generally the most abundant. Other important monosaccharides include fructose, arabinose, xylose, and galactose.

Carbohydrates are readily assimilated and metabolized by bacteria, hence their concentration in water is probably regulated by bacterial activity. A number of studies have demonstrated diurnal variations in carbohydrate concentration, with minimums occurring in the early morning and the maximum concentration occurring in the late afternoon. These results suggest active production of dissolved



Figure 14.21. Abundances of dissolved carbohydrates in seawater as a percentage of total dissolved organic carbon. After Thurman (1985).

carbohydrate by phytoplankton and rapid consumption by bacteria. Thus the residence time of a carbohydrate molecule in solution may be as short as a few hours in many instances (Thurman, 1985).

Amino acids are present in natural waters both as free molecules and combined in peptides and proteinaceous substances and are also important constituents of humic substances. The concentrations of free amino acids are typically a factor of four or so lower than the concentrations of combined amino acids. The simpler amino acids, glycine, serine, alanine, lysine, and valine, are the ones most abundant as free amino acids. Typical concentrations of free and combined amino acids are 15-200 μ g/l in seawater, 15-250 in groundwater, 40-800 in rivers and lakes, 250-5000 in eutrophic lakes and marshes, and 400-9000 in interstitial waters of sediments (Thurman, 1985). Since amino acids are essential to all life, the sources of amino acids are quite varied.

A variety of hydrocarbons are also present in natural waters. These may be divided into volatile and non-volatile hydrocarbons, depending on the vapor pressure and boiling point. Short-chained hydrocarbons tend to be volatile and this volatility limits their abundance in natural waters. Longer chained hydrocarbons are not volatile. Their abundance in natural waters is often limited by their solubility, which is generally low. Hydrocarbons may originate both from natural and anthropogenic sources. Here we focus on the naturally occurring hydrocarbons; anthropogenic hydrocarbons, i.e., pollutants, will be dealt with in a subsequent chapter.

Of the volatile hydrocarbons, methane is by far the most abundant. Some methane can be produced in the digestive tracks of higher animals, but the principal source of natural methane is methanogenic bacteria in reducing sediments. The amount of methane produced in such environments (swamps, reducing marine and lake sediments, etc.) is a significant, though minor, part of the global carbon cycle. The amount of methane released to the atmosphere from natural sources annually, primarily reducing sediments in wetlands, is 1.5×10^{14} g; anthropogenic activities release about 3.6×10^{14} g (Graedel and Crutzen, 1993). The abundance of methane in natural waters is summarized in Table 14.2. The solubility of methane in water is $800 \ \mu g/l$ (at 20°C and 0.1 MPa). At concentrations above this level, methane bubbles will form and this process is undoubtedly important in the transport of methane from sediment interstitial waters of swamps and shallow lakes to the atmosphere. Some of this methane redissolves in the lake water and is oxidized by methanotrophic bacteria.

Among the most common non-volatile and semi-volatile hydrocarbons in seawater are C_{15} and C_{17} n-alkanes (pentadecane and heptadecane) and isoprenoids. C_{15} and C_{17} originate by decarboxylation of C_{16} and C_{18} fatty acids (palmitic and stearic acids), which are derived from zoo- and phytoplank-

Geochemistry

CHAPIER 14: ORGANIC GEOCHEMISTRY

ton. The isoprenoids, primarily pristane and phytane, are derived from the phytol chain in the chlorophyll molecule. *Halogenated hydrocarbons*, i.e., hydrocarbons where one or more hydrogens are replaced by a halogen, are of particular interest because of their toxicity. Most halogenated hydrocarbons in natural waters are anthropogenic, having been directly manufactured and discarded or leaked Others, such as chloroform and into natural waters. trichloromethane, can arise indirectly through chlorination of drinking water. However, some halogenated methanes do occur naturally in seawater at very low abun- From Thurman (1985). dance levels.

TADLE 14.2.	Dissolved Methane	iN
NATURAL WA	TERS	

	Concentration µgC/1
Groundwater	10-10,000
Seawater	10-100
Lake Water	10-10,000
Interstitial Water	100-10,000

HUMIC SUBSTANCES

Humic substances are high molecular weight (> 500 daltons) compounds that are produced by partial degradation of complex biomolecules and recombination of these with simple biomolecules and their breakdown products. They are complex molecules: their exact structures are not known and in any case variable. As we noted earlier, humic substances dominate dissolved organic matter in natural waters. They are also important contributors to particulate organic matter.

The exact definition of humic substances is analytical and varies somewhat between dissolved humic substances and soil humic substances. Aquatic and marine humic substances are defined as colored (yellow, brown, or black), polyelectrolytic acids that can be removed from solution through absorption on weak-base ion exchange resins (e.g., the acrylic-ester resin XAD-8) or through some similar procedure. Humic substances tend to be highly refractory in the biological sense, that is, they resist decomposition by organisms. As a result, their residence time in water is on the scale of weeks to thousands of years. Although their molecular structures remain poorly known, the chemical composition and functional groups of humic substances are well characterized. The most important of these are carboxyl, phenolic, alcoholic, carbonyl, amino, and sulfhydryl (SH) groups. Because of the predominance of the carboxyl groups, humic substances are acidic (i.e., they act as proton donors). Hy*drophilic* acids are closely related to humic substances: they are molecules too complex to fully describe, but are not absorbed as readily onto ion exchange resins. They are slightly colored, highly branched, and highly substituted organic acids. They appear to have lower molecular weight and a greater number of acid functional groups than humic substances.

Humic substances in waters are divided into *fulvic acid* and *humic acid*. The definition of these two is again analytical. Humic acids are defined as those humic substances that precipitate when the solution is acidified with HCl to a pH of 1. Fulvic acids are those substances remaining in solution at this pH.

The concentration ranges of humic substances are given in Table 14.3. As a proportion of dissolved organic carbon, hydrophilic acids predominate in seawater

and groundwater, constituting about 50% of DOC. They are only slightly less important in lake waters. In wetlands, streams and rivers, fulvic acids are the most important fraction of dissolved organic carbon, constituting 60% of wetland DOC and 40% of stream and river DOC. Humic acids are always less abundant than fulvic and hydrophilic acids. Humic acids constitute 15% of wetland DOC and less than 10% of DOC in other waters (Thurman, 1985).

While the definition of humic and fulvic acids is based on their acid solubility, there is nevertheless a compositional difference between the two. Table 14.4 compares the compositions of humic and fulvic acids from several envi- From Thurman (1985).

TADLE 14.3.	Humic	SUDSTANCES	in
NATURAL WA	TERS		

	Concentration mg C/l
Groundwater	0.03-0.10
Seawater	0.06-0.60
Lake Water	0.5-4.0
Rivers	0.5-4.0
Wetlands	10-30

CHAPTER 14: ORGANIC GEOCHEMISTRY

ronments. Humic acid is usually richer in carbon and poorer in oxygen than fulvic acid; it is also on average poorer in hydrogen and richer in nitrogen. These compositional differences are better developed in soil humic substances than in dissolved ones.

Fulvic acid tends to have a lower average molecular weight (800-2000 daltons) than humic acid (>2000 daltons). In addition, fulvic acid has a higher content of carboxyl acid groups than does hu-

SOLVED IN INATURAL WATERS								
	С	Н	0	Ν	Р	S	Ash	
Groundwater								
Biscayne Aquifer Fulvic	55.4	4.2	35.4	1.8			0.04	
Biscayne Aquifer Humic	58.3	3.4	30.1	5.8			10.4	
Seawater								
Saragasso Sea Fulvic	50.0	6.8	36.4	6.4		0.46		
Lake Water								
Lake Celyn Humic	50.2	3.1	44.8	1.9				
Lake Celyn Fulvic	43.5	2.7	51.6	2.2				
Stream Water								
Ogeechee Stream Fulvic	54.6	4.97	38.2	0.87	0.62	0.74	0.86	
Ogeechee Stream Humic	55.9	4.19	36.5	1.27	0.25	0.93	1.13	

Table 14.4. Composition of Fulvic and Humic Acids Dissolved in Natural Waters

Data from Aiken et al. (1985). Concentrations in weight percent.

mic acid, while humic acid is richer in phenolic groups (Table 14.5). The higher content of carboxyl groups in fulvic acid and the greater abundance of aromatic structures in humic acid helps to explain the higher solubility of fulvic acid. In general, dissolved humic substances have higher concentrations of carboxyl groups than soil humic substances. The average concentration of carboxyl groups in dissolved fulvic acid, 5.5mM/g, corresponds roughly to one carboxyl group per 6 carbon atoms. In contrast, dissolved humic acid has about one carboxyl group per 12 carbon atoms.

In addition to functional groups, humic substances incorporate a variety of biomolecules in their structures. Carbohydrates account for 1% of the carbon in dissolved humic substances, and their abundances appear to be higher in humic than in fulvic acids. Many of these carbohydrates appear to be bound to the humic structure only through hydrogen bonds. Amino acids are also present in humic substances, although only in trace amounts. The most abundant are glycine, aspartic acid, glutamic acid, and alanine. Aquatic humic acids contain about 120 nM/mg amino acid; aquatic fulvic acids generally contain less: 15-120 nM/mg. Soil humic substances contain higher amino acid concentrations: roughly 500 nM/mg for soil humic acids and 150 nM/mg for soil fulvic acids. Hydrogen bonding appears to play a role in molecular aggregation.

Soil humic substances have core structures that are highly aromatic (20 to 70% of total C). It was originally thought that this was also true of aquatic humic substances as well. However, a variety of studies over the last 20 years have led to the view that the core structures of aquatic and marine humic substances are primarily aliphatic rather than aromatic (Malcom, 1985; Thurman, 1985, Stein-

berg and Muenster, 1985; Harvey and Boran, 1985). Nuclear magnetic resonance studies suggest only 15-20% of the carbon in aquatic fulvic acids and about 30% of the carbon in aquatic humic acids are aromatic. Figure 14.22 illustrates one proposed hypothetical structure of aquatic fulvic acid. This molecule has a molecular weight of about 1000 daltons. Marine humic acids appear to have even a smaller proportion of aromatic carbon, and marine fulvic acids have essentially none. Since aromatic structures are far more common in terrestrial plants than

TADLE	14.5	Functional	GROUPS	in	Dissolved
Fulvic	and H	lumic Acids	-		

	Carboxyl	Phenolic
	meq/l	
Groundwater Fulvic	5.1-5.5	1.6-2.0
Groundwater Humic		2.5
Seawater Fulvic	5.5	-
Lake Fulvic	5.5-6.2	0.5-2.1
Lake Humic	5.9	3
Stream & River Fulvic	5.5-6.4	1.5-2.1
Stream & River Humic	4.0-4.7	1.9-2.0

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY



Figure 14.22. Hypothetical structure of aquatic fulvic acid. After Bergmann (1978).

marine ones, this is not surprising.

The shape of humic substances is strongly influenced by pH. At low pH, humic and fulvic molecules are fibrous. At neutral pH, these fibers tend to mesh together to form a sponge-like structure. This mesh structure is capable of trapping smaller molecules. For example, fatty acids and acyclic alkalenes, which are found in most humic and fulvic acids, may be simply trapped rather than bonded to the molecular backbone. Under alkaline conditions the structure becomes plate-like.

The origin of dissolved humic substances is not well understood. At one time it was widely held that aquatic humic substances consist of soil fulvic acid that had been leached or eroded from soils. However, most recent works on the subject express the view that at least a substantial fraction of aquatic humic substances is autochthonous, that is, produced within the body of water itself. The evidence favoring this view includes the compositional and structural differences between aquatic and soil humic substances. Aquatic humic substances are primarily aliphatic whereas soil humic substances are primarily aromatic. This observation in itself does not exclude derivation of aquatic humic substances from soil (since leaching may preferentially remove aliphatic-rich molecules), but the difference is consistent with soil humic substances being derived primarily from higher plants (which are rich in aromatic molecules such as lignin and tannins), and aquatic humic substances being derived primarily from algae and zooplankton (which are poor in aromatic molecules). An additional question is the degree to which humic substances are composed of partially degraded macrobiomolecules or are condensed from simpler fragments. The latter view is supported by several observations. One is that clay particles, metal oxides, and metal cations all can promote polymerization of organic molecules. Bacteria, or enzymes released by them, have also been shown to promote condensation and polymerization of organic molecules. Hydrophilic acids may be precursors of the more complex fulvic and humic acids. As we noted, soil humic substances are highly aromatic, and this suggests that lignins and tannins derived from higher plants are important contributors to soil humic substances. Lignins and tannins are relatively resistant to decomposition, although they are probably modified to some degree by decomposers before incorporation in humic molecules. Molecular by-prod-
Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY



Figure 14.23. Possible mechanism for generation of marine fulvic and humic acids from triglycerides. After Harvey and Boran (1985).

ucts of decomposers and their remains as well as waxes from higher plants are probably the primary contributors to the aliphatic components of soil humic substances (Killops and Killops, 1993).

Marine fulvic acids may arise by autoxidative cross-linking of polyunsaturated lipids, perhaps catalyzed by light and transition metals (Harvey and Moran,1985). Olefinic carbons (i.e., those doubly bonded to other carbons) may be particularly susceptible to autoxidatation. Figure 14.23 illustrates this process. A number of laboratory experiments have demonstrated the plausibility of production of marine humic substances in this manner. In this model of humic substance generation, aromatic units in marine humic acid arise from cross-linking of fatty acid chains. Another possibility is that aromatic units in marine humic acids are derived from terrestrial sources: degradation products of lignin and tannin washed into the sea. Most likely, however, terrestrial material is only a locally important contributor to marine humic substances.

ORGANIC MATTER IN SOILS

Organic matter can constitute a third or more of the mass of poorly drained soils, though fractions from 6% to 10% are more common in well-drained soils. Concentrations of organic matter are highest in the surface layers (O and A) and low in the deeper layers (C), as we saw in Chapter 13. In the soil solution, a range of 2 to 30 mg/l DOC is common (Thurman, 1985). Soil organic matter, collectively called *humus*, includes biomolecules as well as humic substances. Biomolecules can be released from

CHAPTER 14: ORGANIC GEOCHEMISTRY

dead cells, excreted from living ones, or leached from foliage and rinsed into the soil by rain (i.e., throughfall). In addition, soil organic matter includes a variety of substances, called *exudates*, released by macro- and microorganisms expressly to breakdown both organic and inorganic components in soils so that they may be as-Among the most similated. important of the exudates are simple carboxylic acids such as acetic and oxalic acid (Fig. 14.6), formic acid (HCOOH), tartaric acid (2,3,-dihydroxybutanedioic acid: COOH(CH (OH))₂COOH) and citric acid, After Schnitzler (1978).

	Humic Acid		Fu	lvic Acid	
	Mean	Range	Mean	Range	
Elemental Comp). ——		-wt %		
C	56.	53-59	45.7	40.7-50.6	
Н	4.6	3.0-6.5	5.4	3.8-7.0	
Ν	3.2	0.8-5.5	2.1	0.9-3.3	
0	35.5	3238.5	44.8	3950.	
S	0.8	0.1-1.5	1.9	0.1-3.6	
Functional Groupsmeq/g					
Total Acidic Gro	ups	5.6-8.9		6.4-14.2	
Carboxyl	3.6	1.5-6.0	8.2	5.2-11.2	
Phenolic OH	3.1	2.1-5.7	3.0	0.3-5.7	
Alcoholic OH		0.2-4.9		2.6-9.5	
Quinoid/keto C=O		0.1-5.6		0.3-3.1	
Methyloxy OCH ₃		2.1-5.7	3.0	0.3-5.7	

Table 14.6. Composition of Soil Humic and Fulvic Acids

and a variety of phenolic acids. The OH groups of phenols may also dissociate, and hence these compounds also contribute to soil acidity and rock weathering, though less so than the carboxylic acids. These simple organic acids are commonly present in relatively high concentrations around plant roots in soils, though on average their concentration is less than 1mM in the soil solution (Drever and Vance, 1994). Because of the presence of both these simple acids and the more complex fulvic and humic acids, most soils are slightly acidic. These acids contribute both directly (through surface complexation reactions) and indirectly (as proton donors, as by increasing the solubility of cations through complex formation) to weathering of rocks, though, as we found in Chapter 13, the overall extent to which weathering is accelerated by plants is unclear. As most biomolecules are readily metabolized by bacteria, their residence time in the soil is likely to be quite short, a matter of days or less. Concentrations of these molecules are maintained by continuous production by the biota. In contrast, the residence time of refractory humic substances can in soils can be as long as thousands of years.

The definition of soil humic substances differs slightly from that of aquatic and marine humic substances. Soil humic substances are divided into fulvic acid, humic acid, and humin based on their solubility. By definition, the organic material that remains insoluble when $500 \text{ mol/m}^3 \text{ NaOH}$ is added to soil is called humin. The material dissolved by this procedure may be separated into fulvic and humic acids by adjusting the pH to 1 by addition of HCl: the material precipitated at that point is humic acid, that remaining in solution is fulvic acid (the procedure for separating fulvic and humic acid is the same for aquatic and marine humic substances).

As was the case for aquatic humic substances, there is a systematic compositional difference between soil humic and fulvic acids; indeed the compositional differences between fulvic and humic acid appears greater in soils than in water. As is summarized in Table 14.6, humic acid is richer in carbon and poorer in oxygen and sulfur than fulvic acid, and poorer in hydrogen and richer in nitrogen as well. The higher content of polar groups such as carboxyl accounts for the higher solubility of fulvic acid. Soil fulvic acids tend to have molecular weights $<2\times10^3$, while the molecular weight of soil humic acids can exceed 10⁶.

Based on these compositions, the approximate chemical formula for average humic acid is $C_{187}H_{189}O_{89}N_9S$ and that of fulvic acid is $C_{68}H_{91}O_{48}N_3S$. Compared to the composition of living organisms, humic and fulvic acid have substantially lower ratios of hydrogen, oxygen, and nitrogen to carbon. Since humic substances are ultimately derived from biomolecules, this indicates that H, O, and N are lost in the humification process. Humin, the insoluble organic matter in soil, tends to be even

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY



Figure 14.24. Possible structure of soil humic acid. After Stevenson (1982).

richer in carbon and nitrogen and poorer in sulfur than either fulvic or humic acid. From these formulas we can deduce that there is a substantial degree of unsaturation (i.e., double carbon bonding) in humic substances, which partially accounts for their stability and biologically refractory nature. Stevenson and Vance (1989) estimated the average content of functional groups in soil humic and fulvic acids as 7.2 and 10.3 meq/g respectively. One possible structure for soil humic acid is illustrated in Figure 14.24.

The origin of soil humus is not precisely known. The relatively high proportion of aromatic units suggests the most important contributors to humic substances are lignins and tannins. As we noted earlier, these are polyaromatic substances that are quite refractory in a biological sense. These are partially degraded by soil microbes. Monomer or smaller polymer units may then condense, perhaps catalyzed by clays, metal ions, or bacteria. To this basic structure other components, which include amino acids, carbohydrates, and alkanes derived from fatty acids may be attached. Soil microbes are probably the primary contributors of these units.

As we found in Chapter 13, organic compounds, particularly carboxylic acids such as oxalic acid, in soils play an important role in podzolization; i.e., the depletion of Fe and Al in the upper soil horizons and their enrichment in lower horizons. This occurs as a result of the ability of carboxylic acids to form soluble complexes with Fe and Al. Fe and Al carboxylate complexes form in the upper soil layers, where organic acid concentrations are high, then are carried to deeper levels by water flow. At deeper levels, the carboxylate is oxidized by bacteria, and the Fe and Al precipitate as hydroxides. We examine metal-organic complexation further in the following section.

CHEMICAL PROPERTIES OF ORGANIC MOLECULES

Acid-Base Properties

As we noted above, the carboxyl group can dissociate to give up a hydrogen atom:

$$RCOOH \rightleftharpoons H^+ + RCOO^-$$
 14.2

(we use R here as a general representation for the remainder of the molecule). Like other acids, organic acids will increasingly dissociate as pH increases. As for other reactions, we can write an equilibrium constant expression:

$$K_a = \frac{a_{H^+} a_{RCOO^-}}{a_{RCOOH}}$$
 14.3

The equilibrium constant is commonly reported as pK_a , which, analogous to pH, is the negative of the logarithm of the equilibrium constant. The Henderson-Hasselback equation relates pK_a , activity quotient, and pH:

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

$$pK_a = -\log K_a = -\frac{a_{RCOO^-}}{a_{RCOOH}} - pH$$
 14.4

Expressed in this way, the pK_a is the pH where half of the acid molecules are dissociated and half undissociated. Most carboxylic acids have pK_a values between 1 to 5, which is below the pH of most streams and lakes, and similar to or lower than the pH range of soils (generally 3 to 8). Thus carboxylic acids will be dissociated in most environments. As we have seen, carboxylic acids (both simple and as functional groups on humic substances) are important components of dissolved and soil organic matter. The effect of increasing concentrations of organic molecules will thus be to increase the concentration of protons, lowering the pH of natural waters.

The OH groups of phenols can also dissociate and hence contribute to solution or soil acidity. Phenols are, however, much weaker acids than carboxylic acids. One measure of this is the pK_a . Whereas the pK_a values of carboxylic acids are typically 3 or so, pK_a for phenols are more typically 8. Thus at pH typical of most natural waters and soils, phenols will be only partially dissociated or not dissociated.

In contrast to carboxyl and phenol groups, nitrogen-containing groups, such as amino groups, are basic as they tend to bind free protons. By analogy to pK_a , defined above, we can define a pK_b , which is the pH will half the molecules or groups will be protonated and half unprotonated. Aliphatic amines have the strongest basicity, with typical pK_b values of 10 to 12. At pH values below this, they will be protonated. Thus in most natural waters and soils they will bear a positive charge and behave as cations. Because particle surfaces are typically negatively charged, these organic cations are readily bound to particle surfaces and removed from solution. Aromatic amines are typically weaker bases, with pK_b values around 4 to 6. They will be protonated only in acidic waters and soils.

As we have seen, humic and fulvic acids are generally the most abundant organic substances in natural waters and soils. They often contribute significantly to the acidity of waters and soils. Un-

der some circumstances, such as lowland tropical rivers, swamps, etc., they are the principle negative ions present. They typically contain 10^{-2} eq/g ionizable acid groups per weight of organic carbon. Carboxyl groups are most common, but other functional groups are also present. As a result, humic substances cannot be characterized by a single pK_a. Their titrations curves (Figure 14.25) typically have а "smeared out" appearance, a result both of the variety of functional groups present, and electrostatic interactions between these groups.

Complexation

Another important geochemical property of organic molecules is their ability to form complexes with metals,



Figure 14.25. Comparison of titration curves of salicylic and humic acid. The salicylic acid shows two end-points, corresponding to its carboxyl and phenol functional groups. The humic acid shows a smeared out titration curves corresponding to a continuously changed pK_a . From Morel (1983).

M. White

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

especially transition metals and aluminum. Complexation between metal ions and organic anions is similar, for the most part, to complexation between metals and inorganic anions. One important difference is that many organic compounds have more than one site that can bind to the metal. Compounds having this property are referred to as *multidentate*. Complex formation with multidentate ligands is called *chelation*, and the complexes formed are called *chelates*. A simple example is the oxalate ion, (COO)²⁺, which consists of 2 carboxyl groups (Figure 14.7) and is *bidentate*. Citric acid and glutamic acid (an amino acid) are tridentate ligands (although the amino group will not dissociate and take on a positive charge, it nevertheless has an electron pair available to share). A second important difference is that organic functional groups are only weakly acidic, meaning they have a high affinity for protons and are often only partially dissociated at the pH of natural waters. As a result, metal-organic complexation is strongly pH dependent.

A metal-oxalate complex results in the formation of a ring (Figure 14.26), with the two oxygens that are singly bonded to carbon each binding to the metal. In the oxalate complex, the ring has five members; a 6 member ring would be formed in a metal-malonate; 7 member rings are formed by phtalate or succinate chelates. A metal glycine complex forms 2 rings on opposing sides of the metal (Figure 14.26). Salicylate is another example of a bidentate ion (Figure 14.27a). In this case, the binding sites are formed by two different functional groups: carboxyl and hydroxyl. Not all organic anions are multidentate, of course. Benzoate, acetate, and phenol are examples of unidentate organic anions.

Multidentate ions can very strongly bind trace metals. A few natural chelators that are specific for Fe have been characterized. One, endobactin, is illustrated in Figure 14.27b. These Fe specific chelators may have stability constants in excess of 10³⁰, so they are indeed very strong chelators. However, they are at sufficiently low concentration that they usually do not play a significant role in trace metal chemistry. However, they may do so in special circumstances such as plankton blooms. It is also possible for a single metal ion to bind to more than one organic ligand, as illustrated in Figure 14.27c.

Just as for inorganic metal-ligand complexes, we can define stability constants (β) and apparent stability constants (β^*) for metal-organic ligand complexes. Table 14.7 lists some examples of stability made. First, as we saw for inorganic lig- separate benzoate and acetate groups.





Cu glycine

Figure 14.26. Examples of rings formed by chelates.



constants for metal-organic complexes Figure 14.27. a. Copper salicylate complex. The Cu ion is that we might expect to find in natural bound to both the carboxylic and phenol groups. b. Enterowaters. Several generalizations may be bactin, a natural iron chelating agent. c. Ni complexed by

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

ands, the higher valence state of the metal, the stronger the complex. Thus in Table 14.7, Fe^{3+} forms stronger complexes with all listed ligands than does Fe^{2+} . Second, there is high degree of correlation between the equilibrium constants of all organic ligands for a given metal. For the divalent metals, stability of metal-organic complexes follows the Irving-Williams series (Chapter 6), $Pb^{2+} > Cu^{2+} > Ni^{2+} > Co^{2+} > Zn^{2+} > Cd^{2+} > Fe^{2+} > Mn^{2+} > Mg^{2+}$.

It is the functional groups of organic molecules that are primarily responsible for metal ion complexation. Amines, azo compounds (compounds containing a -N=N- group linking two other groups), ring nitrogen, carboxyl, ether, and ketone are all important in complex formation. Tetrapyrrole pigments, or porphyrins, such as chlorophyll are very strong metal ion complexing agents, particularly for transition metals such as Zn and Ni. In the case of chlorophyll and similar molecules, complex formation occurs through replacement of Mg with a transition metal ion. In this instance, the metal is bound to two nitrogens (see Fig. 14.15).

Multidentate complexes are generally more stable than corresponding unidentate ones. We can see this in Table 14.7, where the stability constants for citrate, with 3 carboxyl binding sites are higher than those for acetate, with one carboxyl binding site. Another interesting property of multidentate ligands is that the degree of complexation decreases less strongly with dilution than from monodentate complexes.

As was the case for inorganic complexes, the stability of metal-organic complexes are invariably strongly pH dependent. The reason for this is simple: hydrogen ions will be major competitors for sites on functional groups. Indeed, we can write the complexation reaction as:

$$m \mathrm{M} + l \mathrm{HL} \rightleftharpoons \mathrm{M}_m \mathrm{L}_l + \mathrm{H}^+$$

the equilibrium constant for this reaction is then:

$$K = \frac{[M_m L_l] [a_{H^+}]}{[M]^m [HL]^l}$$
 14.5

or in log form: $\log K = \log [M_m L_l] - m \log [M] - l \log[L] - pH$ 14.6

Table 14.7. Log Equilibrium Constants for Metal Ion-Organic Ligand Complexation

			L	ωgβ		
	Glycine ¹	Glutamate ²	Acetate ³	Citrate ⁴	Malonate ⁵	Salicylate ⁶
H^+	9.78	9.95	4.76	6.4	5.7	13.74
Na^+				1.4	0.7	
K^+				1.3		
Mg^{2+}	2.7	2.8	1.3	4.7	2.9	
Ca ²⁺	1.4	2.1	1.2	4.7	2.4	0.4
Al^{3+}			2.4			14.2
Ba^{2+}	0.8	2.2	1.1	4.1	2.1	0.2
Fe^{3+}	10.8	13.8	4.0	13.5	9.3	17.6
Fe^{2+}	4.3	4.6	1.4	5.7		7.4
Ni ²⁺	6.2	6.5	1.4	6.7	4.1	7.8
Cu^{2+}	8.6	8.8	2.2	7.2	5.7	11.5
Zn^{2+}	5.4	5.8	1.6	6.1	3.8	7.7
Pb ²⁺	5.5		2.7	5.4	4.0	
Hg^{2+}	10.9		6.1	12.2		
Ag ⁺	3.5		0.7			
¹ NH ₂ CH	I ₂ (COO)⁻			⁴ (HOOC)CH	H ₂ C(OH)(COO	H)CH ₂ COO ⁻
² (HOOC)(CH ₂) ₂ CHN	H ₂ COO ⁻		⁵ CH ₃ CH ₂ CO	-00	
${}^{3}C_{2}H_{5}CC$	00-			⁶ HOC ₆ H ₄ C	-00	
From Mo	rel and Herii	ng (1993).				

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

We should also note that for multidentate ions, mixed hydrogen-metal and hydroxide-ligand complexes are possible. These will become increasingly important at low and high pH respectively. For example, at pH values below about 3, the CuHCitrate complex will be dominate over the simple Cucitrate complex. In another example, the FeOHGlycolate complex will be more important at all pH values than the simple Fe-glycolate complex. For clarity, we have omitted stability constants for these mixed complexes from Table 14.7. Nevertheless, as these examples show, these mixed complexes must often be considered in speciation calculations, particularly at high and low pH. A more complete compilation of stability constants for metal-organic complexes may be found in Morel and Hering (1993).

Humic and fulvic acids are, as we have seen, usually the most abundant dissolved organic compounds. The functional groups on these substances are capable of complexing metals. A variety of experiments suggest the majority, 50 to 90%, of these functional groups are carboxyl groups, followed in importance by phenol and alcohol groups. When two or more of these functional groups are present in close proximity on the molecule, humic substances will act as chelators and complex metals in a manner similar to simpler multidentate organic molecules. Salicylic acid, with a carboxylic and phenolic group, can be used a simple analog for humic substances in this respect. Considering the stability constants for salicylate complexes listed Table 14.7, we would expect humic substances to strongly complex trace metals. This appears to be the case, although the situation is somewhat complicated be-

Example 14.1. Speciation of Organic Ligands in Fresh Water

Using the stability constants in Table 14.7, and the calculated free ion activities for major cations in Example 6.7, calculate the speciation of glycine, citrate and salicylate. Assume total activities of glycine, citrate, and salicylate of 1.25×10^{-8} , 5×10^{-8} , and 1×10^{-8} M/l respectively.

Answer: For each ligand, we can write a conservation equation:

$$\Sigma L = L^{-} + HL + AL + BL + CL + \dots$$
 14.7

where L^- is the free ligand, HL is the undissociated acid, and AL, BL, CL, etc. are the various metal ligand complexes. For each species we may also write:

$$[ML] = \beta \times [M] \times [L^{-}]$$
 14.8

where [ML], [M], and $[L^{-}]$ are the concentrations of the complex, free metal ion or proton, and free ligand respectively. Substitutiong 14.8 into 14.7, we have:

$$\Sigma L = [L^{-}] + \beta_{HL} [H] [L^{-}] + \beta_{AL} [A] [L^{-}] + \beta_{BL} [B] [L^{-}] + \dots$$
 14.9

Rearranging, we have:

$$[L^{-}] = \frac{\Sigma L}{1 + \beta_{AL}[A] + \beta_{BL}[B] + \dots} \quad 14.10$$

Since the concentrations of the organic ligands are much lower that those of the major cations, we can assume that organic complexation does not affect activities of the major cations. Equation 14.10 gives us the free ion concentration. From that, we can calculate the concentration of each of the complexes using 14.8. The result is shown in the adjacent table. We see that at the pH of this example (8), glycine and salicylate are essentially completely undissociated. Citrate is almost

Speciation of Organic Ligands			
	Glycine	Citrate	Salicylate
Н	98.24%	0.11%	100.00%
Na	0.00%	0.03%	0.00%
Κ	0.00%	0.01%	0.00%
Mg	0.11%	31.77%	0.00%
Ca	0.01%	63.55%	0.00%
free ligand	1.63%	4.53%	0.00%
activity of			
freeligand	2.04×10 ⁻¹⁰	2.26×10 ⁻⁰⁹	1.82×10^{-14}

completely dissociated, but is 95% complexed by Mg and Ca.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

cause the different binding sites of humic substances have differing metal affinities. Experiments suggest that overall stability constants for trace metal–humic substance complexes are in the range of 10^5 (for Zn²⁺, Cd²⁺, and Ni²⁺) to 10^{10} (for Cu²⁺) (e.g., Buffle et al., 1980 Zuehlke and Kester, 1983).

It is not possible, however, to assign specific values to stability constants (or acidity constants for that matter) for humic substances. This seems to be a reflection of several factors, including electrostatic interaction between neighboring functional groups and distortions of the molecule that result from cation binding and neighboring charges in solution. The latter leads to a strong dependence of apparent stability constants on ionic strength. There has been some success in modelling humic and fulvic acids using simple empirical models (e.g., Cabaniss and Shuman, 1988) as well as more theoretical ones that assume the binding properties can be described by combining the complexation the properties of a few simple organic compounds (e.g., acetate, malonate, and catechol) together with a consideration of the coulombic attraction of neighboring groups (Bartschat, et al., 1992; Morel and Hering, 1993).

The degree to which dissolved trace metals in natural waters are complexed by organic ligands has been much debated. There is a fair body of experimental evidence to support the idea that a large fraction of at least some trace metals (e.g., Cu, Zn, Pb) in natural waters is complexed by organic compounds. This evidence consists of bioassay and electrochemical measurements that show the "reactive" concentrations (i.e., free ion) are well below total concentrations (e.g., Hering et al., 1983; Donat and Bruland, 1990). While conceding that organic ligands do indeed have a strong preference for some metals, particularly transition metals, others have argued on more theoretical grounds that in most natural waters, organic ligands complex only a small fraction of these trace metals (e.g., Nürnberg and Valenta, 1983). This argument is based on the following observations. First, concentrations of dissolved organic ligands are generally quite low. Second, the organic ligands that are present will be almost entirely complexed by the major metal ions (e.g., Ca^{2+} , Mg^{2+}) and H⁺ (Example 14.1). While these complexes have lower stability constants, these ions are much more abundant. Third, trace metals will be largely complexed by the more abundant inorganic ligands. Again, while these complexes are less stable, this is more than compensated for by the much greater abundance of inorganic ligands. However, even proponents of this view concede that in certain circumstances, such as estuaries, organic-rich tropical rivers and areas of high productivity in the oceans, where water has high natural DOC concentrations or waters polluted by strong chelators such as EDTA, some trace metals will be almost completely complexed by organic ligands. In other areas, it is possible that the difference between these perspectives reflects the presence of very strong complexing agents known as siderophores, such as trihydroxamate siderophore desferriferrioxamine B, that are produced by organisms to (1) acquire essential trace metals, such as Fe and Zn, and (2) defend themselves against the toxicity of others, such as Cu and Pb. The distribution of these substances, which can have stability constants in excess of 10^{30} , remains poorly known.

Example 14.2 demonstrates that glycine, a common amino acid, and citrate, a common hydroxycarboxylic acid, and salicylate, a common phenolate, will complex only a small fraction of the total Cu in fresh water with typically low concentrations of these substances. We see that this is due to several factors. First, at this pH, most of the glycine and salicylate are undissociated (Example 14.1), and therefore unavailable to bind Cu, and 95% of the citrate is complexed with Ca and Mg. Second, the greater abundance of inorganic anions such as hydroxyl and carbonate results in their dominating the speciation of Cu. However, one should avoid the drawing the conclusion that organic trace metal complexes are inevitably insignificant. We considered only 3 species in this example, and while they strongly bind copper, all are at fairly low concentration. Other organic anions, particularly including humates, are often present a sufficient concentration to complex a significant fraction of some trace metals. Problem 14.5 and 14.6 at the end of this chapter illustrates that situation.

The complexing behavior of humic substances is, well, complex and cannot be characterized by a single stability constant. This is true for several reasons. First, different functional groups can be present on a single molecule, each of which will have a different intrinsic stability constant and pK_a . Second, these functional groups are close enough to one and other that the electrostatic charge one site

CHAPTER 14: ORGANIC GEOCHEMISTRY

can affect the complexing properties of an adjacent one (recall that this was also true of surfaces: see Chapter 6). In particular, protonation and deprotonation changes the charge on the humate molecule, which will affects the attraction of metal ions for it. As a result, metal-humate stability constants vary as function of pH. A full treatment of this problem is beyond the scope of this book, but may be found in Morel and Hering (1993).

Example 14.2. Speciation of Cu in Fresh Water

Use the adjacent stability constants as well as those for glycine, citrate, and salicylate in Table 14.7 to calculate the speciation of Cuin the water sample analysis in Example 6.7, assuming $\Sigma Cu = 10^{9}$ M. Use the calculated free ion concentrations of anions in Examples 6.7 and 14.1.

CU STADILITY CONSTANTS		
	logβ	
$CuOH^+$	6.3	
$Cu(OH)_2$	11.8	
CuCl ⁺	0.5	
CuCO ₃	6.7	
CuSO ₄	2.4	

Answer: In calculating trace element speciation, it is common to assume that complexation with trace metals does not reduce the free ion concentrations of the anions. For this assumption to be

valid, the free ion concentrations of the anions should greatly exceed that of the trace metal. This condition is met in this case for the inorganic anions, but not for the organic ones. Nevertheless, we will proceed by making this assumption initially and subsequently examine its validity and make the necessary corrections. We proceed much as we did in Example 14.1; that is we write a conservation equation for copper:

$$\Sigma Cu = [Cu^{2+}] + [CuOH^{+}] + [Cu(OH)_{2}] + [CuCl^{-}] + [CuCO_{3}] + [CuSO_{4}] + [CuGly] + [CuCit] + [CuSal] 14.11$$

For each species, we also write a mass action equation, for example:

$$CuCit = \beta_{CuCit} \times [Cu^{2+}] \times [Cit]$$
 14.12

Substituting the mass action equations into 14.11 and solving for $[Cu^{2+}]$, we have:

$$[Cu^{2+}] = \frac{\Sigma Cu}{1 + \sum_{i} \beta_{CuL_{i}}[L_{i}]}$$
 14.13

We can then calculate the concentrations of the individual species using 14.12. The results are shown in the adjacent table. We see that Cu is dominantly complexed by hydroxyl and carbonate. The three organic complexes account for only about 1% of the total copper.

Now let's examine our initial assumption that Cu speciation does not reduce the free ion activities of the anions. With the exception of copper salicylate, the concentration of each species is far less than the free ion concentration of the corresponding anion. In the case of salicylate, however, the concentration exceeds the total free ion concentration of salicylate, a clear indication that our initial assumption was invalid. We could address this problem by performing an iterative calculation such as that used in Example 6.7. However, an examination of the situation reveals a

Calculated Copper Speciation			
	Conc	%	
CuOH ⁺	2.12×10^{-10}	21.22%	
$Cu(OH)_2$	6.71×10^{-11}	6.71%	
CuCl ⁺	7.03×10^{-14}	0.01%	
CuCO ₃	5.97×10^{-10}	59.69%	
$CuSO_4$	4.41×10^{-12}	0.44%	
CuGly	8.63×10^{-12}	0.86%	
CuCit	3.82×10^{-12}	0.38%	
CuSal	6.12×10^{-13}	0.06%	
Cu^{2+}	1.06×10^{-10}	10.63%	

simpler approximate solution. The concentration of free saliclyate is far below that of free copper. Furthermore, the stability constant for copper salicylate is very large. In these circumstances, all available salicylate will be complexed with free copper, so we may replace our calculated CuSal concentration with that of the free salicylate concentration we calculated in Example 14.1, 1.82×10^{-14} M. This is a trivial fraction of the total copper. Stream and lake water is likely to contain trace concentrations of other metals that are strongly bound by salicylate, such as Fe. This would further reduce the copper saliclyate activity.

CHAPTER 14: ORGANIC GEOCHEMISTRY

Adsorption Phenomena

The Hydrophobic Effect and Hydrophobic Adsorption

Water molecules near large nonpolar molecules such as long chain hydrocarbons cannot orient their polar OH bonds as they normally would (Figure 14.28). Since water molecules normally orient themselves in a manner that reduces electrostatic repulsions and minimizes interaction energy, the presence of a large nonpolar molecule is energetically unfavorable. As a result, solution of such substances, called *hydrophobic*



Figure 14.28. Disruption of water molecules by a large non-polar organic molecule, in this case a C_{15} n-alkane.

substances, in water is associated with a large ΔH_{sol} and large ΔG_{sol} . Thus one characteristic of hydrophobic substances is limited solubility in water. A second characteristic is that when they are present in solution, they are readily absorbed onto nonpolar surfaces, such as those of organic solids.

Hydrophobic adsorption differs from other types of adsorption phenomena in that adsorption occurs not a result of an affinity of the surface for the solute, but because incompatibility of the hydrophobic compound with water. When a hydrophobic molecule is located on a surface, water molecules are present on one side only, and there is less disruption of water structure than when water molecules are located on both sides. Thus the interaction energy is lower when the substance is located on a surface rather than in solution. Other types of adsorption involve electrostatic or van der Waals interactions or formation of bonds between the surface and the solute. While electrostatic and, particularly, van der Waals interactions generally contribute to hydrophobic adsorption, they are of secondary importance compared to the minimization of interaction energy between the solute and water.

Hydrophobic adsorption can be described by a simple model of partitioning of the hydrophobic species between water and an absorbent. The adsorption partition coefficient, K_{P} , is defined as:

 $K_P = \frac{\text{moles sorbate/mass solid}}{\text{moles solute/volume solution}}$ 14.14

and is typically expressed in units of liters/kilogram. The magnitude of the adsorption partition coefficient for hydrophobic species is related in a simple way to the solubility of the species in water, as illustrated in Figure 14.29a: the least soluble compounds are most strongly adsorbed. The aqueous solubility of such species may be further related to the octanol–water partition coefficient (Fig. 14.29b). Octanol is a largely nonpolar molecule, so that there is little structure or ordering of molecules in liquid octanol as there is in water. Thus there is no disruption of solvent molecules when a nonpolar solute is dissolved in octanol. The octanol/water partition coefficient is thus a measure of the "hydrophobicity" of organic molecules. The adsorption coefficient for hydrophobic substances on organic substrates may be empirically estimated by the following relationship:

$$K_{om} = b (K_{O/W})^a$$
 14.15

where K_{om} is the partition coefficient between organic solids and water, $K_{O/W}$ is the octanol/water partition coefficient, and *a* and *b* are empirical constants, with the value of *a* being around 0.8. From this, a more general expression for mixed organic/inorganic surfaces may be derived:

$$K_{\rm P} = b f_{\rm OC} (K_{\rm O/W})^{\rm a}$$
 14.16

where f_{oc} is the fraction of organic matter in the solid. Comparing 14.15 and 14.16, we see that $K_{om} = K_P / f_{OC}$ 14.17

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

In general, the solubility of organic molecules decreases with increasing molecular weight. This observation, known as *Traube's Rule*, is apparent from Figure 14.29. Small polar molecules such as phenoxyacetic acid and benzoic acid have higher solubilities and lower octanol/water partition coefficients than do large nonpolar ones such and DDT and PCB's (polychlorinated biphenols). It is easy to understand why this should be so: the larger the molecule, the greater volume of water whose structure is disrupted. In addition, and the tendency of a molecule to be absorbed and the strength of this adsorption increases with atomic weight. In part, this is true for the same reason solubility decreases: a greater volume of water is disrupted by large molecules. However, as we noted above, van der Waals interactions between the adsorbed substance and the surface also contribute to hydrophobic

adsorption. These interactions increase with increasing size of the molecule. Van der Waals interactions contribute a surface binding energy of roughly 2.5 kJ/mol per CH_2 group on the surface. Clearly, the more CH_2 groups involved, the more strongly the substance will be bound to the surface. For this reason, polymers are readily adsorbed to surfaces even if the adsorption free energy per segment is small. Adsorption of large polymers can be virtually irreversible.

Hydrophobic molecules are adsorbed preferentially to organic surfaces, which are largely nonpolar, rather than inorganic ones. Thus the degree to which hydrophobic substances are absorbed will depend of the fraction of organic matter that makes up solid surfaces. This is illustrated in Figure 14.30.

Other Adsorption Mechanisms

Many naturally occurring organic molecules contain both a polar and a non-polar Such molecules are called amphipart. patic. A good example is fatty acids, which, as we have seen, consist of hydrocarbon chains with a carboxyl groups attached to one end. The hydrocarbon chain is nonpolar and hydrophobic. The carboxyl group, however, is quite polar upon dissociation. The carboxyl group itself is readily soluble in water (as demonstrated by the high solubilities of simple carboxylic acids such as formic acid and acetic acid) and is hence *hydrophilic*. Carboxyl groups are therefore not subject to hydrophobic adsorption except at very low pH, where they are undissociated. They can, however, bind to polar solid surfaces in much the same way as inorganic ions. These include reactions such as *ligand* exchange:



Figure 14.29. (a) Soil/water adsorption partition coefficients for a variety of organic compounds as a function of water solubility determined by Chiou et al. (1979). (b) Octanol/water partition coefficients as a function of water solubility of organic compounds determined by Chiou et al. (1979).

CHAPTER 14: ORGANIC GEOCHEMISTRY



Figure 14.30. Adsorption partition coefficient for 1,4 dichlorobenzene plotted as a function of fraction of organic carbon in the solid absorbent. Other hydrophobic molecules show similar relationships. After Schwarzenbach and Westall (1980).

Geochemistry

S-OH + COOH(CH₂)_nCH₃ 14.18 \Rightarrow S-COO(CH₂)_nCH₃ + H₂O

where the carboxyl group, less its hydrogen, exchanges for an OH group bound to surface S. Polar function groups or organic anions may also bind to surfaces through *water bridging*, in which complexation with a water molecule solvating an exchangeable cation at a surface occurs:

$$S-M^+(H_2O) + COOHR$$
 14.19

$$\rightleftharpoons$$
 S–M⁺(H₂O)–COOHR

This mechanism is most likely to occur where M is strongly solvated (Mg^{2+} for instance). Where M is not strongly solvated, *cation bridging*, in there is which a direct bond between the acid functional group and the metal, can occur:

$$S-M^+ + COO^-R \rightleftharpoons$$

 $S-M^+-COO^-R$ 14.20

For cationic functional groups, such as quarternized nitrogen, *cation ion exchange* reactions such as:

$$S-M^++NH_3^+R \rightleftharpoons S-NH_3^+R+M^+$$
 14.21

where an organic cation replaces a metal cation at a surface. For anionic functional groups, such as carboxylic acids, anion ion exchange can occur. This is the analogy of reaction 14.18 with the signs reversed, e.g., a carboxyl group in anion form replacing a surface OH⁻ group.

All the above reactions may occur at either organic or inorganic surfaces. *Hydrogen bonding* in which a hydrogen is shared between a surface O atom and an O atom in a dissolved organic such as a carboxyl or phenol group, can occur at organic surfaces, for example:

$$S-H^+ + COO^-R \rightleftharpoons S-H^+ - COO^-R$$
 14.22

Hydrogen bonding is not restricted to acids. Organic bases, notably those containing nitrogen groups such as amines and pyridines, can also form hydrogen bonds with a hydrogen at a solid surface. Hydrogen bonding between dissolved organics and mineral surfaces is less important because the oxygens of mineral surfaces are not as electronegative as in organic compounds.

Many organic compounds will thus be subject to several types of adsorption: non-polar parts may be adsorbed to surfaces through hydrophobic bonding, while polar groups may bind through the mechanisms just described.

Dependence on pH

Figure 14.31 shows the effect of pH on the adsorption of humic acid on Al_2O_3 : the extent of adsorption is greatest at a pH of about 3 and is generally greater at low pH than at high pH. This pH dependence arises because the availability of hydrogen ions in solution will affect the charge on a solid surface in contract with that solution. At pH below the isoelectric point of a mineral, mineral surfaces will be protonated and will carry a positive charge; at higher pH's the mineral surface will bear a negative charge. Furthermore, dissociation and protonation of organic functional groups, which will affect the extent of adsorption through the mechanisms discussed above, is pH dependent.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY



Clearly, pH will also affect the mechanism of adsorption. Carboxyl acids groups of a humic acid molecule might bind to a surface through cation bridging at high pH where the surface has a net negative charge. At low pH, carboxyl groups will bind to a protonated surface through hydrogen bonding. At a pH close to that of the isoelectric point of a mineral, it surface will be neutral, in which case a humic acid would be subject to hydrophobic adsorption through its nonpolar parts. Thus the mechanism of adsorption and the strength of

Figure 14.31. Adsorption of humic acid on δ -Al₂O₃ as a function of pH. After Stumm (1992).

the bond formed between adsorbent and adsorbate will be influenced by pH.

Role in Weathering

In previous chapters we saw that adsorption and the formation of surface complexes plays a key role in weathering reactions. Organic acids can play a an important role in accelerating weathering

reactions in several ways: (1) by forming surface complexes, particularly surface chelates that weaken metal-oxygen bonds in the crystal and thus promote removal of metals from the surface, (2) by forming complexes with metals in solution, reducing the free ion activities and increasing ΔG of the weathering reaction, and (3) lowering the pH of solution (Drever and Vance, 1994; Bennett and Casey, 1994). In addition, organic substances serve as electron donors in the reductive dissolution of Fe and Mn oxides and hydroxides. These effects have been demonstrated in a variety of laboratory experiments (e.g., Furrer and Stumm, 1986; Zinder et al., 1986) and electron microscopy of minerals exposed to high concentrations of organic acids in both natural and laboratory situations (e.g., Bennett and Casey, 1994).

Furrer and Stumm (1986) investigated the effect of a variety of simple organic acids on dissolution of δ -Al₂O₃ and demonstrated a first order dependence of the dissolution rate on the surface concentration of organic complexes, i.e.:

$\Re = k [S \equiv L]$

where [S=L] is the surface concentration of organic complexes. Bidentate ligands that form mononuclear surface complexes seemed particularly effective in increasing dissolution rate. (There appears to be some evidence that formation of polynuclear surface



Figure 14.32. Rate of ligand-promoted dissolution of δ -Al₂O₃ as a function of surface concentration of organic ligands. Chelates forming five and six-member rings, such as those formed by salicylate, produced faster dissolution rates than 7-member rings, such as those formed by phthalate. Unidentate ligands, such as benzoate, have only a small effect on dissolution rate. From Furrer and Stumm (1986).

CHAPTER 14: ORGANIC GEOCHEMISTRY

complexes retards dissolution; Grauer and Stumm, 1982.) Five and six-member chelate rings were more effective in enhancing dissolution rate than seven member rings (Figure 14.32). Though monodentate ligands such as benzoate were readily adsorbed to the surface, they had little effect on dissolution rate. Similarly, Zinder et al. (1986) demonstrated a first order dependence of the dissolution rate of goethite (FeOOH) on oxalate concentration. Field studies show that high concentrations of organic acids, either natural or anthropogenic, clearly accelerate weathering (Bennett and Casey, 1994). However, in most circumstances, the concentrations of organic acids are low, and probably have only a small effect on weathering rates (Drever and Vance, 1994). Organic acids dissolved in formation waters of petroleum-bearing rocks may also enhance porosity by dissolving both carbonates and silicates (Surdam et al., 1984). This enhanced porosity is essential to the migration and recovery of petroleum.

SEdimentary Organic Matter and Coal and Oil Formation

Essentially all bodies of water harbor life, and therefore the production of organic carbon in aquatic and marine environments is ubiquitous. Most sedimentary rocks, however, contain rather little organic matter (a fraction of a percent is typical). This is a testament to the efficiency of life: virtually all organic carbon produced by autotrophs is subsequently oxidized by respiration, a process called *remineralization*. Indeed, most of the organic carbon synthesized in a body of water never reaches the sediment: it is consumed within the water column. Organic carbon that does manage to reach the bottom is subject to consumption by organisms living on and within the sediment. Although macrofauna play a role in remineralization, it is bacteria that are responsible for most of it (in soils, by contrast, fungi are often the dominant consumers of organic matter). Concentrations of bacteria in the surface layers of marine sediments are typically in the range of 10⁸ to 10¹⁰ cells per gram dry weight (Deming and Baross, 1993). The role of bacteria in the cycling of carbon, nitrogen, and sulfur is summarized in Figure 14.33.

These observations raise the question of why any organic matter survives. Why do most sediments contain some organic matter? How does it escape bacterial consumption? And why do some sediments,



Figure 14.33. Role bacteria in the cycling of carbon, nitrogen, and sulfur between inorganic and various organic forms. After Killops and Killops (1993).

CHAPTER 14: ORGANIC GEOCHEMISTRY

particularly those that give rise to exploitable petroleum and coal, contain much more organic matter? What special conditions are necessary for this to occur?

Organic matter preserved in ancient sediments, and particularly coal, gas, and oil, differ chemically from living organisms. Since these resources derive from the remains of once living organisms, we might ask how these chemical differences arise. Are the differences due to chemical transformations of simple organic molecules or selective preservation of more complex ones? Do the differences arise early, during the diagenesis of still young, poorly compacted sediment, or late, under the influence of heat and pressure?

We explore these questions in the following sections, where we examine sedimentary organic matter, its diagenesis, and the formation of petroleum, gas, and coal deposits.

FORMATION AND DIAGENESIS OF ORGANIC-RICH SEDIMENTS

PRESERVATION OF ORGANIC MATTER

It is primarily the particulate remains of phytoplankton that form organic matter in most marine and many aquatic sediments. Factors that affect preservation of these remains include the flux of organic matter to the sediment, bulk sediment accumulation rate, grain size, and availability of oxygen (Henrichs, 1993). The flux of organic matter to the sediment depends in turn on its rate of production in surface waters (biological productivity) and the depth of the overlying water column. Free floating single-celled autotrophs (algae and photosynthetic bacteria), collectively called phytoplankton, are responsible for almost all the primary production of organic carbon in marine ecosystems, as well as many fresh water ones. Productivity depends mainly on the availability of nutrients, which in the ocean depends on the proximity to coasts and ocean circulation (we will discuss nutrient distribution and productivity in more detail in the following chapter). Organic matter falling though the water column from the upper photosynthetic zone (200 meters at most) is rapidly remineralized by bacteria and animals in the water column. Hence the greater the water depth, the less organic matter reaches the sediment. In marginal marine environments, that is, those adjacent to continents, such as river deltas, bays, estuaries, and marginal seas, the land-derived remains of higher plants constitute a significant fraction of the accumulating organic matter. Such material is said to be *allochthonous* (i.e., derived from sources external to the water body). Organic matter produced within the immediate water body is called *autochthonous*.

Organic carbon concentrations are inversely correlated with grain-size for several reasons. First, low density organic particles can only accumulate where water velocities are low enough to allow finer particles to settle out. Second, a significant fraction of the organic matter in sediments may be present as coatings on mineral grains (Mayer, 1993). Small grains have higher surface areas per unit mass or volume, and therefore would have higher organic content. Mayer (1993) also argues that adsorbed organic matter is more refractory than that in discrete particles, meaning it is more likely to survive consumption by heterotrophs in the sediment. Third, the permeability of fine-grained sediments is lower than that of coarse-grained ones. Where permeability is low, the flux of oxygen into the sediments will also be low.

The availability of oxidants, and particularly oxygen, is, as one might expect, among the most important factors in the survival of organic matter. Simply put, the preservation of significant amounts of organic matter in sediment requires that the burial flux of organic matter exceed the flux of oxidants. The flux of oxidants depends on sedimentation rate, bioturbation, and diffusion, and their availability in the overlying water. Where the burial flux of organic carbon exceeds the downward flux of oxygen, the latter will ultimately be completely consumed and conditions will become reducing. At that point aerobic respiration must cease. This may occur either within the sediment, or within the water column itself. Situations where deep water becomes anoxic are rare in the modern ocean (indeed, in most of the deep ocean conditions do not become anoxic even in the sediment); it occurs only in a few basins where circulation of deep water is restricted, such as the Black Sea. However, anoxicity appears to have been more common at certain times in the geologic past, such as the

CHAPTER 14: ORGANIC GEOCHEMISTRY

Cretaceous, when ocean circulation was different. Anoxicity is perhaps more common in lakes, where the abundance of nutrients is higher than in the open ocean.

Whether preservation of high organic matter concentrations in sediments requires anoxic bottom water is a matter of debate. Calvert and Pederson (1992) point out that sediments accumulating in oxic and anoxic basins have similar organic carbon contents. They also argue that extent of decomposition of marine organic matter is similar under oxic and anoxic conditions, though terrestrial organic matter to be degraded less by sulfate reducers. On the other hand, Killops and Killops (1993) point out that ancient lipid-rich sediments of the sort likely to give rise to petroleum are generally finally laminated, implying a lack of bioturbation and therefore anoxic conditions at the sediment-water interface.

Diagenesis of Marine Sediments

Diagenesis in the context of organic matter refers to biologically induced changes in organic matter composition that occur in recently deposited sediment. Actually, these changes begin before organic matter reaches the sediment as organic matter sinking through the water column is fed upon by both the macrofauna and bacteria. Indeed, a significant proportion of the organic matter reaching the sediment does so in the form of fecal pellets of everything from zooplankton to whales. Decomposition continues once the organic matter reaches the sediment surface. Burial by subsequently accumulating sediment eventually isolates it from the water. Where the burial flux of organic matter is high enough, oxygen is eventually consumed and as the organic matter is buried to progressively greater depth, it is attacked by a series of bacterial communities utilizing a progression of electron receptors (oxidants) at decreasing pɛ. We can predict the order of the use of these oxidants from the ΔG of the redox reactions involved, shown in Table 14.8. Thus, moving downward in a column of accumulating sediment, we expect to see, following consumption of free oxygen, a series of zones where nitrate, Mn(IV), Fe(III), sulfate, and nitrogen reduction occur. Each of these zones will be colonized by a bacterial flora adapted for conditions in that zone[§]. Once all oxidants are consumed, respiration continues through fermentation.

The bulk of the organic matter in sediments exists in solid form, yet only dissolved compounds can cross cell membranes and be a useful source of nutrition to microbes. For this reason, bacteria release exoenzymes that first break insoluble complex organic molecules into smaller soluble ones. Complex organic molecules usually cannot be oxidized completely by a single organism, because no single organism is likely to produce all the necessary enzymes. Instead, macromolecules are broken down by consortia of bacteria. In each step, some energy is released and smaller molecules are produced as waste; these are subsequently attacked by other bacteria. Thus proteins, carbohydrates, and lipids are broken down into amino acids, simple sugars, and long-chain fatty acids. These smaller molecules can be attacked by fermenting bacteria that produce acetic acid, other short chained carboxylic acids, alcohols, hydrogen, and CO_2 . In the final step, these are converted to methane (CH₄) by methanogenic bacteria. During this process, the remains of bacteria themselves can become a significant part of the sedimentary organic matter.

The stepwise oxidation results in an interdependence between the various bacterial species within each community, as many species are dependent on the "waste" products of other species. There is also a more general interdependence between communities in sediments. For example, anaerobic communities depend on aerobic ones to produce an anoxic environment. Reduced compounds, for example sulfide, ammonia, and methane, which are waste products of anaerobic communities in the deep levels, diffuse upward into the oxic zone where they are oxidized by various photosynthetic, chemosynthetic, and methyltrophic (methane-oxidizing) bacteria.

[§] Examined at the microscopic level, separation of bacterial species is not quite this simple or complete. For example, within the oxic zone, there are anoxic microenvironments where anerobic bacteria flourish.

CHAPTER 14: ORGANIC GEOCHEMISTRY

Both the abundance of organic matter and of bacteria decrease with depth in marine sediments, the highest concentrations of both being found in the upper 10 cm. There is also evidence that decomposition rates decrease when conditions become anoxic (summarized in Henrichs, 1993). Thus most remineralization occurs in the uppermost 1 or 2 meters and the bulk of the organic matter buried beneath this depth is preserved long-term (Henrichs, 1993).

TABLE 14.8 FREE ENERGY CHANGES FOR BACTERIAL REACTIONS			
Reaction	$\Delta G (kJ/mol CH_2O)$		
$CH_2O + O_2 \rightarrow CO_{2(aq)} + H_2O$	-493		
$5CH_2O + 4NO_{\overline{3}} \rightarrow 2N_2 + 4HCO_{\overline{3}} + CO_{2(aq)} + 3H_2O_{2(aq)} + 3H_2$	O -472		
$CH_2O + 3CO_{2(aq)} + H_2O + 2MnO_2 \rightarrow 2Mn^{2+} + 4HCO$	$\bar{3}$ -348		
$^{+3}CH_2O + 4H^{+} + 2N_2 + 3H_2O \rightarrow 3CO_{2(aq)} + 4NH_4^{+}$	-125		
$CH_2O + 7CO_{2(aq)} + 4Fe(OH)_3 \rightarrow 4Fe^{2+} + 8HCO_{\overline{3}} + 3$	3H ₂ O -103		
$2CH_2O + SO_4^{2-1} \rightarrow H_2S + 2HCO_{\overline{3}}$	-99		
$2CH_2O \rightarrow CH_4 + CO_{2(aq)}$	-88		
$^{+}3CH_{2}O + 2N_{2} + 7H_{2}O \rightarrow 3CO_{2(ag)} + 4NH_{4}(OH)$	-54		

Modified from Berner (1981).

^tBecause the speciation of ammonia is pH dependent, the ΔG of the nitrogen fixation reaction depends strongly on pH.

What molecules are preserved? As we might expect, the simple organic molecules such as amino acids, sugars, and short-chained carboxylic acids are rapidly decomposed by bacteria (time scales of days to weeks). More complex molecules, such as polysaccharides and fatty acids appear to decompose over a few months to a few years (Henrichs, 1993). Certain classes of compounds, principally ones that serve as cellular structural materials (e.g., components of cell walls), appear to be particularly resistant to bacterial decomposition and form the bulk of the preserved organic matter. Examples of these resistant materials are algaenans, which are found in the cell walls of marine algae, and phlorotannins (De Leeuw and Largeau, 1993). Allochthonous material derived from higher plants may also contribute a number of resistant aromatic-rich compounds (see below) to sediments in marginal marine environments. However, a small fraction of readily metabolized compounds is also preserved. Even older sediment, in which there has been ample opportunity for bacterial decomposition, contains low concentrations of such compounds. These molecules may survive because they are located in micro-environments that shield them from bacterial enzymes. Thus labile molecules packaged within resistant structures (e.g., spores, pollen) can be preserved. Adsorption to inorganic particulates may also afford a degree of protection. Enzyme-catalyzed hydrolysis often requires a precise and unique physical alignment of the enzyme and reactant. The part of the surface of an organic molecule adsorbed onto an inorganic surface will not be accessible to the enzyme. Organic molecules partly or wholly contained within micropores on the solid surface will be even more protected. Similarly, we might expect proteinaceous material in carbonate shells to be somewhat protected from bacterial enzymes.

Diagenesis of Aquatic Sediments

On the whole, diagenesis in fresh water sediment is similar to marine diagenesis. As is also the case in marine sediments, most of the organic detritus in aquatic environments originates from plants, animals contributing less than 10%. Perhaps the principal difference in diagenesis between large lakes and the ocean is the much lower sulfate concentrations in lakes. Sulfate is important both as an oxidant and because sulfur can incorporated into organic molecules (primarily lipids) during early diagenesis, a process known as "natural vulcanization". Because fresh waters have low sulfate concentrations, the zone of sulfate reduction is restricted and vulcanization does not occur. Otherwise, the same sequence of oxidant usage and decomposition occurs, and most of the remineralization occurs near the sediment-water interface.

In large lakes, the bulk of the organic matter reaching the sediment may be autochthonous (i.e., produced within the lake itself, primarily by phytoplankton), as is the case in marine environments. Often, however, allochthonous organic matter derived from terrestrial plants constitutes a substantial part of the organic flux to aquatic sediment. Higher plants living within the water may also

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

contribute organic matter, and such material is dominant in swamps and marshes. The significance of this is that higher plants contain a greater abundance of aromatic compounds than algae. We found earlier in the chapter that aromatic compounds are often particularly stable. Thus it is no surprise to find that aromatics such as lignins, tannins, gums, curtans, and suberans, all produced by higher plants, are particularly resistant to bacterial decomposition and hence are more easily preserved in sediment.

Coal is formed by the compaction and diagenesis of organic-rich sediment, called *peat*, deposited in swamps. In contrast to petroleum, which can form in sediments containing only a few percent of organic matter, coal forms from sediments in which organic content is the dominant constituent. There are many examples of modern environments where such organic-rich is now accumulating. Production of peat in these environments is a consequence of a number of factors. The first of these is productivity. Wetlands are generally characterized by high biological productivity, hence there is a high flux of organic matter to the sediment. The second factor is hydrology. Peat formation occurs where there is an excess of inflow and precipitation over outflow and evaporation. This maintains a water-logged soil as peat accumulates. Water-logged conditions restrict the flux of oxygen into the sediment, resulting in conditions becoming anoxic immediately below the sediment-water interface. The third factor is the abundance of dissolved organic acids, some resulting from decomposition, others exuded by mosses and bacteria. These acids lower pH and inhibit the activity of decomposing bacteria. Finally, the primary producers in such environments are bryophytes (mosses) and vascular plants. As we noted above, these contain relatively high concentrations of aromatic compounds, which are more resistant to decomposition than the aliphatic compounds that predominate in algae and bacteria. Nevertheless, less than 10% of the organic production in these environments is preserved as peat, the rest being exported or recycled.

At present, the largest peat-forming environments are high-latitude (>45°) moors and bogs. These moors are typically dominated by a few species of moss (*Sphagnum*) that account for most of the accumulating organic matter. Other modern peat producing environments include coastal swamps, such as the Mahakam Delta of Indonesia, and temperature and tropical lowland swamps.

SUMMARY OF DIAGENETIC CHANGES

Changes in sedimentary organic matter occurring as a result of diagenesis can be summarized as follows:

- Functional groups, such as carboxyl, amigos, and hydroxy, are preferentially removed from their parent molecules.
- Loss of functional groups such as COOH and OH decreases the oxygen, and to a lesser degree, the hydrogen, content of the organic matter.
- The abundance of readily metabolized organic compounds decreases. Nucleic acids and amino acids and related compounds appear to be the most labile (most readily destroyed), followed by carbohydrates, particularly simple ones and those synthesized for energy storage (e.g., starch) rather than structural (e.g., cellulose) purposes. The simple molecules in these groups (e.g., amino acids, glucose) are most labile of all. Lipids appear to be somewhat less labile.
- Unsaturated compounds decrease in abundance compared to their saturated equivalents due to hydrogenation of double carbon bonds.
- Aliphatic compounds decrease in abundance compared to aromatic ones. This results partly from aromatization of unsaturated aliphatic compounds and partly from the more resistant nature of aromatics.
- Short-chained molecules (e.g., alkanes, fatty acids), decrease in abundance relative to their longchain equivalents.
- Hydrolysis of complex molecules produces a variety of molecular fragments that subsequently recombine with other molecules to produce new ones not present in the original biota. For example, phytol, produced by degradation of chlorophyll-a, and phenols, which can be produced by degradation of a variety of aromatic compounds, condense to form phenol-phytol compounds.

CHAPTER 14: ORGANIC GEOCHEMISTRY

- In high sulfur environments, such as marine sediments, addition of H₂S (produced by sulfate-reducing bacteria) is incorporated into carbon double bonds in long-chain compounds such as isoprenoids to produce thiol functional groups. These can subsequently form cyclic structures and ultimately aromatic thiophenyls. This process is known as natural vulcanization.
- Condensation of a variety of molecules and molecular fragments into complex macromolecules.
- All along, bacterial remains are progressively added to the mixture, and are progressively decomposed along with the organic matter originally deposited.

The principal product of these processes is *kerogen*, the name given to the mixture of complex organic compounds that dominates the organic fraction in sediments.

KEROGEN AND BITUMEN

Kerogen is defined as sedimentary organic matter that is insoluble in water, alkali, non-oxidizing acids, and organic solvents (such as benzene/methanol, toluene, methylene chloride). It is usually accompanied by a smaller fraction of soluble organic matter, called *bitumen*. Kerogen, an inhomogeneous macromolecular aggregate, constitutes 90 percent or more of organic matter in sedimentary rocks (much of the remainder being dispersed bitumen). It is interesting that kerogen is by far the most abundant form of organic carbon on Earth; It is three orders of magnitude more abundance that coal, petroleum, and gas, and four orders of magnitude more abundant than the living biomass. Kerogen has the interesting and significant property that upon heating in the laboratory, a procedure known as *pyrolysis*, it breaks down to produce a variety of hydrocarbons similar to those found in natural petroleum. However, kerogen varies widely in its petroleum potential. Kerogen that is rich in aliphatic compounds, generally derived from aquatic and marine algae, has good petroleum potential and is called *sapropelic kerogen*. Kerogen derived principally from the remains of higher plants is rich in aromatic compounds, sometimes called *humic kerogen*, has poor petroleum potential.

Carbon and hydrogen are the main constituents of kerogen. Hydrogen concentrations range from 5 to 18% (atomic), depending on type and degree of evolution. Oxygen concentrations typically range from 0.25 to 3%, again depending on type and degree of evolution. Besides C, H, and O, kerogen typically

contains 1-3% N and 0.25-1.5% S (though the latter can be higher). A variety of trace metals, notably V and Ni, are also found in kerogen.

The structure of kerogen and the manner in which it forms is only partly understood. It appears to consist of nuclei cross-linked by chain bridges. The nuclei consist of stacks of two to four sheets of condensed aromatic rings, with roughly 10 rings per sheet, giving them a dimension of less than 10Å. A variety of functional groups and alkyl chains are attached to the sheets. The bridges linking the nuclei may consist of linear or branched aliphatic chains, oxygen or sulfur functional bonds (e.g., ketones, esters, thiols, etc.). The bridges may also have functional groups attached to them. This structure appears to act as a "molecular sieve" and can trap compounds such as lipids within it. It was once widely believed that kerogen, like humic substances, originated principally by condensation of low molecular weight biomolecules (amino acids, sugars, fatty acids, phenolics, etc.) produced by bacterial decomposition of more complex ones. In this view, humic substances are viewed as the precursors to kerogen. More recently, Tegelaar et al. (1989) proposed that the principal contributors to kerogen are highly resistant molecules (such as tannins, algaenans, etc.) that constitute only a



Figure 14.34. Compositional difference between kerogen maceral groups liptinite, exinite, vitrinite, and inertite. A plot of the H/C ratio vs. the O/C ratio such as this is commonly called a *van Krevelen diagram*. Modified from Tissot and Welte (1984).

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

small proportion of the original organic matter. This hypothesis, known as *selective preservation*, has gained wide acceptance. The difference between these views, however, is not as great as it might initially seem, as it is widely agreed that even highly resistant molecules experience some bacterial alteration (e.g., defunctionalization, hydrogenization) and that some degree of recondensation is involved in kerogen formation.

Microscopic examination reveals that kerogen consists of identifiable plant remains, amorphous material, and rarer animal remains. The amorphous material in kerogen may occur as mottled networks, small dense rounded grains, or clumps. The microscopically identifiable constituents are called *macerals*. Schemes for classifying macerals were first developed to describe coal and later applied to kerogen. Unfortunately, there a number of classifications in use (see Whelan and Thompson-Rizer, 1993, for a summary), which can lead to considerable confusion. Here we follow Killops and Killops (1993) and divide them into four groups. These groups differ in both composition (Figure 14.34) and origin. The *inertite* group consists of carbonized remains formed by rapid oxidation under aerobic conditions. One mechanism by which inertite forms is probably wild fires in peat-producing environments[†]. Inertite may include the carbonized remains of just about anything: woody tissue, fungi, spores, cuticles, resins, algae, etc. Inertite has low H/C and O/C ratios and, as its name implies, is rather inert. Vitrinite is preserved woody tissue. There are two common macerals in this group: telenite and collinite. Telenite has a defined cell structure while collinite is colloidal, derived from solidified humic gels. Exinite includes lipid-rich materials derived from leaf cuticle, spores, pollen, algae, plant waxes, resins, fats, and oils. The fourth group, liptinite, is similar in many respects to exinites, but whereas exinites have recognizable shapes, liptinites are amorphous

bodies. Liptinites are derived primarily from algal remains and usually have higher H/C ratios than exinites. These four maceral groups react differently to heating: vitrinite produces a fused carbon residue, inertites show no visible change, and exinites and liptinites transform into gas and tar.

Kerogen Classification

Kerogen is usually classified into one of three types, based on bulk H/C and O/C ratios (Figure 14.35). Type I kerogen has a high H/C (atomic) ratio (≥ 1.5) and a low O/C (atomic) ratio (<0.1). It is rich in lipids, especially long-chain aliphatics, and has high petroleum potential. It is derived primarily from algal and bacterial remains, often deposited in aquatic or estuarine environments. Kerogen found in the Eocene Green River Shale of the western U.S. is a good example. Type II kerogen, the most common type, has intermediate H/C (~ 1.25) and O/C (<2.0) ratios. It is derived primarily from planktonic and bacterial remains deposited in marine environments (though remains of high plants can contribute as well). Because of its marine origin, it is often sulfur rich. Its lipid content and oil potential are somewhat lower than Type I kerogen. Type III kerogen has low H/C ratios (<1.0) and high O/C ratios



Figure 14.35. H/C and O/C ratios of the three types of Kerogen. Open triangles: Type I, closed red circles: Type II, crosses: Type III. Arrows show the direction of compositional evolution during diagenesis and subsequent thermal maturation (catagenesis and metagenesis). Dashed lines show boundaries between regions of diagenesis, catagenesis, and metagenesis. After Tissot and Welte (1984).

⁺ One such modern environment is the Okefenokee Swamp in southern Georgia (USA). Wildfires often follow major droughts that occur at ~25 year intervals. These fires may burn the peat to a depth of 30 cm.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

(<0.3). It is rich in aromatic and poor in aliphatic structures. It is formed principally from the remains of vascular plants. Its oil potential is poor, but can be a source of gas (particularly methane). A comparison of Figures 14.34 and 14.35 shows that Type I kerogen is related to liptinite macerals, Type II to exinites, and Type III to vitrinite. High sulfur Type II kerogen (denoted TypeII-S) can contain 10% or more sulfur by weight. A fourth kerogen type (Type IV), which more or less corresponds to the inertite maceral group, is sometimes also defined. However, inertite has no petroleum potential, so there is less interest in this type.

Bitumen

The fraction of sedimentary organic matter that is soluble in carbon disulfide is called bitumen and includes solids, liquids, and gases. At the end of diagenesis, bitumens generally constitute less than 3 to 5 percent of the total organic carbon (the remainder being kerogen), though this figure is occasionally higher. During subsequent thermal evolution, however, the fraction of bitumens increases at the expense of kerogen (see below). Bitumen consists primarily of 3 fractions: *asphaltenes*, *resins*, and *pe-troleum*. These fractionations are defined, like humic substances, by their solubility. *Maltenes* are soluble in light hydrocarbons such as hexane, whereas asphaltenes are not. Maltenes can further be separated in to petroleum, which consists of a variety of hydrocarbons, and resins. Resins and asphaltenes, unlike hydrocarbons, are rich in heteroatoms such as N, S, and O. Resins tend to be somewhat richer in hydrogen (H/C atomic ~ 1.4) and poorer in N, S, and O (7-11 wt %) than asphaltenes (H/C atomic ~ 1.2, N, S, O ~ 8 - 12 %). Both have molecular weights greater than 500 and commonly several thousand. Asphaltenes appear to be structural similar to kerogen, consisting mainly of aromatic nuclei link by aliphatic units. They can be thought of as small fragments of kerogen.

The hydrocarbon fraction consists of both aliphatic and aromatic components. The aliphatic component can further be divided into acyclic alkanes, referred to as *paraffins*, and cycloalkanes, referred to as *naphthenes*. The lightest hydrocarbons, such as methane and ethane, are gases at room temperature and pressure; heavier hydrocarbons are liquids whose viscosity increases with the number of carbons. The term *oil* refers to the liquid bitumen fraction. *Pyrobitumens* are materials that are not soluble in CS_2 but break down upon heating (pyrolysis) into soluble components.

THERMAL EVOLUTION OF ORGANIC MATTER AND PETROLEUM GENERATION

As sedimentary organic matter is buried, it experiences progressively higher temperatures and pressures. Although most bacterial decomposition occurs quickly, in the upper meter or so, it may continue at a much slower pace almost indefinitely. Indeed, bacteria have been found in subsurface rock at temperatures of up to 75° C and depths of nearly 3 km. As bacterial activity ceases, a number of new reactions begin as the organic matter attempts to come to equilibrium with higher temperature and pressures. These reactions, in which kerogen breaks down into a variety of hydrocarbons and a refractory residue, are collectively called *catagenesis*. As temperatures in the range of 100 to 150° C are reached, a complex mixture of hydrocarbons, *petroleum*, is produced, along with less amounts of other bitumens, asphaltenes and resins. Collectively, this bitumen fraction is called oil or *crude oil* and is, of course, of great economic interest. At temperatures above 150-175° C, methane and graphite are the ultimate products, created in a process called *metagenesis*.

One of the principal effects of diagenesis is the condensation of the complex macromolecules from simpler ones. During catagenesis, this process is reversed as kerogen disproportionates into comparatively simple hydrogen-rich molecules (hydrocarbons) and a hydrogen depleted carbon residue. The hydrogen-rich phase is mobile and will migrate out of the source rock if a migration pathway exists. The refractory carbon-rich residue is immobile and remains in place.

Whereas diagenesis is a result of microbial metabolic activity, catagenesis is a physical (i.e., thermodynamic) response to increasing temperature and pressure. Kerogen molecules undergo rearrangement to take on a more ordered and compact structure. As this occurs, the alignment of nuclei, each composed of two or more aromatic sheets, becomes increasingly parallel, the number of sheets per nucleus increases, and the space between them decreases. Aliphatic units that are peripheral to

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

the aromatic nuclei as well as those that bridge nuclei are progressively eliminated, with longer chains eliminated preferentially. Since most of the remaining functional groups in kerogen are attached to these aliphatic units, these are also eliminated. Heteroatoms, N, S, and O, are also eliminated in this process. Aromatic units increase in abundance relative to aliphatic units. This results from aromatization of cyclic aliphatic structures as well as elimination of aliphatic structures. Unsaturated n-alkanes have 2 or more hydrogens per carbon atom, whereas aromatic units have 1 or fewer hydrogens per carbon. Thus the compositional effect of catagenesis on kerogen is a decrease in the H/C ratio, as well as a continued decrease in the O/C ratio. This compositional evolution is illustrated by the arrows in Figure 14.35.

The degree of thermal maturation of kerogen can be monitored from its H/C and O/C ratios. In the "oil window", the point where maximum hydrocarbon generation occurs, the H/C ratio is less than 1 and the O/C ratio less than 0.1. Kerogen with H/C ratios lower than 0.5 is over-mature, i.e., it has already entered the metagenesis stage where methane is the principal hydrocarbon product. Kerogen maturity can also be monitored by measuring *vitrinite reflectance*. Kerogen in the diagenetic stage reflects light only weakly, but as its structure becomes more dense and ordered during catagenesis, more incident light is reflected. Vitrinite reflectance is determined by polishing a specimen and then comparing the fraction of incident light reflected to that of a calibrated standard using a microscope photometer. Vitrinite reflectance increases from about 0.2% in recent sedimentary organic matter to 4% or more in over-mature kerogen. In the oil-generating stage of catagenesis, vitrinite reflectance is typically in the range of 0.6 to 1.3%.

Figure 14.36 summarizes the generation of oil and gas as a function of temperature. During catagenesis, heteroatom bonds are the first to be broken as they are generally weaker than carbon-carbon bonds. Hydrocarbons released during this stage are thus those attached to the kerogen structure with heteroatoms or merely trapped within it. Thus the hydrocarbon fraction of bitumens in immature kerogen is dominated by "geochemical fossils" or biomarkers, i.e., molecules that have lost their functional groups but whose basic skeleton is preserved. A few examples include:



Figure 14.36. General scheme for hydrocarbon generation as a function of depth and temperature. Composition of the hydrocarbons generated is shown in the graphs to the right. Temperature and depth scales assume a geothermal gradient of 40° per km. After Tissot and Welte (1984).

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

- Odd-carbon numbered n-alkanes in the range of C₂₅ to C₃₃, which are derived from even-carbon numbered fatty acid and alcohol components of cuticular waxes of higher plants.
- C₁₅ and C₁₇ n-alkanes derived from hydrocarbons and corresponding fatty acids in algae.
- Isoprenoids, most notably pristane (C₁₉) and phytane (C₂₀) derived from phytol in chlorophyll.
- Pentacyclic triterpenoids and their derivatives, including those derived from higher plant resins and hopanoids, derived from bacteria.

As temperature increases, carbon-carbon bonds are also broken, a process called *cracking*. Carbon-carbon bonds in the center of chains are slightly weaker than those on the ends. As these begin to break, hydrocarbon fragments are released that progressively dilute biomarkers. Also because of this effect, the size of hydrocarbons evolved decreases with increasing maturity. Thus the hydrocarbons generated in the "oil window" of catagenesis show a maximum abundance at relatively low carbon number ($\sim C_{10}$) and steadily decreasing abundance with increasing carbon number.

As temperatures approach and exceed 150° C, even smaller hydrocarbons (\leq C₅) become dominant. These are gases at surface temperature and pressure. Dissolved in them, however, are lesser amounts of longer chains (\geq C₆). These condense to liquids upon reaching the surface and hence are called *condensates*. Hydrocarbons that are gas-dominated yet contain a significant amount of longer hydrocarbons are called *gas condensates*, and this stage of catagenesis, corresponding roughly to 150 to 180°C, is called the "wet gas zone". At higher temperatures, the liquid hydrocarbons are completely eliminated by C-C bond breaking. Eventually, all C-C hydrocarbon bonds are broken, leaving methane as the sole hydrocarbon, accompanied by a nearly pure carbon residue. This stage of evolution is referred to as metagenesis or the "dry gas zone".

Rates of reactions involved in catagenesis show an exponential temperature dependence, as we might expect. Reaction rates roughly double for every 5 to 10° increase in temperature. Because of this, catagenesis depends on not just on temperature, but on time as well, or more specifically, on the heating rate. Heating rate in turn depends on (1) the burial rate and (2) geothermal gradient. The burial rate primarily on rate at which the sedimentary basin subsides. The geothermal gradient at the surface of the Earth varies widely, from $10^{\circ}/\text{km}$ to $80^{\circ}/\text{km}$ or even higher in geothermal areas. Values at the low end of this range are typical of old continental shields; higher values are typical of rifts and oceanic crust. Petroleum deposits often occur in subsiding basins associated with tectonic activity, thus geothermal gradients can be high. Gradients of from 25 to 50°C/km are perhaps most typical for petroleum producing environments. As a result of kinetics and variations in burial rate and geothermal gradient, the time required for petroleum generation will vary. In western Canada, Devonian sediments were slowly buried in a region of low geothermal gradient and oil generation followed deposition by 300 million years. In contrast, 10 million year old upper Tertiary sediments in the Los Angeles Basin are already generating petroleum. The temperature required for the onset of petroleum generation varies inversely with time. For example, this threshold is about 60° C in Lower Jurassic sediments of the Paris Basin, but is 115° in Mio-Pliocene sediments of the Los Angeles Basin. The temperature and depth scales in Figure 14.36 correspond to a relatively high geothermal gradient (40°C/km). Reaction rates also depend on the type of kerogen involved. Labile reactive kerogen (Type I) reacts at relatively low temperatures; refractory Type III can require substantially higher temperatures for petroleum generation (as high as 250°C). Since long aliphatic chains are unstable at these temperatures, the principal product of such refractory kerogen is methane.

Migration and Post-Generation Compositional Evolution

Most petroleum source rocks are fine-grained. Subjected to the pressure of burial, their porosities are typically quite low, hence liquid and gaseous hydrocarbons are expelled once the source rock becomes saturated. The mechanisms of migration of hydrocarbons are not fully understood, but probably involve both passage through microfractures and diffusion through the kerogen matrix. Migration will continue until the petroleum reaches either an impermeable barrier, a "trap", or the surface. From the standpoint of economic recovery, the ideal situation is a trap, such as a clay-rich sediment, overlying a porous and permeable "reservoir" rock such as a sandstone. Expulsion efficiencies vary

CHAPTER 14: ORGANIC GEOCHEMISTRY

with kerogen type. In Type I kerogen, nearly all the oil can be expelled from the source rock. In Type III kerogen and coal, however, most or all of the oil may remain trapped in the source rock and be ultimately cracked to gas.

The quantity and quality of the petroleum generated depends largely on the type of organic matter. Since petroleum tends to migrate out of the source rock as it is created, it is difficult to judge the

amount of petroleum generated from field studies. However, both mass balance calculations on natural depth sequences and laboratory pyrolysis experiments on immature kerogen give some indication of the petroleum generation potential 1984, (Tissot and Welte, Rullkötter, 1993). Type I kerogen yields up to 80% light hydrocarbons upon pyrolysis. Mass balance studies of Type II kerogen indicate a hydrocarbon generation potential of up to 60%. Type III kerogens yield much less hydrocarbon upon pyrolysis (<15%).

Chemical changes may occur in several ways during and after migration. Fractionation during migration can occur as a result of the differing diffusivity and viscosity of hydrocarbons: light hydrocarbons are more diffusive and less viscous than heavy ones. As a result, they will migrate more readily and the hydrocarbons in a reservoir are often enriched in the light fraction compared to the source rock. Polar constituents in oil, asphaltenes and resins, may be absorbed by mineral surfaces and are less readily expelled from the source rock, resulting in a depletion in these components in oil in reservoir rocks compared to source rock bitumen. The more water soluble components of petroleum may dissolve in water, either flowing through a reservoir or encountered by migrating petroleum. This process, called water washing, will depleted the petroleum in these water-



Figure 14.37. Ternary diagrams representing the composition of crude oils. (a) Isofrequency contours of hydrocarbons boiling above 210°C in 541 crude oils divided between aromatics, cyclic-, and nand isoalkanes. (b) Isofrequency contours of saturated hydrocarbons, aromatic hydrocarbons, and NSO compounds (wt percent in the fraction boiling above 210° C) in 636 crude oils. From Tissot and Welte (1984).

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

soluble components. Aerobic bacteria encountered by petroleum can metabolize petroleum components, a process called *biodegradation*. Long, unbranched alkyl chains are preferentially attacked, followed by branched chains, cycloalkanes, and acyclic isoprenoids. Aromatic steroids are the least affected. Finally, further thermal evolution can occur after migration, resulting in an increase in methane and aromatic components at the expense of aliphatic chains.

Composition of Crude Oils

Figure 14.37 summarizes the compositions of crude oils. Average "producible" crude oils contain 57% aliphatic hydrocarbons (with a slight dominance of acyclic over cyclic), 29% aromatic hydrocarbons, and 14% resins and asphaltenes. On an elemental basis, it consists approximately of 82-87% C, 12-15% H, 0.1-5% each of S and O, and 0.1-1.5% N. The distribution of n-alkanes of differs widely between various types of crudes as may be seen Among cycloalkanes, Figure 14.38. those with 2 to 4 rings generally pre-Alkylated dominate. compounds dominated the aromatic fraction; those with one to three additional carbons are most common. Aromatics decrease in abundance with increasing number of rings, so that benzene derivatives (one ring) are most common, followed by naphthalenes (two rings),



Figure 14.38. Distribution of n-alkanes in different crude oils. From Tissot and Welte (1984).

etc. Molecules containing both saturated and unsaturated rings (napthenoar-omatics) are also present, typically in an abundance of 5 wt. %. Upon primary distillation, typical crude oil yields 27 volume percent gasoline (C_4 - C_{10} compounds), 13% kerosine (C_{11} - C_{13}), 12% diesel fuel (C_{14} - C_{18}), 10% heavy gas oils(e.g., heating oil) (C_{19} - C_{25}), and 20% lubricating oil (C_{26} - C_{40}) (Royal Dutch Shell, 1983). The ratio of these products can be changed by further refining processes such as solvent extraction, thermal cracking, catalytic cracking, etc., so that, for example, the gasoline yield can be as high as 50%.

Compositional Evolution of Coal

Coal, as we noted earlier, forms from organic-rich sediments typically deposited in swamps. Two types of coals are recognized: sapropelic and humic. Humic coals are by far the most common. They are bright, usually stratified, rich in aromatics and composed primarily of the remains of higher plants. Less common sapropelic coals are dull, rarely stratified, and derived from lipid-rich organic matter such as the remains of algae (boghead coals or torbanites) or spores (cannel coals). The primary maceral group of humic coals is vitrinite, that of sapropelic coals is exinite.

The evolution of coal, illustrated in Figure 14.39, is generally broken into two phases: *peatification* and *coalification*. Coalification is subdivided into *biochemical* and *geochemical* stages. Together, peatification and the biochemical stage of coalification are equivalent to diagenesis, while the geochemical stage of coalification is comparable to catagenesis. During peatification, bacterial and fungal attack results in depolymerization and defunctionalization of the original biomolecules. This process is begun by aerobic organisms and continued by anaerobic bacteria once conditions become reducing. This is accompanied by the evolution of various gases (e.g., CO₂, NH₃, N₂, CO₂, H₂O) and condensation of the degradation products into humic substances. As in diagenesis, the concentrations

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

of the most labile components decrease, while those of more refractory ones increase. The latter include lignins and tannins, and lipids derived from leaves, spores, pollen, fruit, and resin. Another important process during peatification is compaction and expulsion of water.

During biochemical coalification, continued loss of functional groups drives the O/C ratio to lower values with only a slight decrease in H/C ratio. Remaining labile components are metabolized and refractory material continues to condense to aromatic-dominated structures. The final product of the diagenetic phase is *brown coal*, which contains 50-60% C and 5-7% H. This material may be accompanied by a small bitumen fraction, derived primarily from lipid components.



Figure 14.39. Chemical evolution of coals. After Killops and Killops (1993).

Temperature and pressure increase with burial and this initiates the geochemical stage of coalification. Coal at this stage contains 1-2% N and generally less than 1% S. Continued compaction results in a continued decrease in the water present. Loss of functional groups produces further reduces the O/C ratio with only minor decrease in the H/C ratio. By the time the O/C ratio reaches 0.1, most of the functional groups have been lost. The resulting material is now called *bituminous coal*, which requires temperatures in the range of 40 to 100° C. Bituminous coal has a fairly bright appearance and contains 75% or more C and the water content has decreased to less than 10%. At this point, vitrinite reflectance reached 0.5% and 70% or more the carbon is in aromatic. During peatification and the early stages of coalification, CO₂ and H₂O are the primary volatile species produced.

With further heating, aromatization of cycloalkyl structures becomes the dominant process, releasing methane. Aromatization and loss of methane reduce the H/C ratio, which decreases rapidly upon further heating. In the temperature range of 100 to 150° C, *anthracite* is formed as the H/C ratio decreases below 0.5%. Anthracite is characterized by vitrinite reflectance of >2.5% and a carbon content of greater than 90%. Ninety percent or more of this carbon is in aromatic structures. As in kerogen, these aromatic structures initially take the form of randomly ordered nuclei. During the geochemical stage of coalification, these nuclei become increasingly ordered, so that by the anthracite stage, they are arranged as approximately parallel sheets, progressing toward the arrangement in graphite.

ISOTOPE COMPOSITION OF Hydrocarbons

The isotopic composition of sedimentary organic matter and its derivatives, such as coal and oil, depend on (1) the isotopic composition of the originally deposited organic matter and (2) isotopic fractionations occurring during diagenesis and subsequent thermal evolution. The ultimate source of carbon in sedimentary organic matter is atmospheric CO₂ or marine HCO₃. The $\delta^{13}C_{PDB}$ value of the former is about -7‰, while the $\delta^{13}C_{PDB}$ in average surface ocean water is +2.2‰ (both of which vary somewhat). As we found in Chapter 9, isotopic fractionation during photosynthesis results in organic carbon being substantially lighter (lower $\delta^{13}C$) that either atmospheric or dissolved CO₂. Terrestrial C₃ plants typically have $\delta^{13}C$ of -25 to -30‰, C₄ plants have $\delta^{13}C$ of -10 to -15‰, while marine plants are somewhat more variable in isotopic composition (-5 to -30‰), though most have isotopic compositions of -20 to -28‰, i.e., slightly heavier than terrestrial C₃ plants. There is some further frac-

Geochemistry

W. M. White

CHAPTER 14: ORGANIC GEOCHEMISTRY

tionation of carbon isotopes as other organic molecules are synthesized, with lipids being isotopically lighter than carbohydrates and proteins. While these differences are small compared with the fractionation during photosynthesis, they do appear to persist through diagenesis.

Most living organisms have δD_{SMOW} in the range of -60 to -150‰. Within this range, hydrogen isotope ratios vary due to hydrogen isotope fractionation in the hydrologic cycle (Chapter 9). Terrestrial plants tend to be more deuterium depleted than marine ones, and terrestrial plants from cold climates are particularly depleted. Lipids are depleted in δD relative to bulk organic matter by 60‰ or more. Most kerogen, coal, and oil show about the same range in δD as do organisms. As might be expected, however, lipid-rich kerogen and oil can have substantially lower δD .

The situation with nitrogen isotopes is similar: nitrogen isotope ratios in sedimentary organic matter generally reflect that of the biomass from which it is derived, with terrestrial plants having slightly lower average $\delta^{15}N_{ATM}$ than marine plankton (Fogel and Cifuentes, 1993). $\delta^{15}N$ generally decreases somewhat during diagenesis due to bacterial utilization of short chain peptides following peptide bond hydrolysis (Macko, et al., 1993).

For the most part, isotopic fractionation of carbon during diagenesis of organic matter is small. As a result, the δ^{13} C of sedimentary organic matter is typically within a few permil of the δ^{13} C of the biomass from which it is derived. Sedimentary organic matter and humic substances in soil and water tend to be slightly more depleted (by 2 to 3 per mil) in ¹³C than



Figure 14.40. Isotopic composition of methane from various sources. "Biogenic" methane is methane produced by methanogens during diagenesis, "Oil Associated" is methane associated with oil, "C" indicates typical composition of methane in gas condensates, "Metagenic" is methane produced during metagenesis, "A" is the composition of abiogenic methane from mid-ocean ridge hydrothermal systems. Modified from Schoell (1984).

the organisms from which they are derived, though cases where the opposite is true have been observed. There are several possible causes for this (reviewed in Tissot and Welte, 1984, Hoefs, 1987, and Macko et al., 1993). First, functional groups, such as carboxyl, tend to be relatively ¹³C rich. Loss of functional groups during diagenesis or condensation of humic substances, will drive the residual organic carbon to lower δ^{13} C. Second, the appears to be a kinetic fractionation involved in condensation of humic substances and kerogen-like molecules. Third, preferential remineralization of proteins and carbohydrates leaves a lipid-rich residue, which will be isotopically light.

Fractionation of carbon isotope ratios during thermal evolution through the oil generation stage is small. In immature kerogen, bitumens are depleted in δ^{13} C compared to kerogen, but this difference decreases with increasing maturity (Schoell, 1984). Sofer (1984) found that δ^{13} C of the oils are within 2‰ of the isotopic composition its source kerogen. Conkright and Sackett (1992) found that δ^{13} C of organic carbon in sediment cores from DSDP (Deep Sea Drilling Project) Site 368 near the Canary Islands decreased with proximity to an intrusive diabase sill by 2 to 3‰. They attributed the decrease to thermal maturation and loss of isotopically light methane caused by heating from the sill.

Because the isotopic composition of oil is similar to that of its parent kerogen, isotope ratios are a widely used exploration tool the petroleum industry. Isotope ratios of bulk organic matter, however, cannot be used to discriminate between depositional environments as the isotopic differences between

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

marine and terrestrial organic matter are neither sufficiently large nor sufficiently systematic. Interestingly, δD in oil derived from shale source rocks are generally lower than those derived from carbonate source rocks (Schoell, 1984). The cause of this difference is unclear.

There can be significant fractionation of carbon and hydrogen isotopes in the generation of methane. As Figure 14.40 shows, methane produced by methanogenic bacteria, called "biogenic methane" during diagenesis is highly depleted in ¹³C. Methane produced during catagenic is depleted in both ¹³C and deuterium compared to associated oil and kerogen. These fractionations reflect the lower strength of ¹²C–¹²C bonds compared to ¹³C–¹²C bonds and therefore the greater ease with which the former are broken. As the metagenesis stage is entered, however, the isotope fractionation between methane and residual kerogen decreases and the isotopic composition methane generated during this stage approaches that of kerogen. This is just what we would expect from both the inverse relationship between the fractionation factor and temperature and the decreasing fractionation as reactions proceed to completion (see Chapter 9).

Methane in mid-ocean ridge hydrothermal vent fluids has systematically higher δ^{13} C (-15 to - 20‰) than biogenic and thermogenic methane derived from sedimentary organic matter (Schanks et al., 1995), demonstrating its abiogenic origin. The isotopic fractionation between methane and CO₂ in these fluids suggests equilibration at temperatures in the range of 600 to 800° C (Whelan and Craig, 1983).

References and Suggestions for Further Reading

- Aiken, G. R., D. M. McKnight, R. L. Wershaw and P. MacCarthy (eds.). 1985. *Humic Substances in Soil, Sediment, and Water*. New York: Wiley Interscience. 692 p.
- Bartschat, B. M., S. E. Cabaniss and F. M. M. Morel. 1992. Oligoelectrolyte model for cation binding by humic substances. *Environ. Sci. Techol.* 26: 284-294.
- Bennett, P. C. and W. Casey.1994. Chemistry and mechanisms of low-temperature dissolution of silicates by organic acids. In *Organic Acids in Geological Processes*, ed. E. D. Pitman and M. D. Lewan. 162-200. Berlin: Springer Verlag.

Bergmann, W. 1988. Zur Struckturaufklärung von Huminsäuren aus Abwasser, PhD Thesis, 136 pp.

Berner, R. A. 1981. Early Diagenesis, A Theoretical Approach. Princeton: Princeton Univ. Press.

- Buffle, J., P. Deladoey, F. L. Greter and W. Haerdi. 1980. Study of the complex formation of copper (II) by humic and fulvic substances. *Anal. Chim. Acta*. 116: 255-274.
- Cabaniss, S. E. and M. S. Shuman. 1988. Copper binding by dissolved organic matter: II. Variation in type and source of organic matter. *Geochim. Cosmochim. Acta*. 52: 195-200.
- Calvert, S. E. and T. F. Petersen. 1992. Organic carbon accumulation and preservation: how important is anoxia. In Organic Matter: Productivity, Accumulation, and Preservation in Recent and Ancient Sediments, Vol. J. K. Whelan and J. W. Farmington. ed., pp. 231-263. New York: Columbia University Press.

Chiou, C. T., L. J. Peters and V. H. Freed. 1979. A physical concept of soil-water equilibria for nonionic organic compounds. *Science*. 206: 831-832.

- Conkright, M. T. and W. M. Sackett. 1992. Stable carbon isotope changes during maturation of organic matter. In Organic Matter: Productivity, Accumulation, and Preservation in Recent and Ancient Sediments, Vol. J. K. Whelan and J. W. Farmington. ed., pp. 403-414. New York: Columbia University Press.
- de Leeuw, J. W. and C. Largeau. 1993. A review of the macromolecular organic compounds that comprise living organisms and their role in kerogen, coal, and petroleum formation. In *Organic Geochemistry: Principles and Applications, Vol.* M. H. Engel and S. A. Macko. ed., pp. 23-72. New York: Plenum.
- Deming, J. W. and J. A. Baross. 1993. The early diagenesis of organic matter: bacterial activity. In *Organic Geochemistry: Principles and Applications, Vol.* M. H. Engel and S. A. Macko. ed., pp. 119-144. New York: Plenum.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

- Donat, J. R. and K. W. Bruland. 1993. A comparison of two voltametric techniques for determining zinc speciation in northeast Pacific Ocean water. *Mar. Chem.* 28: 301-323.
- Drever, J. I. and G. F. Vance.1994. Role of soil organic acids in mineral weathering processes. in *Or*ganic Acids in Geological Processes, ed. E. D. Pitman and M. D. Lewan. 138-161. Berlin: Springer Verlag.
- Engel, M. and S. A. Macko, *Organic Geochemistry: Principles and Applications*, Plenum Press, New York, 861 p, 1993.
- Fogel, M. L. and M. L. Cifuentes.1993. Isotope Fractionation during Primary Production. in *Organic Geochemistry: Principles and Applications,* ed. M. H. Engel and S. A. Macko. 73-98. New York: Plenum.
- Furrer, G. and W. Stumm. 1986. The coordination chemistry of weathering I. Dissolution kinetics of δ -Al₂O₃ and BeO. *Geochim. Cosmochim. Acta.* 50: 1847-1860.
- Graedel, T. E. and P. J. Crutzen. 1993. *Atmospheric Change: An Earth System Perspective*. New York: W. H. Freeman.
- Grauer, R. and W. Stumm. 1982. Die Koordinationschemie oxidisher Grenzflächen und ihre Auswirkung auf die Auflösingskinetic oxidisher Festphasen in wässrigen Lösungen. *Colloid. Polymer Sci.* 260: 959-970.
- Harvey, G. R. and D. Boran. 1985. Geochemistry of humic substances in seawater, in *Humic Substances in Soil, Sediment, and Water*, edited by G. R. Aiken, D. M. McKnight, R. L. Wershaw and P. Mac-Carthy, p. 233-247, New York: Wiley Interscience.
- Henrichs, S. M.1993. Early diagenesis of organic matter: the dynamics (rates) of cycling of organic compounds. In Organic Geochemistry: Principles and Applications, Vol. M. H. Engel and S. A. Macko. ed., pp. 101-117. New York: Plenum.
- Hoefs, J. 1987. Stable Isotope Geochemistry. Berlin: Springer Verlag.
- Killops, S. D. and V. J. Killops. 1993. Organic Geochemistry,, Essex: Longman Scientific and Technical. 265 p.
- Macko, S. A., E. M. H. and P. L. Parker. 1993. Early diagenesis of organic matter in sediments: assessment of mechanisms and preservation by the yse of isotopic molecular approaches. In *Organic Geochemistry: Principles and Applications, Vol.* M. H. Engel and S. A. Macko. ed., pp. 211-223. New York: Plenum Press.
- Malcolm, R. L., Geochemistry of stream fulvic and humic substances, in *Humic Substances in Soil*, *Sediment, and Water*, edited by G. R. Aiken, D. M. McKnight, R. L. Wershaw and P. MacCarthy, p. 181-209, Wiley Interscience, New York, 1985.
- Mason, S. F., Chemical Evolution: Origin of the Elements, Molecules, and Living Systems, Oxford University Press, Oxford, 317 p, 1991.
- Mayer, L. W. 1993. Organic matter at the sediment—water interface. In Organic Geochemistry: Principles and Applications, Vol. M. H. Engel and S. A. Macko. ed., pp. 171-184. New York: Plenum.
- McKnight, D. M., W. E. Pereira, M. L. Ceazan and R. C. Wissmar. 1982. Characterization of dissolved organic materials in surface waters within the blast zone of Mount St. Helens, Washington. *Organic Geochem.* 4: 85-92.

Morel, F. M. M. 1983. Principles of Aquatic Chemistry. New York: Wiley Interscience.

- Morel, F. M. M. and J. G. Hering. 1993. *Principles and Applications of Aquatic Chemistry*. New York: John Wiley and Sons.
- Nürnberg, H. W. and P. Valenta.1983. Potentialities and applications of voltammetry in chemical speciation of trace metals in the sea. in *Trace Metals in Sea Water*, ed. C. S. Wong, E. Boyle, K. W. Bruland, J. D. Burton and E. D. Goldberg. 671-697. New York: Plenum Press.
- Purves, W. K., G. H. Orians and H. C. Heller. 1992. *Life: The Science of Biology*. Sunderland, MA: Sinauer Associates.
- Pitman, E. D. and M. D. Lewan (ed.). 1994. Organic Acids in Geological Processes, Berlin: Springer Verlag.
- Royal Dutch Shell. 1983. The Petroleum Handbook. Amsterdam: Elsevier.

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

- Rullkötter, J. 1993. The thermal alteration of kerogen and the formation of oil. In *Organic Geochemistry: Principles and Applications, Vol.* M. H. Engel and S. A. Macko. ed., pp. 101-117. New York: Plenum.
- Schoell, M. 1984. Stable isotopes in petroleum research. In *Advances in Petroleum Geochemistry*, 1, ed. J. Brooks and D. Welte. 215-245. London: Academic Press.
- Schwartzenbach, R. P. and J. Westall. 1981. Transport of nonpolar organic compounds from sufrace waters to groundwater. Laboratory sorption studies. *Environ. Sci. Tech.* 15: 1360-1367.
- Schanks, W. C., J. K. Böhlke and R. R. Seal.1995. Stable isotopes in mid-ocean ridge hydrothermal systems: interactions between fluids, minerals, and organisms. In *Seafloor Hydrothermal Systems*, *Geophysical Monograph Vol.* 91, S. E. Humphris, R. A. Zierenberg, L. S. Mullineaux and R. E. Thomson. ed., pp. 194-221. Washington: AGU.
- Sofer, Z. 1984. Stable carbon isotope compositions of crude oils: application to source depositional environments and petroleum alteration. *Am. Assoc. Petrol. Geol. Bull.* 68: 31-49.
- Sposito, G., The Chemistry of Soils, Oxford University Press, New York, 277 p., 1989.
- Stevenson, F. J. 1982. Humus Chemistry. New York: John Wiley and Sons.
- Stevenson, F. J. and G. F. Vance. 1989. Naturally occurring aluminum-organic complexes. In *The environmental Chemistry of Aluminum, Vol.* G. Sposito. ed., pp. 117-145. Boca Raton, FL: CRC Press.
- Steinberg, C. and U. Muenster, Geochemistry and ecological role of humic substances in lake water, in *Humic Substances in Soil, Sediment, and Water*, edited by G. R. Aiken, D. M. McKnight, R. L. Wershaw and P. MacCarthy, pp. 105-145, Wiley Interscience, New York, 1985.
- Stumm, W. 1992. Chemistry of the Solid-Water Interface. New Yokr: Wiley Interscience.
- Stumm, W. and J. J. Morgan. 1996. Aquatic Chemistry. New York: Wiley Interscience.
- Surdam, R. C., S. W. Boese and L. J. Crossey. 1984. The geochemistry of secondary porosity. in *Clastic diagenesis, Am. Assoc. Petrol. Geol. Mem* 37, ed. R. A. McDonald and R. C. Surdam. 127-149. Tulsa: Am. Assoc. Petrol. Geol.
- Tegelaar, E. W., S. Derenne, C. Largeau and J. W. de Leeuw. 1989. A reappraisal of kerogen formation. *Geochim. Acta.* 53: 3103-3107.
- Thurman, E. M., Organic Geochemistry of Natural Waters, Martinus Nijhoff/Dr W. Junk Publishers, Dordrecht, 497 p., 1985.
- Tissot, B. P. and D. H. Welte. 1984. Petroleum Formation and Occurrence. Berlin: Springer Verlag.
- Welhan, J. A. and H. Craig. 1983. Methane, hydrogen and helium in hydrothermal fluids at 21°N on the East Pacific Rise. In *Hydrothermal Processes at Seafloor Spreading Centers, Vol.* P. Rona, K. Boström, L. Laubier and K. L. Smith. ed., pp. 391-409. New York: Plenum.
- Whelan, J. K. and J. W. Farrington (ed.). 1992. Organic Matter: Productivity, Accumulation, and Preservation in Recent and Ancient Sediments. New York: Columbia University Press.
- Whelan, J. K. and C. L. Thompson-Rizer. 1993. Chemical methods for assessing kerogen and protokerogen maturity. In Organic Geochemistry: Principles and Applications, Vol. M. H. Engel and S. A. Macko. ed., pp. 101-117. New York: Plenum.
- Zinder, B., G. Furrer and W. Stumm. 1986. The coordination chemistry of weathering: II Dissolution of Fe(III) oxides. *Geochim. Cosmochim. Acta*. 50: 1861-1869.
- Zuehlke, R. W. and D. R. Kester. 1983. Copper speciation in marine waters. in *Trace Metals in Sea Water*, ed. C. S. Wong, E. Boyle, K. W. Bruland, J. D. Burton and E. D. Goldberg. 773-788. New York: Plenum Press.

Problems

- 1. Sketch the structure of the following:
 - a.) Citric acid: HOC(CH₂CO₂H)₂CO₂H
 - b.) Tartaric acid: HO₂CCH(OH)CH(OH)CO₂H (2,3,-dihydroxybutanedioic acid).
- 2. Write the chemical formula and sketch the structure of 2-hydroxy-propanoic acid (lactic acid).

Geochemistry

CHAPTER 14: ORGANIC GEOCHEMISTRY

3. Suppose you could follow the pathway of individual atoms during photosynthesis. While this is not possible, something similar can be done by isotopic labelling of water and CO_2 . If ¹⁸O-labelled water is added to a suspension of photosynthesizing chloroplasts, which of the following compounds will first show enrichment in ¹⁸O: ATP, NADPH, O_2 , or 3-phosphoglycerate? If you repeat the experiment with ²H-labelled water and ¹³C-labelled CO_2 , which of these molecules will first show enrichment in these isotopes?

4. The the first and second acidity constants of oxalic acid $((COOH)_2)$ are $pK_{a1} = 1.23$ and $pK_{a2} = 4.19$. What is the pH of a solution formed by dissolving 1 mole of oxalic acid in 1 kg of water? If this solution is titrated with 1 M NaOH, how will pH change as a function of the amount of base added? Make a plot of pH vs. amount of base added.

5. Assume that humic acid may be modelled as consisting of two ligands: a bidentate carboxylic ligand (e.g., malonate) and a bidentate phenol one (e.g., catechol), and that the effective concentrations of these are 10^{-3} mol/g humate and 5×10^{-4} mol/g humate respectively. Using the following apparent stability constants, calculate the fraction of copper complexed if the humate concentration is 10 mg/l, the pH 8, and the total copper concentration is 10^{-8} M. Assume that copper and humate are the only species present.

Apparent stability constants:

"Malonate":	H_2L	$pK_{a1} = 1.9$
	HL-	$pK_{a2} = 1.9$
	CuL	$pK_{Cu} = 7.7$
"Catechol":	H_2L	$pK_{a1} = 6.8$
	HL-	$pK_{a2} = 10.0$
	CuL	$pK_{Cu} = 17.3$

6. Repeat the calculation in Problem 5 above, but for pH 5.5. At this pH, the apparent stability constants are:

"Malonate":	H_2L	$pK_{a1} = 2.9$
	HL-	$pK_{a2} = 2.9$
	CuL	$pK_{Cu} = 6.7$
"Catechol":	H_2L	$pK_{a1} = 7.8$
	HL-	$pK_{a2} = 11.0$
	CuL	$pK_{Cu} = 16.3$
		-

7. The adjacent table lists organic solid/water (K_{OM}) and octanol/water (K_{OC}) partition coefficients for some non-polar compounds. Are these data consistent with equation 14.15? What values do you determine for constants *a* and *b*? (*HINT*: Use linear regression.)

8. Sediment from a highly eutropic lake was p found to have an organic carbon fraction of 5.8%. d Using the adsorption partition coefficient for DDT listed in problem 7, predict the concentration of DDT in the sediment if the lake water has a

	K _{OW}	K _{OM}
acetophenone	38.90	42.66
benzene	128.82	83.18
tetrachloroethylene	398.11	208.93
napthalene	2290.87	1288.25
parathion	6456.54	1148.15
pyrene	151356.12	83176.38
chlorobenzene	512.86	389.05
DDT	1548816.62	138038.43
2,4,5,2',4',5'-PCB	5248074.60	218776.16

DDT concentration of 3 μ g/l and the sediment contains 5.8% organic matter.

CHAPTER 14: ORGANIC GEOCHEMISTRY

9. The rate of bond cleavage during the thermal maturation of kerogen approximately doubles for every 10° rise in temperature. Based on this and assuming that these reaction rates show an Arrhenius temperature dependence (equation 5.24), estimate the activation energy for these reactions.

10. Astrophysicist Thomas Gold has suggested that most petroleum deposits are formed by abiogenic organic carbon (mainly in the form of methane) diffusing out of the mantle. There are few, if any, geochemists that agree. Describe at least 3 *geochemical* observations that support the "conventional" theory that petroleum is formed from sedimentary kerogen, which in turn is derived from the remains of once living organisms.

CHAPTER 15: OCEANS

CHAPTER 15: THE OCEANS AS A CHEMICAL SYSTEM

INTRODUCTION

ntoine Lavoisier* called seawater "the rinsings of the Earth." Given the tenuous understandntoine Lavoisier cancer scanned set the time (the late 18th century), uns is a terminal insightful observation. Most of the salts in the oceans are derived from weathering of the continents and delivered to the oceans by rivers. But the story of seawater is more complex than this. Some components of seawater are derived from hydrothermal metamorphism of the oceanic crust. Other components in seawater, most notably the principal anions as well as water itself, are derived from neither weathering nor hydrothermal reactions. These so-called "excess volatiles" are derived from volcanic degassing. Furthermore, salts do not simply accumulate in seawater. This point was overlooked by John Jolly in his attempt to estimate the age of the Earth, described in Chapter 8, from the mass of salts in the sea and the amount added annually by rivers. His result, 90 million years, is a factor of 50 less than the actual age of the Earth. The ocean is a dynamic, open system, and it is ultimately the balance between addition and removal of an element that dictates it concentration in the ocean. This was recognized by Georg Forschhammer in 1865 when he wrote: "The quantity of different elements in seawater is not proportional to the quantity of elements which river water pours into the sea, but is inversely proportional to the facility with which the elements are made insoluble by general chemical or organo-chemical actions in the sea." One of our objectives in this chapter will be to examine the budget of dissolved matter in the oceans; that is, to determine the sources and sinks and the rates at which salts are added and removed from the oceans.

Elements also cycle between different forms within the oceans; these include both organic and inorganic solids as well as various dissolved species. This internal cycling is intimately tied to the various physical, geological, and biological processes occurring within the ocean. The biota plays a particularly crucial role both in internal chemical cycling and in controlling the overall composition of seawater. A second objective of this chapter will be to examine how elements and compounds are distributed within the ocean, and how they cycle between various forms.

Lavoisier's statement also reminds us that the oceans are part of a grander geochemical system. Sediments deposited in the ocean provide a record of that system. On human time scales at least, the ocean appears to be very nearly in steady state. It is tempting to apply Lyell's principal of uniformitarianism and assume that the composition of the seawater has also been constant on geologic time scales. There is, however, strong evidence that some aspects of seawater composition do change over time, as we found in Chapters 8 and 9. Precisely because these variations are related to changes in other geological processes, such as plate tectonics, climate, life, and atmospheric chemistry, they can tell us much about the Earth's history and the workings of the planet. Interpreting these past changes begins with an understanding of how the modern ocean works and the controls on its composition. This understanding is our main goal for this chapter.

Some Background Oceanographic Concepts

Salinity, Chlorinity, Density, and Temperature

A useful concept in oceanography is *salinity*. Salinity can be thought of as the total dissolved solids in seawater. More precisely, salinity is defined as: *the weight in grams of the dissolved inorganic matter in one kilogram of water after all the bromide and iodide have been replaced by the equivalent amount of chloride and all carbonate converted to oxide* (CO_2 driven off). This unfortunate defi-

^{*}Antoine Lavoisier, born in France in 1843, is often called the father of modern chemistry. He died at the guillotine in 1794.

Geochemistry

CHAPTER 15: OCEANS

nition has an interesting historical basis. Robert Boyle⁺ found that he could not reproduce total dissolved solid measurements by drying and weighing. The salinity definition is due to Thorensen, who would bubble Cl_2 gas, which substitutes for Br and I, through seawater. The salt could then be heated, converting carbonate to oxide, and a constant weight achieved. Salinity is now determined by measuring electrical conductivity, which increases in direct relation to the concentrations of ions in water, and hence with salinity. Another useful definition is *chlorinity*, which is *the halide concentration in grams per kilogram measured by titration with silver and calculated as if all the halide were chloride* (total halides are actually 0.043% greater than chlorinity). Chlorinity can also be measured by conductivity. As we shall see, Cl is always present in seawater as a constant proportion of total salt, and therefore there is a direct relationship between chlorinity and salinity. By definition:

$S_{\infty} = 1.80655 \text{ Cl}_{\infty}$ 15.1

"Standard seawater", which is close to average seawater, has a salinity of 35.000 parts per thousand (ppt or ‰) and a chlorinity of 19.374 ‰. Open ocean water rarely has a salinity greater than 38‰ or less than 33‰.

Temperature, along with salinity, determines the density of seawater. Since density differences drive much of the flow of ocean water, these are key oceanographic parameters. Temperature in the oceans can be reported as *potential*, or *in situ* temperature, but the former is the most commonly used. In situ temperature is the actual temperature of a parcel of water at depth. Potential temperature, denoted θ , is the temperature the water would have if brought to the surface. The difference between the two is thus the temperature difference due to adiabatic expansion. Since water cools when it expands, potential temperature is always less than *in situ* temperature (except for surface water, where there is not difference). The difference is small, on the order of 0.1°C. While this difference is important to oceanographers, it is generally negligible for our purposes. Temperature and salinity, and therefore also density, are conservative properties of seawater, which is to say that they can be changed only at the surface.

The density of seawater is 2 to 3 percent greater than that of pure water. Average seawater, with a salinity of 35% and a temperature of 20°C, has a density of 1.0247 g/cc. Density is usually reported as the parameter σ , which is the per mil deviation from the density of pure water (1 g/cc). Thus if density is 1.0247 g/cc, σ is 24.7. Again, one can distinguish between *in situ* and potential density, potential density being the density water would have if brought to the surface, and is always lower than *in situ* density. The difference is small, a few percent, and generally negligible for our purposes.

Circulation of the Ocean and the Structure of Ocean Water

The concentrations of dissolved elements vary both vertically and horizontally in the ocean. To fully understand these variations, we need to know something about the circulation of the ocean. This circulation, like that of the atmosphere, is ultimately driven by differential heating of the Earth: solar energy is gained principally at low latitudes and lost at high latitudes. Because the mechanisms of surface and vertical circulation in the oceans are somewhat different, it is convenient to treat them separately.

Surface Circulation

Surface circulation of the ocean is driven primarily by winds; hence the surface circulation is sometimes also called the *wind-driven circulation*. Figure 15.1 is a simplified map of the wind-driven circulation. The important features are as follows:

• Both north and south of the climatic equator, known as the Inter-Tropical Convergence, or ITC, water moves from east to west, driven by the Trade Winds. These currents are known as the North

⁺ Robert Boyle (1627-91) was another of the founders of modern chemistry. He defined the chemical element as the practical limit of chemical analysis, and deduced the inverse relationship between the pressure and volume of gas, a version of the ideal gas law.

CHAPTER 15: OCEANS



Figure 15.1. Surface and Deep Circulation of the oceans. Arrows show the direction of wind driven currents. Gray areas are regions of upwelling. Red stippled areas are regions of deep water production. In the Indian Ocean, black arrows show current directions in northern hemisphere winter, red arrows show current direction in northern hemisphere summer.

and South Equatorial Currents. Between these two currents, the Equatorial Counter Current runs from west to east.

• Two large gyres operate in both the Atlantic and Pacific Oceans, one in the northern and one in the southern hemisphere. Rotation is clockwise in the northern hemisphere and counter-clockwise in the southern hemisphere. The Coriolis Force, an apparent force that results from the Earth's rotation, is largely responsible for this circular current pattern. These currents are most intense in along the western boundaries of ocean basins, a phenomenon, also due to the Earth's rotation. Examples of intense western boundary currents are the Gulf Stream and Kuroshio Current.

• The circulation in the Indian Ocean is similar, but undergoes radical seasonal changes in response to the Monsoons. In northern hemisphere summer, the North Equatorial Current reverses and joins the equatorial countercurrent to become the Southwest Monsoon Current. The Somali Current, which flows to the southwest along the African Coast in northern hemisphere winter, reverses direction to flow northeastward in northern hemisphere summer.

• Water moves from west to east in Southern Ocean (the globe-encircling belt of ocean south of Africa and So. America). This is called the Antarctic Circumpolar Current or West Wind Drift. Directly adjacent the Antarctic coast, a counter current, called the Polar Current, runs east to west.

Density Structure and Deep Circulation

The deep circulation of the oceans is driven by density differences. Seawater density is controlled by temperature and salinity, so this circulation is also called the *thermohaline circulation*. Most of the ocean is stably stratified; that is, each layer of water is denser than the layer above and more dense than the layer below. Where this is not the case, a water mass will move up or down until it reaches a level of equilibrium density. *Upwelling* of deeper water typically occurs where winds or currents create a *divergence* of surface water. *Downwelling* occurs where winds or currents produce a *convergence* of surface water. Wind and current-driven upwelling and downwelling link the surface and deep circulation of the ocean.

Geochemistry

CHAPTER 15: OCEANS

In the modern ocean, temperature differences dominate density variations and are principally responsible for deep circulation. This may not have always been the case, however. During warmer periods in the past, such as the Cretaceous, deep circulation may have been driven principally by salinity differences.

Figure 15.2 shows an example of how temperature, salinity, and density vary with depth in temperate and tropical regions. Both are usually nearly constant in the upper hundred meters or so as a result of mixing by waves (the actual depth of the mixed layer varied both seasonally and geographically, depend largely on wave height). Below the upper mixed layer is a region, called the *thermocline*, where temperature decreases rap-



Figure 15.2. Temperature, salinity, and density variations at GEOSECS station 25 at 58° N in the North Atlantic. The station was occupied in September, and summer heating has extended the thermocline and pycnocline nearly to the surface. Gray area shows the position of the permanent thermocline and pycnocline. An inversion in the salinity profile near the surface indicates an excess of precipitation over evaporation.

idly. Salinity may also change rapidly in this region; a region where salinity changes rapidly is called a *halocline*. The temperature changes cause a rapid increase in density with depth, and this region of the water column is called the *pycnoline*. Below the pycnoline, temperature and salinity vary less with depth. Temperature and salinity are *conservative* properties of a water mass, which is to say that they can only be changed at the surface. Hence within the deep ocean, temperature and salinity vary only because of mixing of different water masses. In polar regions, water may be essentially isothermal throughout the water column.



Figure 15.3. The two box model of the ocean and the fluxes between them.

The pycnocline represents a strong boundary to vertical mixing of water and effectively isolates surface water from deep water. This leads to a sometimes useful chemical simplification: the two-box model of the ocean. In this model, the ocean is divided into a box representing the surface water above the pycnoline, and one representing the deep water below it (Figure 15.3). Fluxes between these boxes can occur both because of advection of water (upwelling and downwelling) and because of falling particles, both organic and inorganic. The upper box exchanges with the atmosphere and receives all the riverine input. All photosynthetic activity occurs in the upper box because light effectively does not penetrate below 100m (only 0.5% of the incident sunlight penetrates to a depth of 100 m, even in the clearest water). On the other hand, the flux out of the ocean of both particles and dissolved solids occurs through the lower box. Since the depth of the surface layer varies in the ocean and the density boundary is gradational rather than sharp, any definition of the size of the boxes is rather arbitrary. The depth of the boundary between the surface and deep layer may be variably defined, depending on the particular problem at hand. In Example 15.1, for
Geochemistry

CHAPTER 15: OCEANS

instance, we define it as 1000 m.

Water flows across the pycnocline only a few limited regions; we can divide these into regions of "intermediate water" formation and "deep water" formation (formation refers to a water mass acquiring its temperature and salinity characteristics at the surface and sinking through the pycnocline). Intermediate waters do not usually penetrate below depths of 1500 m; deep water may pene-

Example 15.1. Replacement time of Deep Ocean Water

Use the simple two-box model in Figure 15.4 together with the following to estimate the residence time of water in the deep ocean. Take the boundary between the surface and deep water to be 1000m. Assume the system is at steady state and that $^{14}C/C$ ratio in deep water is 10% lower than in of surface water (Stuvier, et al., 1983).

Answer: Since we can assume the system is at steady state, the upward and downward fluxes of both water and carbon must be equal. Denoting the water flux by F_W we may write the mass balance equation for carbon in the deep water reservoir as:

$$F_w C_p = F_w C_s + F_{CP}$$
 15.2

where C_D , C_S , and C_P are the concentrations of carbon in deep waterand shallow water respectively, and F_{CP} is the flux of carbon carried by sinking particles. The sinking particle flux in thus just:

$$F_{\rm CP} = F_{\rm W}(C_{\rm D} - C_{\rm S})$$
 15.3

In other words, the sinking particle flux must account for the difference in carbon concentration between the surface and deep water.

We may now write a mass balance equation for ¹⁴C in deep water by setting the loss of ¹⁴C equal to the gain of ¹⁴C. ¹⁴C is lost through the upward flux of water and radioactive decay and gained by the downward flux of water and the sinking particle flux.

$$F_{W}C_{D}(^{14}C/C)_{D} - \lambda V_{D}C_{D}(^{14}C/C)_{D} = F_{W}C_{S}(^{14}C/C)_{S} + F_{CP}(^{14}C/C)_{S}$$
 15.4

where $({}^{14}C/C)_D$, and $({}^{14}C/C)_S$ are the ${}^{14}C/C$ ratios in deep and shallow water respectively, V_D is the volume of deep water, and λ is the decay constant of ${}^{14}C$. We have implicitly assumed that sinking particles have the same ${}^{14}C$ activity as surface water. Subsituting 15.3 into 15.4, we have

$$F_{\rm w}C_{\rm D}({}^{14}{\rm C/C})_{\rm D} + \lambda V_{\rm D}C_{\rm D}({}^{14}{\rm C/C})_{\rm D} = F_{\rm w}C_{\rm S}({}^{14}{\rm C/C})_{\rm S} + F_{\rm w}(C_{\rm D} - C_{\rm S})({}^{14}{\rm C/C})_{\rm S}$$
 15.5

Rearranging and eliminating terms we have

$$\lambda V_{\rm D} ({}^{14}{\rm C/C})_{\rm D} = F_{\rm W} ({}^{14}{\rm C/C})_{\rm S} - F_{\rm W} ({}^{14}{\rm C/C})_{\rm D}$$
 15.6

Another rearrangement and we arrive at:

$$V_{\rm D}/F_{\rm W} = [1 - ({}^{14}{\rm C/C})_{\rm D}/({}^{14}{\rm C/C})_{\rm S}]/[\lambda ({}^{14}{\rm C/C})_{\rm D}/({}^{14}{\rm C/C})_{\rm S}]$$
 15.7

As we shall see later in this chapter, we define steady state residence time as the amount in a reservoir divided by the flux into it or out of it. Thus the above equation gives the residence time of water in the deep ocean (notice it has units of time). Substituting 0.1209×10^{-3} yr⁻¹ for λ (Table 8.5) and 0.9 for ($^{14}C/C$)_D/($^{14}C/C$)_s, we calculate a residence time of 920 years. This is somewhat longer than the residence time arrived at by Stuvier et al. (1983) through a more sophisticated analysis.

We can also use this equation to calculate the average upward velocity of water. Rearranging 15.7, we have:

$$F_{\rm W} = V_{\rm D} \lambda ({}^{14}{\rm C/C})_{\rm D} / ({}^{14}{\rm C/C})_{\rm S} / [1 - ({}^{14}{\rm C/C})_{\rm D} / ({}^{14}{\rm C/C})_{\rm S}]$$
 15.8

If we express the volume of the deep ocean as the average depth, d, times area, A, we have:

$$F_{\rm W} = {\rm Ad}\lambda ({\rm ^{14}C/C})_{\rm D} / ({\rm ^{14}C/C})_{\rm S} / [1 - ({\rm ^{14}C/C})_{\rm D} / ({\rm ^{14}C/C})_{\rm S}]$$
 15.9

Dividing both sides by A, we have:

$$F_{W}/A = d\lambda ({}^{14}C/C)_{D}/({}^{14}C/C)_{S}/[1 - ({}^{14}C/C)_{D}/({}^{14}C/C)_{S}]$$
 15.10

 F_W/A is the velocity. Taking d as 3000 m, we calculate F_W/A as 3.26 m/yr. (This calculation follows a similar one in Broeker and Peng, 1983).

Geochemistry

W. M. White

CHAPTER 15: OCEANS

trate to the bottom. There are four principal regions of intermediate water production. The first is in the Mediterranean where evaporation increases salinity of surface water to 37-38‰. Winter cooling further increases density causes this surface water to sink. It flows out of the Strait of Gibraltar and sinks in the Atlantic to a depth of about 1000 m where This water is known as it spreads out. Mediterranean Intermediate Water (MIW). Another intermediate water, known as Antarctic Intermediate Water, or AAIW, is produced at the convergence at the Antarctic Polar Front at about 50°S. North Pacific Intermediate Water is produced at the convergence Arctic



Figure 15.4. Simple two-dimensional box model of ocean circulation. The volumes of each reservoir are not given in units of 10^{18} m³ (after Stuvier et al., 1983). In this case, the boundary between surface water and deep water is taken as 1500 m.

Polar Front at about 50° N in the Pacific. North Atlantic Intermediate Water is also produced at the Arctic Polar Front at 50° to 60° N. Of these water masses, Antarctic Intermediate Water is the densest and most voluminous.

There are only two regions of deep water production, both at high latitudes. Antarctic Bottom Water (AABW), which is the densest and most voluminous deep water in the ocean, is produced primarily in the Weddell Sea. Cold winds blowing from Antarctica cool it, while freezing of sea ice increases its salinity. The other deep water mass, North Atlantic Bottom Water (NADW), is produced around Iceland in winter when winds cause upwelling and cooling of saline MIW. NADW then sinks and flows southward along the western boundary of the Atlantic. In the Southern Ocean it mixes with and becomes part of the AABW.

Mixing between deep water and water results in a slow, diffuse upward advection through the deep layer and then into the thermocline. Thus whereas the flux from the surface layer to the deep one is focused, the upward flux is diffuse. Final return from the thermocline to the surface occurs in localized zones of upwelling. The principal upwelling zones are those along the equator, where the trade winds create a divergence of surface water, along the west coasts of continents, where winds blowing along the coast drive the water offshore (this is a process known as Ekman transport and is related to the Coriolis force), and at the Antarctic divergence in the Southern Ocean.

With our knowledge of deep water circulation, we can extend our one dimensional model (Figure 15.3) to two dimensions (Figure 15.4). The model illustrates several important features of the deep circulation of the oceans. First, no deep water is produced in either the Pacific or Indian Oceans. Second, the Atlantic exports deep water and imports surface water. Both the Indian and Pacific import deep water and export surface water. Third, all exchanges of deep water take place via the Southern Ocean. This simple picture of deep water transport will allow us to easily understand some of the chemical differences between Pacific and Atlantic Ocean water. This model can also be used, together with ¹⁴C activities, to determine the replacement time, or ventilation time, of deep water. Stuvier et al. (1983) used this model and ¹⁴C activities measured at 124 stations occupied during the GEOSECS program from 1972 to 1978 to determine deep water residence times of 275, 250, and 510 years for the Atlantic, Indian, and Pacific Oceans respectively.

The Composition of Seawater

Table 15.1 lists the concentrations and chemical form of the elements in seawater. Concentrations range over 12 orders of magnitude (16 if H and O are included). From Figure 15.5 we see that the most abundant elements in seawater are those on the "wings" of the Periodic Table, the alkalis, the alkaline earths, and the halogens. In the terminology we introduced in Chapter 6, these elements form "hard" ions that have inert gas electronic structures. Bonding of these elements is predominately covalent; they have relatively small electrostatic energy and large radius (low Z/r ratio), so that in solution they are present mainly as free ions rather than complexes. Elements in the interior of the

Geochemistry

CHAPTER 15: OCEANS

periodic table are generally present at lower concentrations. These elements have higher Z/r ratios, form bonds of a more covalent character, and are strongly hydrolyzed. The latter tendency leads to their rapid removal by adsorption on particle surfaces. A few elements are exceptions to this pattern. These are elements, such as S, Mo, Tl and U, that form highly soluble oxyanion complexes, for example, SO_4^{2-} , MOO_4^{-2} , UO_2^{2+} or soluble simple ions (e.g., Tl⁺).

Although solubility provides a guide to elemental concentrations in seawater, the composition of seawater is not controlled by solubility. Rather, the composition of seawater is controlled by a variety of processes, from tectonism on the planetary scale to surface adsorption/desorption reactions at the atomic scale.

Many of the same processes that remove the elements from seawater, and thus play a role in controlling its composition, also impose vertical, and to a lesser degree horizontal, concentration gradients in the ocean. Table 15.1 also assigns each element to one of three categories based on their vertical distribution in the water column: C: conservative, CG: conservative gas, N: biologically controlled, "nutrient-type" distribution), S: scavenged. In the following sections, we will examine the behavior of each of these groups and the processes responsible for these gradients.

Speciation in Seawater

The wide variety of elements and the relatively high concentrations of ligands in seawater leads to the formation of a variety of complexes. The fraction of each element present as a given species may be calculated if the stability constants are known (Chapter 6). Calculation of major ion speciation requires an iterative procedure, similar to that in Example 6.7. Calculation of trace element speciation is fairly straightforward, as demonstrated in Example 15.2. Table 15.1 lists the principal

Example 15.2. Inorganic Complexation of Ni in Seawater

Using the stability constants (β^0) for Ni complexes and the free ligand concentrations in the adjacent table, calculate the fraction of total dissolved Ni in each form. Assume a temperature of 25°C. Use the following free single ion activity coefficients for the ligands: OH⁻: 0.65, Cl⁻: 0.63, CO₃: 0.2, SO₄: 0.17. Use the Davies equation (equation 3.88) to obtain the remaining activity coefficients.

Complex	$\text{Log }\beta^0$	Log[Cation]
NiOH ⁺	6.3	-5.7
$Ni(OH)_2$	12.1	-5.7
NiCl ⁺	2.8	-2.6
NiCO ₃	13.1	-4.5
$Ni(SO_4)$	2.1	-2.0

Answer: Our first task is to calculate apparent stability constants for seawater, a high ionic strength solution. The ionic strength of seawater is 0.7; using the Davies equation, we calculate a $\log \gamma$ for Ni²⁺ of -0.5, and $\log \gamma$ of -0.125 for singly charged species and a $\log \gamma$ of 0 for neutral species. The apparent stability constants may then be calculated as:

$$\log \beta^* = \log \beta^0 + \log \gamma_{Ni} + \nu \log \gamma_L - \log \gamma_{NiL}$$
 15.11

where L designates the ligand and v is its stiochiometric coefficient (e.g., 2 for Ni(OH)₂, 1 for all others). The concentration of each complexis given by:

$$[NiL_n] = \beta * [Ni^{2+}][L]^n$$

The conservation equation for Ni is:

$$\Sigma Ni = [Ni^{2+}] + [Ni(OH)^{-}] + [Ni(OH)_{2}] + [NiCl^{-}] + [NiSO_{4}] + [NiCO_{3}]$$

We can rewrite this as:

Principal Ni Complexes in Seawater % form $\log \beta^* \quad \log [NiL]/[Ni]$ Ni^{2+} 1 4.4% NiOH⁺ 3.53 -2.170% $Ni(OH)_2$ 8.12 -3.28 0% NiCl⁻ 0.02 -0.24 2.6% 1.15 0.6% NiSO₄ -0.85NiCO₃ 5.82 -1.32 92%

$\Sigma Ni = [Ni^{2+}](1 + \beta^*[Ni^{2+}][OH^+] + \beta^*[Ni^{2+}][OH^+]^2 + \dots)$

A little rearranging allows us to obtain the fraction of Ni present as each species listed in the table. Ni is present predominately as carbonate, with minor amounts of the free ion and as chloride.

CHAPTER 15: OCEANS

species present for each element. The major ions in seawater, Na^{2+} , K^+ , Mg^{2+} , Ca^{2+} , Cl^- , SO_4^{2-} , and HCO_3^- are predominately present (>95%) as free ions. Many of the trace metals, however, are present primarily as complexes.

	Average Open Ocean River Water					River Water
	Seawater Co	oncentration C	oncentration	Principal		Concentration
Elemer	nt (μg/liter)	$(\mu M/l)$	Range	Dissolved Species	Distributio	n (µg/liter)
Н	1.1×10^{8}	55.6×10^{6}		H ₂ O	С	1.1×10^{8}
He	7.2×10^{-3}	0.0018		He	NG	—
Li	185	26.5		Li^+	С	12
Ве	0.0023	$2.5 imes 10^{-5}$	8-50 pM	$BeOH^+$, $Be(OH)_2$	S/N	—
В	4.61×10^{3}	427		$B(OH)_3, B(OH)_4$	С	18
С	2.64×10^{4}	2200		HCO $\frac{1}{3}$, CO $\frac{2}{3}$, MgHCO $\frac{4}{3}$	Ν	—
C(org.)	1×10^{2}	8.3		various	_	—
Ν	8540	610		N_2	CG	—
Ν	430	31		NO ₃	Ν	—
0	$8.9 imes 10^8$	55.6×10^{6}		H ₂ Ô	С	—
0	2870	179	0-200 µM	O ₂	inverse N	—
F	1333	70.13		F^-	С	5.3
Ne	0.164	0.0081		Ne	CG	—
Na	1.105×10^{7}	4.806×10^{5}		Na ⁺	С	5,300
Mg	1.322×10^{6}	5.439×10^{4}		Mg^{2+} , $MgSO_{4}^{2-}$	С	3,100
A1	0.30	0.011	1-150 nM	$Al(OH)_{3'}Al(OH)_{4}$	S	50
Si	2800	100	0-250 µM	H_4SiO_4 , $H_3SiO_{\overline{4}}$	Ν	5,000
Р	62	2.0	0-3.5µM	HPO ^{2–} ₄ , NaHPO [–] ₄ , MgHPO ₄ , PO [–] ₄	3 N	115
S	9.063×10^{5}	2.826×10^{4}		SO_4^{2-} , NaSO $\overline{4}$, MgSO $\frac{2}{4}$	С	2840
Cl	$1.984 imes10^7$	5.596×10^{5}		Cl	С	4700
Ar	636	15.9		Ar	CG	_
Κ	$4.10 imes 10^5$	1.046×10^{4}		\mathbf{K}^{+}	С	1450
Ca	4.22×10^{5}	1.054×10^{4}		Ca^{+2} , CaSO $_{4}^{2-}$	~C	14,500
Sc	0.0006	1.33×10^{-5}		Sc(OH) ₃	S/N	0.004
Ti	0.007	$1.4 imes10^{-4}$	4-560 pM	TiO(OH) ₄	S/N	10
V	1.78	0.035	34-38 nM	HVO_4^{2-} , $H_2VO_{\overline{4}}$	~C	0.8
Cr	0.2	0.004	2.3-5.5 nM	CrO_4^{2-} , NaCrO ₄ , Cr(OH) $\frac{1}{2}$	S/N	1
Mn	0.02	$3.7 imes 10^{-4}$	<0.3-40 nM	Mn^{+2} , $MnCl^+$	S	8.2
Fe	0.03	$5.5 imes 10^{-4}$	0.05 – >6nM	Fe^{2+} , $FeCl^+$, $Fe(OH)_3$	S/N	50
Со	0.002	$3.4 imes 10^{-5}$	7-70pM	Co^{+2} , $CoCl^+$	S	0.2
Ni	0.49	$8.4 imes10^{-3}$	3-12 nM	Ni ⁺² , NiCl ⁺ , NiCO ₃	Ν	0.5
Cu	0.15	$2.4 imes10^{-3}$	0.8-4nM	$CuCO_3$, $Cu(CO_3)_2^{2-}$, $CuOH^+$	S/N	1.5
Zn	0.38	0.006	0.5-9 nM	Zn^{+2} , $ZnCl^{+}$, $ZnSO_{4}$	Ν	30
Ga	0.0012	$1.8 imes10^{-5}$	2-30 pM	Ga(OH) ₃	S/N	0.09
Ge	0.05	0.0007	<5-200 pM	$H_4GeO_4, H_3GeO_{\overline{3}}$	Ν	0.09
As	1.23	0.016	13-27 nM	$As(OH)_{3'}As(OH)_{\overline{4}}$	S/N	1.7
Se	0.159	0.002	0.5-2.5 nM	SeO_{3}^{2-} , SeO_{4}^{2-}	Ν	0.003
Br	$6.9 imes 10^4$	863		Br ⁻	С	20
Kr	0.32	0.0038		Kr	CG	
Rb	124	1.45		Rb^+	С	1.5
Sr	7930	90.5	89-92 μM	Sr^{+2}	~C	60
Y	0.017	$1.96 imes 10^{-4}$	·	YCO $\frac{1}{3}$, Y(CO ₃) $\frac{1}{2}$	S/N	0.008
Zr	0.012	1.6×10^{-4}	10-300 pM	Zr(OH) ¹⁻	S/N	0.09
Nb	0.0046	$5 imes 10^{-5}$	1	NbO = , NBO	S?	_
Мо	11	0.114		MoO_{4}^{-2}	C	0.5
Ru	< 0.005	$<5 \times 10^{-5}$		- 4	?	_
Rh	.08	$8 imes 10^{-4}$			Ν	_

TABLE 15.1. CONCENTRATIONS OF ELEMENTS DISSOLVED IN SEAWATER AND RIVER WATER

Geochemistry

CHAPTER 15: OCEANS

TABLE 15.1 (CONTINUED)

Averag		verage Open Ocean			River Water
	Seawater	Concentration Concentration	Principal	C	Concentration
Eleme	nt (μ g/liter)	$(\mu M/l)$ Range	Dissolved Species	Distribution	(µg/liter)
Pd	4.3×10^{-5}	4×10^{-7} 0.2-0.7 pM	PdCl ^{2–}	Ν	
Ag	0.023	2.2×10^{-5} 1-30 pM	AgCl	Ν	0.3
Cd	0.067	6×10^{-4} 0-1000 pM	$CdCl^+$, $CdCl_2$	Ν	0.02
In	0.00001	9×10^{-8} 0.02-0.15 pM	In(OH) ₃	S	—
Sn	$4.7 imes 10^{-4}$	4×10^{-6} 1.4-40 pM	$SnO(OH)_{\overline{3}}$	S	—
Sb	0.24	0.002 0.7-2 nM	Sb(OH) ₃	~C	0.008
Те	0.0001	8×10^{-7} 0.1-1.7 nM	Te(OH) ₆ , TeO $\frac{2}{3}$, HTeO $\frac{1}{3}$	S	—
Ι	59.5	0.468 0.4-0.6 μM	IO $\overline{3}$	~C	0.05
Xe	0.065	$5 imes 10^{-4}$	Xe	CG	—
Cs	0.3	2.3×10^{-3}	Cs^+	С	0.035
Ва	15	0.11 30-130 pM	Ba ⁺² , BaCl ⁺	Ν	60
La	5.7×10^{-3}	4.13×10^{-5} 4-50 pM	LaCO $\frac{1}{3}$, La(CO ₃) $\frac{1}{2}$	S/N	0.019
Ce	$7.2 imes 10^{-4}$	5.12×10^{-6} 2-8 pM	CeCO $\frac{1}{3}$, Ce(CO ₃) $\frac{1}{2}$	S/N	0.024
Pr	$7.2 imes 10^{-4}$	5.09×10^{-6} 1-10 pM	$PrCO_{3}^{+}$, $Pr(CO_{3})_{\overline{2}}$	S/N	0.005
Nd	3.4×10^{-3}	2.35×10^{-5} 3-40 pM	NdCO $\frac{1}{3}$, Nd(CO ₃) $\frac{1}{2}$	S/N	0.018
Sm	$5.8 imes 10^{-4}$	3.89×10^{-6} 1-8 pM	SmCO $\frac{1}{3}$, Sm(CO ₃) $\frac{1}{2}$	S/N	0.004
Eu	$1.7 imes 10^{-4}$	1.15×10^{-6} 0.1-2 pM	EuCO $\frac{1}{3}$, Eu(CO ₃) $\frac{1}{2}$	S/N	0.001
Gd	9.2×10^{-4}	5.87×10^{-6} 1-10 pM	$GdCO_{3}^{+}, Gd(CO_{3})_{\overline{2}}^{-}$	S/N	0.006
Tb	$1.7 imes 10^{-4}$	1.1×10^{-6} 0.3-2 pM	TbCO $\frac{1}{3}$, Tb(CO ₃) $\frac{1}{2}$	S/N	0.001
Dy	1.12×10^{-3}	6.94×10^{-6} 1.5-13 pM	$DyCO_{3}^{+}$, $Dy(CO_{3})_{\overline{2}}^{-}$	S/N	0.005
Ho	3.7×10^{-4}	2.24×10^{-6} 0.4-3.7 pM	HoCO $\frac{1}{3}$, Ho(CO ₃) $\frac{1}{2}$	S/N	0.001
Er	1.2×10^{-3}	7.35×10^{-6} 1.5-12 pM	$\operatorname{ErCO}_{3}^{+}$, $\operatorname{Er}(\operatorname{CO}_{3})_{\overline{2}}^{-}$	S/N	0.004
Tm	2×10^{-4}	1.21×10^{-7} 0.3-2 pM	TmCO $\frac{1}{3}$, Tm(CO ₃) $\frac{1}{2}$	S/N	0.001
Yb	1.23×10^{-3}	7.11×10^{-6} 1.5-13 pM	YbCO $\frac{1}{2}$, Yb(CO ₂) $\frac{1}{2}$	S/N	0.005
Lu	$2.3 imes 10^{-4}$	$1.35 \times 10^{-6} 0.3 - 2.3 \text{ pM}$	LuCO $\frac{3}{4}$, Lu(CO ₃) $\frac{2}{5}$	S/N	0.001
Hf	$1.6 imes 10^{-4}$	9×10^{-7} 0.4-2.4 pM	Hf(OH) =	S/N	2.5×10^{-3}
Та	2.5×10^{-3}	$1.4 imes 10^{-5}$	Ta(OH)₅	S?	_
W	0.010	$5.4 imes 10^{-5}$	WO 4-	~C	$1.6 imes 10^{-4}$
Re	0.0074	3.96×10^{-5}	$\operatorname{ReO}_{\overline{4}}$	C	0.0004
Os	$1.7 imes 10^{-6}$	$9 imes 10^{-9}$?	_
Ir	1×10^{-6}	6×10^{-9}	Ir ³⁺	S?	_
Pt	5×10^{-5}	$2.6 imes 10^{-7}$	$PtCl_{4}^{2-}$, $PtCl_{6}^{2-}$	~C	_
Au	2×10^{-5}	1×10^{-7} 0-240 fM	AuOH(H ₂ O), AuCl, AuCl $=$	S?	0.0001
He	0.00014	7×10^{-7} 0.2-2 pM	$HgCl_{4}^{2-}$, $HgCl_{7-}$, $HgCl_{2-}$	S/N	0.07
ΤΪ	1.3×10^{-2}	6.5×10^{-5} 58-78 pM	Tl^+ , TlCl	~C	$3.5 imes 10^{-4}$
Pb	2.7×10^{-3}	1.3×10^{-5} 3-170 pM	PbCl ⁻ ,PbCl ₂ , PbCO ₂	S	0.01
Bi	3×10^{-5}	1.4×10^{-7} 10-500 fM	BiO^+ , $Bi(OH) + 2$	S	_
Ро	_	_		_	_
At	_	_	_		_
Rn	_	_	Rn		_
Fr	_	_	Fr ⁺	_	_
Ra	1.3×10^{-7}	5.8×10^{-10}	Ra ⁺²	Ν	_
Ac					_
Th	2×10^{-5}	8.6×10^{-8} 50-650 fM	Th(OH).	S	0.1
Pa					
U	3.3	0.0138	$UO_2(CO_3)_3^{-4}$	С	0.19

Concentrations based on single analyses or only pre-1980 data are shown in italics. Category: C Conservative, N: Nutrient/Biologically Controlled, S: Scavenged CG: conservative gas; NG non-conservative gas. Sources: Seawater Concentrations: modified from Martin and Whitfield (1983), Broecker and Peng (1982), and Quinby-Hunt and Turekian (1983), and the electronic supplement to Nozaki (1997); Speciation: Morel and Hering (1995), Turner and Whitfield (1981), Cantrell and Byrne (1987), Bruland (1983), Erel and Morgan (1991); River Concentrations: Table 12.2, and modified from Martin and Whitfield (1983), Broecker and Peng (1982).

Geochemistry

CHAPTER 15: OCEANS



Figure 15.5. The composition of seawater. The most abundant elements are those on the sides of the periodic table. Elements in the interior tend to be less abundant.

CONSERVATIVE ELEMENTS

The conservative elements share the property of being always found in constant proportions to one another and to salinity in the open sea, even though salinity varies. All the major ions in seawater, except for bicarbonate, are included in this group. Their concentrations are listed in Table 15.2. This constancy of the major ion composition of seawater, which is typically expressed as a ratio to Cl, is sometimes called the Law of Constant Proportions, and has been known for nearly 2 centuries. For most purposes, we may state that concentrations of these elements vary in the ocean only as a result of di-

lution or concentration of dissolved salts by addition or loss of pure water. While chemical and biological processes occur within the ocean do change seawater chemistry, they have an insignificant effect on the concentrations of conservative elements.

The major ions do vary in certain unusual situations, namely (1) in estuaries, (2) in anoxic basins (where sulfate is reduced), (3) when freezing occurs (sea ice retains more sulfate than chloride), (4) in isolated basins where evaporation proceeds to the point where salts begin to precipitate, and (5) as a result of hydrothermal inputs to restricted basins (e.g. red sea brines). Ca and Sr are slight exceptions to the rule in that they are inhomogeneously distributed even in the open ocean, though only slightly. The concentrations of these elements, as well as that of HCO_{3} , vary as a result of biological production of organic carbon, calcium carbonate, and strontium sulfate[†] in the surface water and sinking of the remains of organisms into deep water. Most of these biologically produced particles breakdown in deep water, releasing these species into solution (we explore this in greater detail below). Thus there is a particulate flux of carbon, calcium, and strontium from surface waters to deep waters. As a result, deep water is about 15% enriched in bicarbonate, 1% enriched in Sr, and 0.5% enriched in Ca relative to surface water. As we shall see, these biological processes also create much

TADLE	15.2.	MAIOR	lons	in	SEAWATER
TADIL	I / . Z .	MAJOK	10113		JLAWAILK

Ion	g/kg (ppt)	Percent of
	at 5 = 55700	Dissolved solids
CI⁻	19.354	55.05
SO ²⁻ / ₄	2.649	7.68
HCO ₃	0.140	0.41
$B(OH)_{4}^{-}$	0.0323	0.07
Br ⁻	0.0673	0.19
F-	0.0013	0.00
Na ⁺	10.77	30.61
Mg^{2+}	1.290	3.69
Ca ²⁺	0.412	1.16
K^+	0.399	1.10
Sr^{2+}	0.008	0.03

larger vertical variations in the concentrations of many minor constituents.

Some minor and trace elements are also present in constant proportions; these include Rb, Mo, Cs, Re, Tl, and U. Vanadium is nearly conservative, with a total range of only about $\pm 15\%$. All these elements share the properties that they are not extensively utilized by the biota and form ions or radicals that are highly soluble and not surface reactive.

Dissolved Gases

The concentrations of dissolved gases in the oceans are maintained primarily by

⁺ A class of protozoans called Acantharia build shells of SrSO₄.

Geochemistry

CHAPTER 15: OCEANS

exchange with the atmosphere. These gases may be divided into conservative and non-conservative ones. The noble gases and nitrogen constitute the conservative gases. As their name implies, their concentrations are not affected by internal processes in the ocean. Concentrations of these gases are governed entirely by exchange with the atmosphere. Since they are all minor constituents of seawater, we can use Henry's Law to describe the equilibrium solubility of atmospheric gases in the ocean:

$$C = kp$$
 15.12

where C is the concentration in seawater, k is the Henry's Law constant, or Bunsen absorption coefficient, and p is the partial pressure of the gas in

	Atmospheric Equilibrium Conc. in		
	Partial	Seawate	r (ml/l)
	Pressure	0°C	24°C
He	5.2	4.1×10^{-5}	3.8×10^{-5}
Ne	1.8	1.8×10^{-4}	1.5×10^{-4}
N_2	0.781	14.3	9.2
O ₂	0.209	8.1	5.0
Ar	9.3×10^{-3}	0.39	0.24
Kr	1.1×10^{-6}	9.4×10^{-5}	$8.5 imes 10^{-5}$
Xe	8.6×10^{-8}	1.7×10^{-5}	8.5×10^{-6}
CO ₂	3.6×10^{-4}	0.47	0.24
N_2O	3×10^{-7}	3.2×10^{-4}	1.4×10^{-4}

TABLE 15.3. Dissolved Gases in Seawater

the atmosphere. The equilibrium concentrations of atmospheric gases in seawater at 1 atm (0.1 MPa) are listed in Table 15.3. The conservative gases are not uniformly distributed in the ocean. This is because of the temperature dependence of gas solubility: they are more soluble at lower temperature. Over a temperature range of 0° to 30° C, this produces a variation in dissolved concentration of about a factor of two for several gases. As may be seen in Figure 15.6 and Table 15.3, the temperature dependence is strongest for the heavy noble gases and CO_2 , and weakest for the light noble gases. Thus the concentration of conservative gases in seawater depends on the temperature at which atmosphere-ocean equilibration occurred. Another interesting aspect of the solubilities curves in Figure 15.6 is their non-linearity. Because of this non-linearity: mixing between water masses that have equilibrated with the atmosphere at different temperatures will lead to concentrations above the solubility curves. We also notice in Figure 15.6 that solubility for the different gases ranges over nearly 2 orders of magnitude; the light noble gases are the least soluble; the heavy noble gases and CO_2 are



Figure 15.6. Solubility of gases in seawater as a function of temperature. Adapted from Broeker and Peng (1982).

the most soluble.

Gas solubility is also strongly dependent on salinity and pressure. For example, at 20 ° C the solubility of oxygen is 20% lower in seawater with a salinity of 35‰ than in pure water. Over the range of salinities typical of open ocean water (34‰ to 38‰) solubility of oxygen varies by about 3%. According to equation 15.11, the equilibrium concentration will increase directly with pressure. Atmospheric pressure at the air-sea interface is effectively constant; however, pressure increases rapidly with depth in the ocean, by 1 atm for every 10 meters depth. When actual concentrations are compared with predicted equilibrium concentrations, the surface ocean is oversaturated by a few percent. This is thought to be due to bubbles, produced by breaking waves, being carried to depth in the ocean. Bubbles need only be carried to depths of a few tens of centimeters to account for the observed oversaturation.

 O_2 and CO_2 are the principal non-conservative gases. They vary because of photosynthesis and respiration. Nitrogen is also biologically utilized. However, only a small fraction of the dissolved N_2 is present as "fixed" nitrogen (as NH_4 ,

Geochemistry

CHAPTER 15: OCEANS



Figure 15.7. Temperature, salinity and oxygen distribution in a north-south cross-section of the Atlantic Ocean. The salinity panel also shows water movements and water masses, indicated by numbers: ① North Atlantic Bottom Water, ② Mediterranean Water, ③ N. Atlantic Deep Water, ④ Antarctic Bottom Water, ⑤ Antarctic Intermediate Water.

mosphere. The longer deep water has been away from the surface, the more depleted in oxygen it will be. The second factor is the abundance of organic matter. The abundance of food decreases with depth, so that respiration, and hence oxygen consumption, is highest just below the photic zone and lowest in the deepest water. Thus the vertical distribution of oxygen is characterized by an oxygen minimum that typically occurs within the thermocline. Furthermore, the rate of organic matter production in the surface waters varies geographically (for reasons we will subsequently discuss). Oxygen is more depleted in deep water underlying high biological productivity regions that beneath regions of low productivity.

Figure 15.7 shows the distribution of O_2 in the Atlantic Ocean. Highest O_2 concentrations are found in surface waters at high latitudes, where the water is cold and the solubility of O_2 is highest. The

 NO_{3}^{-} , and $N_{2}O$), hence N_{2} behaves effectively as a conservative element.

Helium is another non-conservative gas because of the input of He to the ocean by hydrothermal activity at mid-ocean ridges. Elevated He concentrations and high ${}^{3}\text{He}/{}^{4}\text{He}$ ratios found at mid-depth, particularly in the Pacific, reflecting this injection of mantle He by mid-ocean ridge hydrothermal systems. In the following sections, we examine the variation of O₂ and CO₂ in the ocean in greater detail.

O_2 Variation in the Ocean

Biological activity occurs throughout the oceans, but is concentrated in the surface water because it is only there that there is sufficient light for photosynthesis. This part of the water column is called the photic *zone*. Both respiration and photosynthesis occur in the surface water, but the rate of photosynthesis exceeds that of respiration in the surface ocean, so there is a net O_2 production the surface layer. Most organic matter produced in the surface ocean is also consumed there, by a small fraction sinks into the deep water. This sinking organic matter is consumed by bacteria and scavenging organisms living in the deep water. Respiration in the deep water and the absence of photosynthesis results in a net consumption of oxygen. Within the deep water, two factors govern the distribution of oxygen. The first is the "age" of the water, the time since it last exchanged with the at-

Geochemistry

CHAPTER 15: OCEANS

minimum O_2 concentrations are found in mid-latitudes at depths of 500 to 1000 m. Oxygen minima at these depths characterize most of the world ocean at temperate and tropical latitudes. At greater depth, as well as at higher latitudes, the concentration of O_2 is higher because the water there is generally younger, i.e., it has more recently exchanged at the surface. Because deep water in the Pacific and Indian Oceans is older than deep water in the Atlantic, oxygen concentrations are generally lower. A particularly strong O_2 depletion occurs beneath high productivity regions of the eastern equatorial Pacific and conditions are locally suboxic (i.e., no free O_2). Anoxic conditions develop in deep water in basins where the connection to the open ocean is restricted. The best example is the Black Sea. The Black Sea is a 2000 m deep basin whose only connection with the rest of the world ocean is through the shallow Bosporus Strait. As a result, water becomes anoxic at a depth of about 100 m. Anoxia is also present in the Curaco Trench, off the northern coast of South America. Here anoxia is a result both of restricted circulation and high productivity in the overlying surface water. Anoxic conditions also develop in some deep fjords.

Distribution of CO_2 in the Ocean

As we found in previous chapters, most CO_2 dissolved in water will be present as carbonate or bicarbonate ion. Nevertheless, we will often refer to all species of carbonate and CO_2 simply as CO_2 . CO_2 cannot be treated strictly as a dissolved gas, as there are sinks and sources of CO_2 other than the atmosphere. For example, much of the dissolved CO_2 is delivered to the ocean by rivers as bicarbonate ion. These properties make the distribution of CO_2 more complicated than that of other gases.

Like oxygen, CO_2 concentrations are affected by biological activity and its solubility is affected by temperature. These factors result in a significant geographic variation in CO_2 in surface water. Surface water is often supersaturated with respect to the atmosphere in equatorial regions as upwelling brings CO_2 -riched deeper water to the surface and warming decreases it solubility. In the subtropical gyres P_{CO_2} is generally maintained below saturation values by photosynthesis. The greatest degree of undersaturation occurs in polar regions, where photosynthesis decreases CO_2 and cooling increases it solubility. The North Pacific is an exception as it appears to be supersaturated both within much of

the North Pacific gyre and at high latitudes (Takahashi, 1989). Thus there is a net flux of CO_2 from the ocean to the atmosphere in low latitudes and a net flux from the atmosphere to the ocean in high latitudes.

Biological activity is responsible for vertical variations in CO_2 in the ocean. Photosynthesis converts CO_2 to organic matter in the surface water. Most of this organic matter is remineralized within the photic zone, but some 5% is transported out of this zone into deep water (mainly by falling fecal pellets, etc.; but vertically migrating zooplankton and fish also transport organic carbon from the surface to the deep layer), depleting surface water in CO₂. Respiration converts most of the falling organic matter back into dissolved CO₂ and only a very small fraction of the organic matter produced is buried in the sediment. This aspect of biological activity thus affects CO₂ distributions in exactly opposite way it affects oxygen. However, a few planktonic organisms, most notably foraminifera (protozoans), pteropods (snails), and coccolithophorids (algae), produce carbonate shells, which results in an additional extraction of CO₂ from surface waters. These shells, or tests as they are properly called, also sink into the deep water when the organisms die. The solubility of calcium carbonate increases with depth, for reasons we will discuss shortly, so that much of the car-



Figure 15.8. Depth profile of total dissolved inorganic carbon (ΣCO_2) and $\delta^{13}C$ of dissolved inorganic carbon in the North Atlantic.

Geochemistry

CHAPTER 15: OCEANS

bonate redissolves. The total amount of dissolved CO₂ converted to carbonate is small compared to that converted to organic carbon. However, a much large fraction of biogenic carbonate sinks out of the photic zone, so that the downward flux of carbon in carbonate represents about 20% of the total downward flux of carbon. A larger fraction of carbonate produced is also buried, so that the flux of carbon out the ocean is due primarily to carbonate sedimentation rather than organic matter sedimentation.

The transport of CO₂ from surface to deep water as organic matter and biogenic carbonate is called the biological pump. As we might expect, the biological pump produces an enrichment of CO_2 in the deep ocean over the shallow ocean, as is illustrated in Figure 15.8. Many vertical profiles of ΣCO_2 show a maximum at the same depth as the oxygen minimum, although the example in Figure 15.8 does not. It occurs for the same reasons as the oxygen minimum: there is more organic matter at this level and hence higher respiration, and deep water is often "younger". As does oxygen enrichment, the extent of enrichment of CO_2 in deep water depends on the age of the water mass and the downward flux of organic matter (and therefore ultimately on the intensity of photosynthesis in the overlying water). It depends additionally on the rate of calcium carbonate dissolution.

Biological activity also produces a variation in the isotopic composition of carbon is seawater. We found in Chapter 9 that photosynthetic organisms utilize ¹²C in preference to ¹³C. Thus photosynthetic activity in the upper layer depletes surface water in 12 C, increasing δ^{13} C. When organic matter is remineralized at depth, the opposite occurs: deep water in enriched in ¹²C. Biological activity therefore imposes a gradient in δ^{13} C on the water column (Figure 15.8). Comparing the Σ CO₂ with the δ^{13} C profile, we see that the latter shows a pronounced maximum while the former does not. Why? The answer to this question is left as a problem at the end of the chapter.

The extent of depletion of ${}^{12}C$ in surface water will depend on biological activity: $\delta^{13}C$ will be higher in productive waters than in unproductive waters. The extent of enrichment of ¹²C in deep water, as does CO₂, depends on the age of the deep water. "Old" deep water will have lower δ^{13} C than "young" deep water.

Seawater pH and Alkalinity

We found in Chapter 6 that the pH of most natural waters is buffered by the carbonate system and this is certainly true of seawater. Compared to other natural waters, seawater has a relatively constant pH, with a mean of about 8, but the variations in dissolved CO_2 do produce pH variations of about ± 0.3 . This variation is largely due to biological activity: removal of dissolved CO₂ by photosynthesis increases pH, while release of CO₂ by respiration decreases it. The reason for this is easy to understand. At the pH of seawater, bicarbonate is the predominant carbonate species. Thus we can describe the dissolution of CO₂ as:

$$\text{CO}_2 + \text{H}_2\text{O} \rightleftharpoons \text{H}^+ + \text{HCO}_{-3}$$
 15.13

Photosynthesis extracts CO₂ from water, so reaction 15.13 is driven to the left, consuming H⁺. Respiration produces CO₂, driving this reaction to the right, producing H⁺. For this reason, the pH of the ocean decreases with depth. In the profile shown in Figure 15.9, we see a minimum in pH at the same depth as the O_2 minimum, reflecting the high rate of respiration at this depth.

pH is also affected by precipitation and dissolution of Figure 15.9. pH profile in the North Pacalcium carbonate. Since bicarbonate is the most abundant carbonate species, the precipitation reaction is effectively:



cific Ocean. Position of the oxygen minimum is shown.

Geochemistry

CHAPTER 15: OCEANS

$$Ca^{2+} + HCO_{\overline{3}} \rightleftharpoons H^{+} + CaCO_{\overline{3}}$$
 15.14

Here is it easy to see that precipitation of calcium carbonate decreases pH while dissolution increases it. Thus production of biogenic carbonate in surface water and its dissolution in deep water acts to reduce the vertical pH variations produced by photosynthesis and respiration.

Another important parameter used to describe ocean chemistry, and one closely related to pH is alkalinity. In Chapter 6 we defined alkalinity as the sum of the concentration (in equivalents) of bases that are titratable with strong acid. It is a measure of acid-neutralizing capacity of a solution. An operational definition of total alkalinity for seawater is:

$$Alk = [HCO_{3}^{-}] + 2[CO_{3}^{2-}] + [B(OH)_{4}^{-}] + [H_{2}PO_{4}^{-}] + 2[HPO_{4}^{2-}] + [NO_{3}^{-}] + [OH^{-}] - [H^{+}]$$
 15.15

Often, particularly in surface water, the phosphate and nitrate terms are negligible (in anoxic environments, we would need to include the HS⁻ ion). Carbonate alkalinity is:

$$CAlk = [HCO_{3}^{-}] + 2[CO_{3}^{2-}] + [OH^{-}] - [H^{+}]$$
 15.16

(which is identical to 6.32). One of the reasons alkalinity is important is that it can be readily determined by titration.

In Chapter 6, we stated that alkalinity is "conservative", meaning that it cannot be changed except by the addition or removal of components. It is important to understand that alkalinity is not conservative in an oceanographic sense, as is, for example, salinity. In an oceanographic sense, we define a "conservative" property to be one that changes only at the surface by concentration or dilution. While addition and removal of components may occur, through precipitation and dissolution, these processes have negligible effects on conservative properties. Concentration and dilution affect alkalinity; indeed, these processes are the principal cause of variation in alkalinity (alkalinity is strongly correlated with salinity). However, precipitation and dissolution in the ocean do significantly affect alkalinity (whereas the affect on salinity is negligible), so alkalinity is not conservative in an oceanographic sense. Indeed, alkalinity typically varies systematically with depth, being greater in deep water than in the surface water.

What causes this depth variation? It might be tempting to guess that photosynthesis and respiration are responsible. However, these processes have no direct effect on alkalinity. When CO_2 dissolves in water, it dissociates to produce a proton and a bicarbonate ion. In the alkalinity equation, these exactly balance, so there is no effect on alkalinity. Production and oxidation of organic matter do affect alkalinity through the uptake and release of phosphate and nitrate, but the concentration of these nutrients is generally small. The main cause of the systematic variation of alkalinity in the water column is carbonate precipitation and dissolution. For every mole of calcium carbonate precipitated, a mole of carbonate is removed and alkalinity increases by 2 equivalents, and visa versa, so the effect is quite significant.

Carbonate Dissolution and Precipitation

From the preceding sections, we can see that precipitation of calcium carbonate in surface waters and its dissolution at depth is an important oceanographic phenomenon. Carbonate sedimentation is also an important geological process in other respects, including its role in the global carbon cycle. Let's examine carbonate precipitation and dissolution in a little more detail. Two forms of calcium carbonate precipitate from seawater. Most carbonate shell-forming organisms, including the planktonic foraminifera and coccolithophorids that account for most carbonate precipitated, precipitate calcite. Pteropods and many corals, however, precipitate aragonite, even though aragonite, the high pressure form of calcium carbonate, is not thermodynamically stable anywhere in the ocean. The surface ocean is everywhere supersaturated with respect to both calcite and aragonite, usually to depths of 1000 m or more¹. Nevertheless, except in some rather rare and unusual situations, carbonate pre-

¹ You might ask how aragonite can be supersaturated if it is not thermodynamically stable. It is supersaturated because aragonite has a lower Gibbs Freee Energy than seawater, but aragonite has a higher Gibbs Free Energy than calcite, so it is unstable with respect to calcite.

Geochemistry

CHAPTER 15: OCEANS

cipitation occurs only when biologically mediated. There are two interesting questions here. First, why does the ocean go from supersaturated at the surface to understaturated at depth, and second, why doesn't calcium carbonate precipitate without biological intervention?

There are three reasons why the oceans become undersaturated with respect to calcium carbonate a t depth. First, increasing P_{CO_2} of deep water drives pH to lower levels, increasing solubility. This might seem counter-intuitive, as one might think that that increasing P_{CO_2} should produce an increase the carbonate ion concentration and therefore drive the reaction toward precipitation. However, increases in P_{CO_2} and ΣCO_2 with depth produce a *decrease* in CO_3^- concentration. This is most easily understood if we express the carbonate ion concentration as a function of P_{CO2} using the solubility and dissociation constants for the carbonate system (equations 12.21 through 12.23):

$$[CO_3^{2-}] = \frac{K_2 K_1 K_{Sp-CO_2} P_{CO_2}}{[H^+]^2}$$
 15.17

This equation shows that while carbonate is proportional to $P_{CO_{\gamma}}$ it is inversely proportional to the square of $[H^+]$. The pH drop resulting from production of CO_2 by respiration is thus dominant. Carbonate ion concentrations drop by over a factor of three from the surface waters to the waters with the highest dissolved CO₂.

The second reason is that the solubility of calcium carbonate increases with increasing pressure. This results from the positive ΔV of the precipitation reaction. Calcite and aragonite are about twice as soluble at 5000 m (corresponding to a pressure of 500 atms) than at 1 atmosphere. Third, the solubility of CaCO₃ changes with temperature, reaching a maximum around 12°C (see Example 15.3). As we might expect, the solubility of calcite is also dependent on salinity (due to the effect of ionic strength on the activity coefficients), but salinity variations are not systematic with depth.

The kinetics of carbonate precipitation are still not fully understood, in spite of several decades of research. Quite a bit is known, however, particularly about the calcite precipitation and dissolution. A number of laboratory studies (e.g., Chou et al., 1989; Zuddas and Mucci, 1994) have concluded that the principal reaction mechanism of calcite precipitation in seawater is:

$$Ca^{2+} + CO_3^{2-} \rightleftharpoons CaCO_3$$
 15.18



5000

5000 m is 50 MPa. Substutiting values, we can construct the graph shown in Figure 15.10. We see that calcite is somewhat more than twice as Figure 15.10. Calculated change of the calcit soluble at a depth of 5000 m than at the surface.



Geochemistry

CHAPTER 15: OCEANS

Example 15.4. Temperature Dependence of Calcite Solubility

For sea water of 35% salinity, the temperature dependence of the apparent calcite solubility product may be expressed as (in units of moles² per kg^2):

 $\log K' = A + BT + C/T + D\log(T)$

where A = -178.34874, B= -0.061176, C= 3894.39267, D=71.595 (Mucci, 1983; Millero, 1995). For seawater of average Ca^{2+} ion concentration (Table 15.2), how does the carbonate ion concentration at which saturation is acheived vary between 0°C and 35°C?

Answer: The equilibrium carbonate ion concentration is given by:

$$[CO_3^{2-}] = \frac{K_{CaCO_3}}{[Ca^{2+}]} \qquad 15.24$$

The concentration of Ca given in Table 15.2 corresponds to a molal concentration of 10.28 mM/kg. Substituting 15.23 into 14.24, we can construct the graph shown in Figure 15.11. Maximum solubility is achieved at about 12°C and decreases at both higher and lower temperatures.



Figure 15.11. Calculated temperature dependence of the concentration of carbonate ion in 35% salinity seawater in equilibrium with calcite.

In other words, this simple stoichiometric expression best represents what actually occurs on an molecular level (however, other mechanisms appear to predominate at lower pH). In Chapter 5, we found that the net rate of reaction can be expressed as:

$$\mathfrak{R}_{\rm net} = \mathfrak{R}_{\perp} + \mathfrak{R}_{\perp} \tag{5.73}$$

If reaction 15.18 is the elementary reaction describing precipitation, then:

$$\Re_{+} = k_{+} \left[Ca^{2+} \right] \left[CO_{3}^{2-} \right]$$
 15.19

where k_{\perp} and k_{\perp} are the forward and reverse rate constants respectively. Taking the concentration of $CaCO_3$ in the solid is 1, the net reaction rate should be:

 $\Re_{-} = k_{-} [CaCO_3]$

$$\Re_{\text{net}} = k_+ [\text{Ca}^{2+}][\text{CO}_3^{2-}] - k_-$$
 15.21

However, Zhong and Mucci (1993) and Zuddas and Mucci (1994) found that under conditions of constant $[Ca^{2+}]$, the overall rate equation is:

$$\Re_{\text{net}} = k_{\text{f}} [\text{CO}_{3}^{2-}]^{3} - k_{-}$$
 15.22

where k_t is an apparent rate constant (incorporating both the rate constant and Ca²⁺ concentration). In other words, the reaction is third order with respect to the carbonate ion, rather than first order as expected. This indicates that other processes must be involved and strongly influence the reaction rate. Exactly what these other processes are is not yet fully understood.

It is known that the presence of Mg^{2+} and SO_4^{2-} ions strongly retard calcite precipitation rates (Berner, 1975; Busenberg and Plummer, 1985). Why is not yet fully understood. One possibility is that the formation of ion pairs, such as $(Mg^{2+})(CO_3^{--})$ and $(Ca^{2+})(SO_4^{2-})$, reduces the availability of reactants. Another possibility is that Mg and SO_4^{2-} are absorbed on the surface and thus block addition of new Ca² and CO $\frac{2}{3}$ to the surface.

15.20

Geochemistry

W. M. White

CHAPTER 15: OCEANS

In contrast to the precipitation reaction, dissolution of carbonate appears to begin close to the depth where undersaturation is reached. However, dissolution is not instantaneous. If we could remove the water from the ocean basins, we would find that mountain peaks in the ocean are covered with white carbonate sediment, but that carbonate is absent from the deeper plains and valleys. This picture is very much reminiscent of what we often see in winter: mountains peaks covered with snow while the valleys are bare. Although the snow line, the elevation where we first find snow, depends on temperature, it does not necessarily correspond to the 0° C isotherm. Indeed, it will generally be somewhat lower than this. The snow line is that elevation where the rate at which snow falls just matches the rate at which it melts. In the ocean, the "carbonate snow line" is called the carbonate compensation depth (CCD), and it is depth at which the rate at which carbonate accumulates just equals the rate at which it dissolves. This is always lower than the depth where it begins to dissolve, a depth known as the *lysocline*. Like the snow line, the depth of the CCD does ultimately depend on thermodynamic factors, and it varies. It is deepest in the



Figure 15.12. Dissolution rate of calcite as a function of depth in the Pacific at 19° N 156° W as determined by the Peterson spheres experiment.

Atlantic, where it as deep as 5500 m, and shallowest in the North Pacific, where it is as shallow as 3500 m. The average depth of the CCD is about 4500 m. Since aragonite is more soluble than calcite, it is restricted to even shallower depths. Pteropod oozes, sediments composed primarily of the aragonitic shells of pteropods, are largely restricted to the mid-ocean ridge crests and tops of seamounts.

The survival of carbonate in sediment is, then, a question of kinetics as much as of thermodynamics. The kinetics carbonate dissolution in seawater have been addressed with both laboratory and field experiments. Figure 15.12 shows the rate of calcite dissolution determined by an ingenious experiment performed by Peterson (1966) three decades ago. Peterson hung carefully weighed spheres of calcite at various depths in the ocean for 265 days. He then recovered the spheres and reweighed them to determine the rate of dissolution as a function of depth. The results showed a rapid increase in dissolution at a depth of about 3500 m. Since then, this experiment has been duplicated several times with increasing sophistication. The results show that the depth at which rapid dissolution begins, the lysocline, corresponds reasonably closely to the depth at which undersaturation is reached. Below that depth, the rate of dissolution increases rapidly, and correlates with the degree of undersaturation. Laboratory studies have shown that at the pH of seawater, the dominant dissolution mechanism also appears to be the reverse of reaction 15.18 (e.g., Chou et al., 1989). If that is the case, then the net rate of dissolution should be:

$$\Re_{\text{net}} = k_{-} - k_{f} [\text{CO}_{3}^{2-}]^{3}$$
 15.25

The strong dependence of the dissolution rate on carbonate ion concentration is consistent with the rapid increase in dissolution rates below the lysocline. Hence field and laboratory investigations appear to yield consistent results, but more experimental work on dissolution mechanisms and rates under conditions relevant to seawater is still needed.

Geochemistry

CHAPTER 15: OCEANS

NUTRIENT ELEMENTS

In addition to CO_2 , many other elements are essential to life and are taken up by phytoplankton in the course of organic matter production. Two of the most important nutrient elements in the ocean are phosphorus and nitrogen. As we found in Chapter 14, nitrogen is a component of proteins and nucleic acids, and phosphorus is involved in many energy transfer reactions. Both are essential for all life. For most primary producers, nitrogen must be "fixed" before it can be utilized. Thus NO_{3} , N_2O , and NH_{4}^{+} are nutrients, but N_2 is not (in the following discussion we shall use N to refer only to fixed inorganic nitrogen). Biological uptake of these nutrients typically leads to their severe depletion in surface waters. Indeed, nitrate levels sometimes fall below detection limits. The uptake of nutrients from surface waters and their release into deep water from falling organic particles imposes a vertical gradient on the concentration on nutrients in the ocean (Figure 15.13). Redfield (1958) found that C, N, and P were present in living tissue in nearly constant proportions of 106:16:1. These are called the *Redfield ratios* (more recent work suggests that marine organic matter is actually richer in carbon and just slightly poorer in nitrogen, with the best current estimate of the Redfield ratios being about 126:16:1). Thus we would expect the concentration of phosphate and nitrate to be highly correlated in seawater, and this is indeed the case.

Several planktonic organisms, most notably diatoms, which are among the most important photosynthesizers in the ocean, build tests of SiO_2 . For such organisms, Si is as important a nutrient as N and P and it is also strongly depleted in surface water (Figure 15.14a). It appears that in many regions of the oceans, the availability of one or all of these three nutrients limits biological productivity. For this reason, these elements are known as *biolimiting*.

Maxima in nutrient profiles sometimes occur in the depth range of 500-1500 m, reflecting relatively high rates of respiration at these depths. Since nutrients accumulate in deep water, we would expect that the oldest waters would have the highest nutrient concentrations. This is indeed the case. Atlantic deep water shows considerably less enrichment in nutrients than the Pacific because it is



Figure 15.13. Depth profiles of two important nutrients, nitrate and phosphate, in the oceans.

Geochemistry



Figure 15.14. Depth profiles of Si (a) and (b) Ge in the Pacific (solid lines) and the Atlantic (dashed red lines). (c) shows the correlation between Si and Ge in the Pacific profile. Modified from Froelich et al. (1985).

younger. The flux of deep water from the Atlantic to the Pacific results in a flux of nutrient elements from the Atlantic to the Pacific and an enrichment the Pacific in nutrient elements compared to the Atlantic.

Upwelling of deep water returns nutrients to surface water. Rivers and wind blown dust deliver nutrients to surface waters from the continents. Thus coastal regions and regions of upwelling, such as in polar regions and along the equator, characteristically have high biological productivity due to this addition of nutrients. External nutrient sources are very limited in the great central gyres, so that these regions have low biological productivity.

Many other elements are also essential for life. B, Na, Mg, S, Cl, K, Ca, and Mo are widely or universally required and F, Br and Sr are required by some species. However, these elements are sufficiently abundant in seawater that biological activity produces no, or very little, variation in their concentration in the ocean. For this reason, these elements are sometimes referred to as biounlimited. In addition to these elements, Mn, Fe, Cu, and Zn are widely or universally required and the elements V, Cr, Ni, Se, and I are required by some species. These elements are in sufficiently low concentrations that biological activity imposes vertical and horizontal gradients in their concentration in the ocean. Some of these elements, such as Fe and Zn, also show severe depletion in the surface water and are also classed as

biolimiting elements. Indeed, Martin and Gordon (1988) argued that the availability of Fe limits phytoplankton growth in high productivity regions such as the Subarctic, Antarctic, and equatorial Pacific, where major nutrients are not severely depleted. In these areas, Fe concentrations are as low as 0.1 nmol/l. That Fe can indeed limit phytoplankton productivity was confirmed in an experiment carried out in the eastern equatorial Pacific in 1993. Fe (as $FeSO_4$) was added to a 100 km² patch of ocean where Fe concentrations were very low. Within days of adding Fe, phytoplankton productivity increased significantly (Martin et al., 1994).

For other elements, such as V, Cr, and Se, only slight surface water depletion occurs. Elements such as these are known as *biointermediate*. An example of one of these biointermediate elements, Se, is shown in Figure 15.15.

The distributions of several of these biolimiting and biointermediate elements are also significantly affected by non-biological processes in the ocean and as a result, they can have vertical concentration profiles that differ from the classic nutrient profile see in Figures 15.13 and 15.14. This can occur as a result of oxidation and reduction, adsorption onto particle surfaces, a process called scavenging, and by other inputs, such as input from hydrothermal vents and sediments of the ocean bottom. We will consider the vertical distribution of these elements, and the processes creating them, in more detail in the next section.

Geochemistry

CHAPTER 15: OCEANS



Figure 15.15. Selenium profiles in the Panama Basin. Circles: Se^{IV}, crosses: Se^{VI}, triangles: total Se. From Measures et al. (1984).

ated decomposition of soft tissues of organisms. Si is incorporated only in the inorganic tests, which undergo dissolution without biological mediation (however, passage of these tests through the gut of higher organisms undoubtedly speeds dissolution). This abiologic dissolution is slower. Elements that are rapidly released are referred to as *labile* elements; those released more slowly, such as Si, are referred to as refractory. Labile nutrients include nitrate, Mn, Cd, and Ni. Since Ge concentrations correlated with those of Si, we include it with the refractory elements. It was once thought that all refractory elements were incorporated in hard tissues, such as tests. However, Collier and Edmond (1984) demonstrated that several refractory elements are present only in soft tissues of plankton, but nevertheless undergo only slow release. These elements include

A number of other elements show nutrient type distribution patterns, that is depleted in surface water and enriched in deep water, even though they have no known biological function. These include As (which shows only slight surface depletion), Sc, Ge, Pd, Ag, Cd, Ba, the rare earths, Pt, and Ra. Ba appears to be precipitated by biological mediated reactions (as $BaSO_4$), but the details are not yet understood. Other elements are inadvertently taken up; that is, organisms appear not to be able to discriminate between them and needed elements. Ge provides a well-documented example. It is chemically similar to Si and apparently is taken up by diatoms in place of Si. Like Si, its concentration approaches 0 in surface water and increases with depth (Figure 15.14). Overall, its concentration is strongly correlated with that of Si. Clearly, biological activity controls the distribution of Ge in the oceans.

Comparing the phosphate and silicate profiles in Figure 15.16, we see that concentrations of phosphate increase more rapidly that that of SiO_2 . Essentially all release of phosphate occurs with in the upper 1000 m, and most within the upper 500 m. Maximum Si concentrations are only reached at depths of 1500 to 2000 m. The reason for these different distributions is straightforward. P is released by biological medi-



Figure 15.16. Comparison of labile and refractory nutrients. a. Concentration profiles of two labile nutrients (Cd: points and phosphate: curves) in the North Atlantic and North Pacific. b. Concentration profiles of two refractory nutrients (Zn: points and silicate: curves) in the North Atlantic and North Pacific. Modified from Bruland (1983).

Geochemistry

CHAPTER 15: OCEANS

Cu, Zn, and Fe. Thus whether an element is labile or refractory apparently depends on how readily its organic host molecule is decomposed, as well as whether it is incorporated in hard or soft tissue.

It is interesting to compare the distribution of Si and Ca, both utilized by organisms to build tests. In both cases, the tests have some tendency to dissolve after the organism dies, and thus there is some recycling of these elements back to seawater. But because the flux of Si to the oceans is low, the concentration is relatively low everywhere and the plankton utilizes essentially all of it in surface waters. Plenty of Ca is supplied to the oceans and its concentration is high; organisms utilizing only a portion of that available. The relative variation in Ca concentration is only about 0.5%. Both Si and Ca provide good demonstrations of the lack of control of equilibrium thermodynamics on ocean chemistry, at least on a large scale. Calcium carbonate is oversaturated in surface water, yet precipitates only in biologically mediated reactions. In contrast, the ocean is everywhere undersaturated with respect to opal (though is locally oversaturated with respect to quartz), yet it is biologically precipitated and redissolved only slowly upon death of the organism. The oceans are certainly an example of a kinetically controlled, rather than thermodynamically controlled, system.

PARTICLE-REACTIVE ELEMENTS

The vertical and horizontal distributions of a number of trace elements in seawater, including Al, Ga, Sn, Te, Hg, Pb, Bi, and Th are controlled by abiologic reactions with particles. These reactions include surface adsorption and desorption, and oxidative and reductive dissolution and precipitation. (Even here, however, the biota plays some role, in the production of particles, production of organic molecules that coat particles and affect their surface properties, and in catalyzing oxidation and reduction reactions.) The distribution of Be, Cu, Ti, Cr, Mn, Fe, Zr, the REE, Hf, and Pt are controlled by both biological processes and abiologic reactions with particles. A common characteristic of these all these elements is that they are strongly hydrolyzed in seawater, i.e., at seawater pH they react

with water to form hydroxo complexes (e.g., $Al(OH)_3$, $Al(OH)_4$). Most particles in the ocean have large numbers of O-donor surface groups. Cations that form hydroxo complexes in solution are readily adsorbed to these surfaces. Hence these elements are highly "particle reactive", meaning they readily absorb to particle surfaces.

Erel and Morgan (1991) proposed that the seawater concentrations of these particle-reactive elements are controlled by two opposing processes: adsorption to and transport on inorganic and organic particle surfaces and desorption driven by complexing with dissolved ligands in seawater. Surface adsorption constants provide a measure of the extent of adsorption; complexation constants a measure of the extent of complexation. The data on surface adsorption constants for marine particles is sparse, but Erel and Morgan point out that constants of adsorption of trace metals to O-donor surface groups of oxide, hydrous oxide, and organic surfaces are linearly related to hydrolysis constants (Dzombak and Morel, 1990). In other words, the adsorption constants can be approximated by:

$$\log K_{ads} = a \times \log(\beta_{MOH}^0) + b \qquad 15.26$$

where

For hydrous ferric oxide, *a* has a value of 1.2 and *b* a

 $\beta^{0}_{\rm MOH} = \frac{[M(OH)^{(n-1)+}]}{[M^{n+}][OH^{-}]}$



Figure 15.17. Log of the ratio of concentration in the ocean to the concentration in the upper crust plotted against the log difference between hydroxo and chloride plus carbonate stability constant for 16 trace metals. Both parameters are normalized to the same parameter for Fe. The nutrient (or "recycled") and particle reactive (or "scavenged") elements define separate correlations. These correlations illustrate the strong control exerted by surface and aqueous complexing on elemental concentrations in the ocean. From Erel and Morgan (1991).

Geochemistry

CHAPTER 15: OCEANS

value of -4.4 (Dzombak and Morel, 1990). In this case, then, the ratio of the dissolved to adsorbed concentration of an element should depend on the ratio of the stability constants of dissolved complexes to the first hydrolysis constant. Chloro- and carbonato- complexes are the dominate complexed species in seawater. Hence dissolved to adsorbed ratio can be approximated by:

$$\frac{[TM]_{diss}}{[TM]_{ads}} = \frac{\beta_{MCl}^{0}[Cl^{-}] + \beta_{MCO_{3}}^{0}[CO_{3}^{2^{-}}]}{\beta_{MOH}^{0}10^{b}[surface sites]/[H^{+}]} = \Delta\beta^{Cl+CO_{3}}$$
15.27

Figure 15.17 is a plot of this ratio (normalized to the same ratio for Fe) against the ratio of the seawater concentration of the element to its upper crustal ratio (again normalized to the same ratio for Fe). It shows a linear correlation for the "nutrient elements", indicating adsorption/complexation are indeed strong controls on elemental concentrations in seawater. The plot distinguished the "nutrient" elements from the scavenged, "or "particle-reactive" ones. The latter fall below the correlation because they are adsorbed by carbonate surface groups in addition to OH⁻ surface groups.

Particle-reactive elements show a variety of vertical concentration profiles. Al, Ga, Te, Hg, Pb, Bi,

and Th are typically enriched in surface water and depleted at depth. Lead provides an example (Figure 15.17). In the case of Al, Ga, Sn, Te, and Pb, the surface enrichment in results input wind-born particulates to the surface water followed by partial dissolution; in the case of Th, it results from riverine input (Whitfield and Turner, 1987). These elements are then progressively scavenged by particles. In the case of Sn, Te, Hg, and Pb, a significant part of the aeolian flux may be anthropogenic, hence the surface maxima may not be permanent features of the ocean.

In another commonly seen profile, concentrations increase with depth after first decreasing to a mid-depth minimum (e.g., Al in Figure 15.19). Again, high concentrations in surface waters result from riverine or aeolian inputs while decreasing concentrations with depth result from particle scavenging. Higher concentrations in deep water reflect diffusion out of sediment pore waters. Pore water enrichment occurs as ions are desorbed from surfaces as a result of a decrease in pH in sediment pore water. This in turn is a consequence of respiration and production of CO₂.

A continuous increase in concentration with depth is seen in the P some profiles of Be, Cu, Ti, Zr, and Hf. Such profiles result from From a combination of aeolian and/or riverine surface input, biological uptake in surface waters, scavenging, and input from sediment pore water to bottom water (e.g., Bruland, 1980). The profile of Ti in Figure 15.20 provides and example.

Mid-depth maxima may occur for several reasons. The first is hydrothermal input. Most hydrothermal activity in the oceans occurs on mid-oceans ridges, who depth is typically in the range of 2500-3500 m. Hydrothermal fluids mix with surrounding water and form slightly warm, and therefore slightly buoyant, plumes, which rise hundreds of meters above the ridge crest and are then transported laterally by ocean currents. Elements strongly enriched in hydrothermal fluids will be enriched in these plumes, creating mid-depth maxima for these elements. He is one such element, Mn is another.

Profiles of Mn, and to a lesser degree Fe, can also show maxima associated with the O minimum. These maxima occur because particle-bound Mn and Fe are reduced in O-poor water and the reduced species, Fe²⁺ and Mn²⁺, are much more soluble than the oxidized species. In reducing con-



Figure 15.18. Depth profiles of Pb in the Pacific and Atlantic Oceans. From Schaule and Patterson (1983).



Figure 15.19. Typical vertical concentration profile for Al. After Bruland (1983).

Geochemistry

CHAPTER 15: OCEANS



Figure 15.20. Profiles of dissolved Ti in the Atlantic and Pacific Oceans. From Orians et al. (1990).

ditions then, Fe and Mn can be released from particulates rather than scavenged. Both these can have higher concentrations in basins where deep circulation is so limited that deep waters become anoxic, e.g., the Cariaco trench, some Norwegian fjords, and the Black Sea. Two vertical Mn profiles are shown in Figure 15.21. The maximum at 1000-1500m in the left profile is associated with the O_2 minimum. This profile also shows surface enrichment due to riverine input. River water is less dense than seawater, hence it tends to mix horizontally rather than vertically. The right profile is located over the Galapagos Spreading Center and shows the effect of hydrothermal input on Mn concentrations.

The One-Dimensional Advection-Diffusion Model

Let's examine these concentration-depth profiles in a bit more detail. Concentration profiles such as these can be readily modeled using a *one dimensional advection-diffusion model* (Craig, 1974). The essential assumption of such a model is that the profile observed is a steady-state feature; that is that the variation with depth is the same today as it was, say, 1000 years ago. Let's begin by considering the sim-

ple case of the vertical variation of conservative property of ocean water, such as salinity, between fixed values of salinity at the top and bottom of the water column. Salinity will vary only because of transport of water (by our definition of "conservative", chemical and biological processes have no effect). Two kinds of transport are of interest: turbulent transport and vertical velocity of the water. Turbulent transport is also known as "eddy diffusion". Its is exactly analogous to chemical diffusion



Figure 15.21. Profiles of Mn in the North Central Pacific (left) and over the Galapagos vent areas (right).



Geochemistry

CHAPTER 15: OCEANS

and may be described by the equation:

$$\frac{\partial c}{\partial t} = K \frac{\partial^2 c}{\partial z^2}$$
 15.28

where c is a concentration (such as salinity), K is the eddy diffusion coefficient and has units of m^2/yr , and z is depth. Notice that equation 15.28 is identical to Fick's Second Law (equation 5.91) except that we have replaced the chemical diffusion coefficient, D, with K (and x with z; we define z as being positive upward). Adding a term for vertical velocity, we have:

$$\frac{\partial c}{\partial t} = K \frac{\partial^2 c}{\partial z^2} - \omega \frac{\partial c}{\partial z}$$
 15.29

Notice that the velocity term in equation 15.29 is exactly analogous to the one in equation 5.159, which we derived from sediment diagenesis. At steady-state $\partial c / \partial t = 0$, so:

$$K\frac{\partial^2 c}{\partial z^2} = \omega \frac{\partial c}{\partial z}$$
 15.30

This is a second-order differential equation with respect to c, the solution depends on the boundary conditions. These are that c is fixed at $c = c_0$ at the bottom (z = 0) and $c = c_z$ at the surface (z=Z). The solution to this equation is:

$$c(z) = (c_z - c_0)f(z) - c_0$$
 15.31

where

$$f(z) = \frac{\exp\left(\frac{Z\omega}{K}\right) - 1}{\exp\left(\frac{Z\omega}{K}\right) - 1}$$
15.32

Since c(z) is a linear function of f(z), equation 15.31 can be used to test the appropriateness of the one dimensional model. If a truly conservative parameter, such temperature of salinity, is plotted against f(z), a straight line should result. Any deviation from linearity would indicate there is significant horizontal advection and that the one dimensional model is not appropriate. Provided horizontal advection is not occurring, we can use equation 15.31 to determine whether a particular species is conservative or not: any deviation from linearity on a plot of c versus f(z) would indicate non-conservative behavior.

Now let's consider a non-conservative species that is actively scavenged from seawater through surface adsorption on particles. We assume that the adsorption rate is proportional to the concentration, i.e., first order kinetics. The change of concentration with time can then be described as:

$$\frac{\partial c}{\partial t} = K \frac{\partial^2 c}{\partial z^2} - \omega \frac{\partial c}{\partial z} - \psi c$$
15.33

where ψ is the scavenging rate constant, which we assume is constant with depth. The sign of ψ is such that positive ψ corresponds to removal from seawater, i.e., adsorption, negative to desorption. ψ has units of inverse time; the inverse of ψ is known as the *scavenging residence time* and is denoted τ_{ψ} .

At steady-state:
$$K \frac{\partial^2 c}{\partial z^2} = \omega \frac{\partial c}{\partial z} + \psi c$$
 15.34

The solution to this equation, which again depends on the boundary conditions, is:

$$c(z) = c_z F(\psi, z) + c_0 G(\psi, z)$$
 15.35

CHAPTER 15: OCEANS

where:

$$F(\psi,z) = \frac{\exp(-[Z-z]\omega/2K)\sinh(Az\omega/2K)}{\sinh(AZ\omega/2K)}$$
$$G(\psi,z) = \frac{\exp(z\omega/2K)\sinh(A[Z-z]\omega/2K)}{\sinh(AZ\omega/2K)}$$

 $A = (1 + 4K\psi/\omega^2)^{1/2}$

and

Finally, suppose that in addition to these processes, there is a steady production of the dissolved species through biological decomposition of organic matter. The production in that case would be independent of concentration. The rate of change of concentration is then:

$$\frac{\partial c}{\partial t} = K \frac{\partial^2 c}{\partial z^2} - \omega \frac{\partial c}{\partial z} - \psi c + J$$
 15.36

where J is the rate of production. For simplicity, we assume that it is independent of depth. At steady state:

$$K\frac{\partial^2 c}{\partial z^2} - J = \omega \frac{\partial c}{\partial z} + \psi c \qquad 15.37$$

J can also be used to represent "zero-order" removal; zero-order implying that the rate of removal is independent of concentration. The solution to equation 15.37 is given by replacing all concentration terms in 15.35 by $(c - J/\psi)$. For $\psi = 0$, i.e., for no scavenging, this solution is clearly undefined. The solution in that case is:



The value of the vertical velocity, ω , can be determined from ¹⁴C analyses. Typical values are in the range of –3 to –4 m/yr. Once this is known, equation 15.32 can be fit to a conservative parameter (typically temperature or salinity) to determine K. K is typically in the range of 2000 to 3000 m²/yr. With values for K and ω , the values of J and ψ may be determined by finding the best fit of equation 15.38 to the data. Figure 15.22 shows an example of the one dimensional advection-diffusion model applied to copper concentrations from GEOSECS station 345 in the eastern Pacific (22°31' N, 122° 12' W) between 200 m and 3800 m. ω was determined from ¹⁴C distributions and K then determined by fitting potential temperature to the model. The black dashed line shows the expected Cu distribution i f Cu were behaving conservatively. Clearly, Cu concentrations are lower than expected in that case. The red line shows the fit of the model for ψ = -0.0012 and J = 0. The ψ implies a scavenging residence time for Cu of about 830 years.

Sources and Sinks of Dissolved Matter in Seawater

Most elements are present in seawater at concentrations far below their equilibrium solubility. Ca is present in surface water at concentrations above solubility. Clearly then, the composition of seawater is not controlled by thermodynamic solubility (only the concentrations of the inert gases are controlled by equilibrium solubility). Rather, the composition of seawater is kinetically controlled; specifically, it is controlled by the rates at which dissolved matter is added to and removed from seawater. Rivers represent the principal "source" of dissolved solids in seawater and sediments represent the principal "sink". There are, however, a variety of other sources and sinks. These are shown in Figure 15.23, which illustrates the marine geochemical cycle of Mn. For any given element,



Figure 15.22. Cu data from GEOSECS station 345 in the eastern Pacific (Boyle et al., 1977). Dashed line shows the expected distribution for a one dimensional advection-diffusion model if Cu is behaving conservatively (K = $2886 \text{ m}^2/\text{yr}$; $\omega = 3.7 \text{ m/yr}$). Solid red line shows the fit for $\psi = 0.0012 \text{ yr}^{-1}$.

Geochemistry

Geochemistry

CHAPTER 15: OCEANS

one of these sources or sinks might dominant, and some or most of the sources might be negligible. In the following sections, we will discuss each of these sources and sinks. Before we do, however, we introduce a useful concept in marine geochemistry, that of residence time.

Residence Time

An important concept in the chemistry of seawater is that of *residence time* (Barth, 1952; Goldberg and Arrhenius, 1958). Residence time, τ , is defined as the ratio of the mass of an element in the ocean divided by the flux to the ocean, i.e.:

$$\tau = \frac{\mathbf{A}}{\mathbf{d}\mathbf{A}/\mathbf{d}\mathbf{t}}$$
 15.39

where A is the mass of the element of interest, and dA/dt is the flux to seawater. Implicit in the residence time concept is the assumption that *the oceans are in steady state*, that is, the composition does not change with time, so that the flux of an element into seawater must equal the flux out of seawater. Thus it does not matter whether we use the flux into or the flux out of the oceans in equation 15.39.

If river water is the principal source of the element, equation 15.39 can be re-expressed as:

$$\tau = \frac{C_{sw}}{C_{RW}} \times \frac{\text{mass of seawater}}{\text{flux of river water}} = \frac{C_{SW}}{C_{RW}} 3.7 \times 10^4 \text{yr}$$
 15.40

where C_{sw} and C_{RW} are the concentrations in seawater and river water respectively. (For water, these two terms are essentially both equal to one, so equation 15.40 that on average a water molecule goes through the hydrologic cycle once every 37,000 years.) For example, rivers are the principal source of Na in seawater. The concentration of sodium is 3.9 mg/kg in river water and 10.77 k/kg in seawater, so we calculate a residence time of 103 Ma for Na. About half the sodium in river water is derived from cyclic salts; i.e., it has simply been cycled through the hydrologic system. If we don't count this cycling in the residence time, then sodium has an ocean residence time of about 200 Ma. Many elements have several sources and the fluxes from these are poorly known; thus their residence times are



Figure 15.23. Marine geochemistry of Mn, illustrating the range of possible sources and sinks, as well as internal processing, of dissolved material in seawater.

Geochemistry

CHAPTER 15: OCEANS

not well known for all elements. Nevertheless, residence times of elements in seawater clearly vary greatly: from 150 Ma for Cl (300 is cyclic salts are not included) to a few tens of years for Fe.

Figure 15.23 shows ocean residence time of the elements plotted against the ratio of their concentration in seawater to the upper crustal concentration, the principal source of most elements in seawater. The two are strongly correlated. Thus residence time is a good indicator of the enrichment (or depletion) of an element in seawater relative to its upper crustal concentration.

Over the last 2 decades, remarkable progress has been main in defining the composition of seawater. However, the fluxes of many elements to (and from) the ocean remain poorly known. As a result, the residence times of most elements are poorly constrained and may be significantly revised in the future.

THE RIVERINE SOURCE

Overall, rivers are the main source of dissolved salts in the ocean, though they are not the main source for all elements. Current estimates of average concentrations in river water are given in Table 15.1. We discussed the factors that controlled river water composition in Chapter 13. These were the chemical composition of the source rock, climate, topography, and the intensity of weathering in the catchment basin.

As we found in Chapter 13, rivers vary widely in composition. Since there are many rivers and each is in some respect unique, the task of estimating the average composition of rivers and their combined flux to the ocean is not an easy one. The largest 20-25 rivers carry only 15% of the flux to the oceans. Relatively few rivers have been subjected to thorough geochemical investigations; for some trace elements there are few data for any rivers. Furthermore, the composition of many rivers has been disturbed by mining, agricultural, industrial, and other activities, so that their modern compositions may not be representative of the composition in the past. Finally, the composition of most rivers var-



Figure 15.24. Log of seawater residence time plotted against the ratio of seawater to upper crustal concentration. Modified from Taylor and McLennan (1985).

ies with river flow. Thus characterizing the composition of a river requires many measurements made over the course of a year or more. For all these reasons, there is considerable uncertainty in the concentrations listed in Table 15.1.

Estuaries

Even if the compositions of rivers were well known, the task of determining the riverine flux to the sea would still be a very difficult one. The reason for this is that estuaries act as flow-through chemical reactors, in which dissolved components are added and removed, and what flows out is not the same as what flows in. This is a consequence of changes in solution chemistry that occur when river and sea water mix.

An estuary is that portion of a river into which coastal seawater is carried by tidal forces; estuaries are thus zones of mixing

Geochemistry

CHAPTER 15: OCEANS

of fresh and sea water. They may be lagoons behind barrier islands (e.g., Pamlico Sound, North Carolina), river deltas (e.g., the Rhine Delta), drowned river valleys (e.g., the Gironde Estuary, France), tectonic depressions (e.g., San Francisco Bay), or fjords (Sannich Inlet, British Columbia). In the typical estuary, there is a downstream flow of fresh water at the surface and an upstream flow of seawater at depth; the fresh water overlies the salt water because of its lower density. Depending on the geometry of the estuary, the strength of the tides, and the strength of the river flow, these two layers will mix to varying degrees. Low river flow and strong tides produce a "well mixed" estuary in which there is little vertical gradient in salinity; strong river flow and weak tides lead to a "salt wedge" estuary in which a strong pycnocline develops at the interface between the layers.

The principal chemical changes that can occur in estuaries are as follows.

- changes in ionic strength,
- changes in the concentrations of major cations, which affects speciation of minor components,
- changes in pH,
- changes in the concentration and nature of suspended matter, and
- changes in the redox state within the water and sediment.
- In the following paragraphs, we consider these changes in greater detail.

As seawater mixes with river water, the resulting increase in ionic strength and change in pH and solute composition induces the flocculation of riverine colloids which dramatically affects trace metal chemistry. As a practical matter, aquatic and marine chemists often consider "particulates" to be the material retained on 0.4 μ m filters and anything passing through such a filter to be "dissolved". Unfortunately, nature is not so neat: in reality there is a continuum between truly dissolved substances and readily recognizable particles. Materials between these are termed *colloids*. Colloids are often defined as having sizes in the range of 10⁻⁹ to 10⁻⁶ m and may be separated by special techniques such as ultrafiltration and membrane filtration. Entities at the small end of this range would consist of ~10³ or fewer atoms and approach the size of larger humic acids, while particles at the large end of the range would be retained on filter paper. Colloids are a surprisingly important component of natural waters. Most of the Fe^{III}, as well as several other readily hydrolyzed metals such as the rare earths, are present as colloids of hydrous ferric oxide and metal humates rather than true dissolved species (e.g., Boyle et al., 1977). A significant fraction of humic acids is probably also colloidal (Sholkovitz, et al., 1978).

The stability of colloids depends strongly on their surface charge. Because they are small, colloids have high surface area to volume ratios. As a result, individual colloids settle very slowly and may remain in suspension indefinitely, behaving essentially as dissolved species. When colloids coagulate, however, the surface area to volume ratio decreases, allowing them to settle out of suspension. Whether or not they coagulate depends on the balance of forces acting between individual particles. As is the case for larger particles, colloids typically have a surface charge, which is balanced by the ions adjacent the surface, the electric double layer (Chapter 6). The double layer produces a repulsive force between particles. The thickness of the double layer is inversely related to the square root of the ionic strength of the solution (equation 6.116). Countering this repulsion is an attraction due to van der Waals interactions. The strength of the van der Waals interactions decreases, as it does when sea water mixes with river water, the diffuse outer layer, or Gouy layer, is compressed, allowing individual particles to approach each other more closely. Once they approach within a critical distance, the van der Waals attraction binds them together. The process eventually produces particles large enough to settle out, a process called *flocculation*.

Surface charge is affected by adsorption of ions to surface sites; it therefore depends on solution chemistry. Surface charge is also pH dependent (e.g., Figure 12.35). Thus changes in pH and solute chemistry during estuarine mixing may also promote coagulation. Eckart and Sholkovitz (1976) and Boyle et al. (1977) found that flocculation of humic acids occurs when Ca and Mg bind to carboxylic acid groups, neutralizing their negative surface charge. This allows them to coagulate and precipitate.

Geochemistry

CHAPTER 15: OCEANS

In a series of experiments, Sholkovitz (1976, 1978) demonstrated that mixing of sea and water from the River Luce of Scotland results in removal of 95% of Fe and lesser amounts of humic acids, Cu, Ni,

Mn, Al, Co, and Cu by colloid flocculation. Boyle et al. (1977) demonstrated that this is a general phenomenon occurring in many estuaries and estimated that 90% of the riverine flux of Fe is removed in this way. The exact nature of colloidal iron in rivers and its precipitation in estuaries is only partly understood. Boyle et al. (1977) point out that the iron-carbon ratio of flocculated material is too high for all Fe to be organically complexed. They suggested instead that the hydrous ferric oxide forms a stable colloidal dispersion with humic acids in rivers.

Subsequent studies revealed that other elements, such as the rare earths, are also removed by this process. For example, Hoyle et al. (1984) found that 95% of the heavy rare earths and 65% of the light rare earths dissolved in river water were removed by flocculating colloids in the estuary of the River Luce, a river particularly rich in organic matter. Loss of rare earths as a function of salinity was similar to that of iron, as shown in Figure 15.25, so that Hoyle et al. concluded that

the rare earths were adsorbed or coprecipitated with Fe-organic colloids. Organic matter colloids appear to be the carriers of rare earths as no REE removal was observed in laboratory experiments in which seawater was mixed with organic-poor river water. Hoyle et al. (1984) found that heavy rare earths were preferentially removed. Subsequent work showed that in most cases, it is the light rare earths that are preferentially associated with colloidal phases and removed by flocculation (Elderfield, et al., 1990; Sholkovitz, 1992).

Flocculation is not restricted to colloids. The electric double layer surrounding larger particles is also condensed as river water mixes with seawater, causing them to flocculate as well. As a result of this process, estuary act as sediment traps. As much as 90% of the suspended load of rivers never reaches the open sea, being deposited in the estuarine or near-shore environment instead (Chester, 1991).

Increasing ionic strength also results in a decrease in activity coefficients, as would be predicted, for example, by the Debye-Hückel equation (equation 3.84). This decreases the *effective* concentration of dissolved species and thus drives precipitation-dissolution and adsorption-desorption reactions toward dissolution and desorption. Mixing of river and seawater also results in changing concentration of major ions. This af-



Figure 15.25. Removal of Fe and REE during mixing of seawater and organic-rich water from the River Luce in Scotland. Removal apparently occurs by coagulation and precipitation of Fe-organic colloids. Modified from Hoyle et al. (1984).



Figure 15.26. Ni speciation as a function of salinity in estuaries. From Montoura et al. (1978).

Geochemistry

CHAPTER 15: OCEANS

fects speciation of trace ions in several ways. First, as the proportion of seawater increases, trace metals are replaced on cation binding sites of organic molecules such as humic and fulvic acids by Mg and Ca. Second, as the concentrations of inorganic ligands increase, metals ions are increasingly complexed by them. These changes are illustrated for nickel in Figure 15.26. The result is a decrease in the free ion concentration of many metals, which in turn results in desorption of surface-bound species. Thus as a result of changing speciation, as well as the increase in ionic strength,



Figure 15.27. Cartoon illustrating the use of concentration-salinity diagrams in estuaries.

the total dissolved concentration of some ions may increase in an estuary. Changes in pH, of course, will also affect speciation.

Whether a dissolved component is added to or removed from solution in an estuary can be readily determined by plotting its concentration against salinity. The concentration of any *conservative* species in an estuary will be a function only of the proportions in which sea and river water are mixed. Salinity is always conservative in estuaries because addition or removal of major species does not occur to a significant extent. The concentration of any other conservative species should therefore be a simple linear function of salinity. As illustrated in Figure 15.27, any deviation from linearity indicates removal or addition of that species.

The behavior of barium in the Yangtze estuary provides an example of addition demonstrated by a concentration-salinity diagram (Figure 15.28). The addition of Ba probably results from replacement



Figure 4.28. Ba concentrations as a function of salinity in the estuary of the Yangtze River compared with a conservative river-seawater mixing line. B a is being added to solution at salinities of about 2 and 14. After Edmond et al. (1985).

of adsorbed Ba on surfaces by major cations as we described above (Edmond et al., 1985). Figure 15.29 shows concentration-salinity diagrams for metal ions in the Gironde estuary, France. Fe concentrations plot along a curve below the conservative mixing line, consistent with its removal by colloid flocculation. Ni plots along a curve above the mixing line, indicating addition. Kraepiel et al. (1997) concluded this results from desorption of Ni from sediment particle surfaces, which in turn is due to a decrease in the free Ni ion concentration due to complexation by inorganic ligands. Essentially all of the Ni addition occurs at salinities lower than 10%, at high salinities Ni behave nearly conservatively. Pb appears to be nearly conservative. Pb is strongly particle-reactive, so truly conservative behavior would be surprising. Its apparent conservative behavior could be due to a combination

Geochemistry

CHAPTER 15: OCEANS



Figure 15.29. Concentrations of dissolved Pb, Ni, and Fe measured in the Gironde estuary, Bordeaux, France. Open symbol is the average concentration of river water. Dashed line is the conservative mixing line, solid line is visual best fit to the data. The curves indicate removal of Fe and Pb and addition of Ni. Modified from Kraepiel et al. (1997).

of addition by desorption from solid surfaces and removal by coprecipitation with colloidal Fe hydroxides and humates.

Redox reactions in estuaries also affect solution chemistry. Increasing pH speeds the oxidation of ferrous to ferric iron, while the presence of suspended matter, increases the rate of oxidation of Mn²⁺, which may be present metastably in rivers. In some estuaries, the development of a strong pycnocline inhibits exchange of gas between the deeper saline layer and the atmosphere. Many estuaries are surrounded by wetlands (salt marshes or wetlands), which export both dissolved and particulate organic matter. Bacterial respiration of this organic matter can lead to anoxic conditions in the deep layer. Anoxia may also develop when algal growth is stimulated by anthropogenic additions of nitrate and phosphate and bacterial decomposition of the algal remains occurs. Chesapeake Bay is a good example of this situation; the bottom waters of which become anoxic in the summer. In such circumstances, species whose solubility depends on their redox state, most notably Fe and Mn, may redissolve in the anoxic water.

Where anoxia does not develop in the water column, it may still develop within the sediment because of high organic sedimentation rates. Fe and Mn may be reduced in the sediment and diffuse back into the overlying water column. Hence reduced sediments may serve as a source of metals ions. A study of Narragansett Bay, Rhode Island by Elderfield et al. (1981) provides an example. Although the bay water is oxygenated, oxygen is rapidly depleted within the uppermost sediment due to the high organic content. As a result, Mn⁴⁺ in

the sediment is reduced to soluble Mn^{2+} , the concentration of which increases to 2 to 7 μ M within the top cm of sediment, well above the 0.5 μ M concentrations found in Bay water. Remobilized Mn then diffuses out of the sediment into Bay water. The resulting benthic flux of Mn into the Bay is 40% of the riverine flux. In contrast, the reducing conditions within the sediment lead to precipitation of Ni, Cd, and Cu sulfides, so that there is a net diffusive flux of these elements from Bay water to the sediment. The Ni and Cu fluxes into the sediment were 60% and 30% of the riverine fluxes respectively.

Reductive dissolution can also remobilize particle-reactive elements scavenged by Mn and Fe oxides and hydroxides. This apparently occurs in Buzzards Bay, Massachusetts, where the sediment becomes reducing at depth. Elderfield and Sholkovitz (1987) found rare earth concentrations in sediment pore waters that were 10 to 50 times local seawater values. Much of the remobilized REE are

Geochemistry

CHAPTER 15: OCEANS

readsorbed or reprecipitated by Mn and Fe oxides in the upper sediment, but Elderfield and Sholkovitz (1987) nevertheless concluded that benthic flux of REE to the bay was of a magnitude similar to the riverine flux.

In summary, chemical processes in estuaries significantly modify the riverine flux to the oceans. This modification occurs primarily through solution-particle reactions that occur as a result of mixing of sea and river water. The results of these processes differ between estuaries because of differences in physical regime, biological productivity, residence time of the water in the estuary, and the nature and concentration of suspended matter. Some generalizations may nevertheless be made. A significant fraction of Fe and other particle-reactive elements such as Al and the REE are removed by colloidal flocculation in estuaries. Though biological productivity removes nutrients (SiO₂, NO₃, PO₄) from solution, these are largely recycled within the estuary so that the net modification to the flux of these elements is probably small (Chester, 1991). The flux of major cations is similarly largely unaffected. Estuaries may act as a source for a number of other elements, such as Ba, Mn, Ni, and Cd through desorption and remobilization in the sediment.

Mid-Ocean Ridge Hydrothermal Systems

One of the most exciting developments in geochemistry in the past 25 years has been the discovery of hydrothermal vents at mid-ocean ridges. Simply the sight of 350° C water, black with precipitate, jetting out of the ocean bottom, surrounded by a vibrant if bizarre community of organisms living in total darkness at depths of 2500 m or more was exciting. But these phenomena were exciting for other reasons as well. Hydrothermal systems are sites of active ore deposition, so scientists were able to directly analyze the kinds of fluids that produce volcanogenic massive sulfide ores. Hydrothermal activity is also an important source for some elements in the oceans, and an important sink for others and has a profound effect on the composition of the oceanic crust. Thus the discovery of hydrothermal vents has provided geochemists with the opportunity to put into place a major piece of the great geochemical puzzle.

The fluids emanating at hydrothermal vents are seawater that has undergone extensive reaction at a variety of temperatures with the oceanic crust. They are reduced (sulfide-bearing), acidic, and rich in dissolved metals. Three kinds of venting has been observed: low temperature diffuse venting, in which hydrothermal solutions that have mixed extensively with seawater solutions in the subsurface diffuse slowly out of the seafloor; "black smokers", in which high temperature (usually >300° C) fluid jets from sulfide "chimneys" and precipitate sulfide and Fe-Mn oxyhydroxide "smoke"; and "white smokers", in which high temperature fluids (200-300°C) jet from anhydrite chimneys and precipitate white anhydrite "smoke".

The Composition of Hydrothermal Fluids

Samples of pure vent fluids have proven difficult to obtain, as vent fluids quickly mix with ambient seawater. The pure vent fluid end-member of the sampled mixture must therefore be calculated. This is straightforward provided the concentration of at least one property of the vent fluid is known. Since the temperature of the vent fluid can be determined, this provides the key to calculating the vent fluid end-member composition. The first vents discovered, on the Galapagos Spreading Center, were diffuse, low temperature vents (<13 ° C). A strong inverse correlation between Mg and temperature was observed, and Edmond, et al. (1979), concluded that the pure hydrothermal fluid had a temperature of 350° C and a Mg concentration of 0. The data from the first high temperature vents discovered, at 21 °N on the East Pacific Rise, extrapolated to a similar temperature and an Mg concentration of 0 (Figure 15.30a; the scatter in the data in this figure is due to the temperature probe flopping about: temperatures were high enough to melt the adhesive holding it in place). Thus Mg appears to be quantitatively extracted from seawater in hydrothermal systems. This was subsequently shown to be true of all other high temperature hydrothermal systems. Having determined that the Mg concentration of the vent fluid is 0, the concentrations of all other species in the vent fluid were easily obtained from the intercept of a plot of the concentration of the species of interest. For example, a plot of sulfate versus Mg extrapolates to 0 sulfate at 0 Mg (Figure 15.30b). Thus these vent fluids also have

Geochemistry

CHAPTER 15: OCEANS

0 sulfate. The same procedure shows the vent waters are also rich in silica and Li relative to seawa-ter (Figure 15.31).

The pure hydrothermal fluid in the first few vents discovered all had relatively homogeneous compositions and similar temperatures. Subsequently discovered vents were more variable. Temperatures of fluids from vents at 19 sites ranges from 220 to 403°C (this range excludes "diffuse flow"). The lowest temperatures occur where vent fluids exist through sediment overlying the basalt (e.g., the Guaymas Basin in the Gulf of California, Escanaba Trough on the Gorda Ridge, and Middle Valley on the Juan de Fuca Ridge). The highest temperatures were found in vents at the site of a recent volcanic eruption (at 9-10° N on the East Pacific Rise). The majority of vents have temperatures in the range of 300°C to 380°C, a surprisingly narrow range. This narrow range of temperatures probably reflects the large density decrease that occurs when seawater is heated beyond these temperatures. Alternatively, it may reflect a sharp decrease in rock porosity beyond these temperatures (L. Cathles, pers. comm.).





Figure 15.30. (a) Temperature vs. Mg concentration for several vents in the 21°N region of the EPR. (b) Mg-sulfate plot for 21°N vents. From Van Damm et al. (1985).

Figure 15.31. (a) SiO_2 vs. Mg for hydrothermal fluids from the 21°N vents. (b) Mg-Li plot for 21°N vents. From Van Damm et al. (1985).

CHAPTER 15: OCEANS

	-					
	21°N	Escanaba	South Cleft	Axial Volcano	TAG	Seawater
	(EPR)	(Gorda)	Plume	Virgin Mound	MAR	
T°C	273-355	108-217	224	299	335-350	—
рН	3.3-3.8	5.4	3.2	4.4	3.7-3.9	8.2
Li µmol/kg	891-1322	1286	1718	184	845	25.6
δ ⁶ Li	-6.6 to -10				-6.3 to -8.5	-32.3
Be nmol/kg	10-13		95		38.3	0.025
B μmol/kg	500-548	1.71-2.16	496	450	518-530	406
$\delta^{11}B$	30.0-32.7	10.1-11.5	34.2		25.6-26.8	+39.5
$CO_2 mmol/kg$	5.7		3.7-4.5	285		
CH ₄ mmol/kg	0.06-0.09		.08209			
NH ₄ mmol/kg	<.01	5.6				—
Na mmol/kg	432-513	560	796	148	510	468
Al μ mol/kg	4.0-5.2				5.0-5.3	0-0.15
Si mmol/kg	15.6-19.5	5.6-6.9	23.3	13.5	18.3	0-0.25
$H_2S mmol/kg$	6.6-8.4	1.1-1.5	3.5	18	5.9	26.9*
$\delta^{34}S$	+1.4-3.4	+7.8	+5.7	+7.3		+21
Cl mmol/kg	489-579	668	1087	176	559	532
K mmol/kg	32.5-49.2	34-40.4	51.6	6.98	23.8	9.96
Ca mmol/kg	11.7-20.8	33.4	96.4	10.2	9.9-10.5	10.3
Mn mmol/kg	.67-1.0	0.01-0.21	3.59	142	659	< 0.04
Fe mmol/kg	.75-2.43	0-0.1	18.7	12	1	< 0.006
Conmol/kg			200			< 0.07
$Cu \mu mol/kg$			1.5	0.4	1.64	< 0.004
$Zn \mu mol/kg$			780	2.2		< 0.009
Ge nmol/kg	130-170		150-260			0-0.2
As nmol/kg	30-452					13-27
Se nmol/kg	<0.6-72		<1			0.5-1.5
Br μ mol/kg	802-929	1179	1832	250	847	839
Rb μ mol/kg	27-33	80-105	37		10.7	1.45
Sr μ mol/kg	65-97	209	312	46	51	87
⁸⁷ Sr/ ⁸⁶ Sr	.70307033	0.7099			0.7028	0.70918
Monmol/kg			6			115
Ag nmol/kg			120			< 0.03
Cd nmol/kg			910			0-1.0
Sb nmol/kg			18			0.7-1.3
I μmol/kg		99				0.2-0.5
Cs nmol/kg	202	6.0-7.7			179	2.25
Ba μ mol/kg	8-16					0.085
Tl nmol/kg			110			< 0.08
Pb nmol/kg			1630			< 0.002

TABLE 15.4. COMPOSITION OF REPRESENTATIVE HYDROTHERMAL VENT Fluids

From compilations of van Damm (1995) and Shanks et al. (1995). EPR: East Pacific Rise, Mar: Mid-Atlantic Ridge. South Cleft and Axial Volcano sites are on the Juan de Fuca Ridge.

* Concentration of *sulfate* in seawater.

Table 15.4 lists compositions of several representative hydrothermal vent fluids (all compositions are calculated on the assumption that the pure vent fluid has an Mg concentration of 0 as described above). The fluids are acidic and reducing (sulfide has replaced sulfate). Most of the vent fluids

Geochemistry

CHAPTER 15: OCEANS

have Cl, Na, and B concentrations that are close (within about 20%) to seawater concentrations. In these cases, these elements appear to behave almost conservatively in the hydrothermal system. The Virgin Mound vent on Axial volcano, a "white smoker", has substantially less Cl than seawater, while the South Cleft fluid has substantially more. In general, the vent fluids have substantially less sulfide than seawater has sulfate, thus there has been a net loss of sulfur from solution. All the fluids have much greater concentrations of Li, Be, Al, C, Si, Ge, Rb, Cs, and transition metals (except for Mo) than seawater. All except the Virgin Mound vent higher concentrations of K than seawater.

The composition of hydrothermal fluids depends on a number of factors (Von Damm, 1990): temperature, pressure, the ratio in which water and rock react (the water/rock ratio), whether the rock is fresh or has previously reacted, whether equilibrium is achieved, or compositions are instead controlled by kinetics, whether fluids have undergone phase separation, and whether mixing between different fluids has occurred. We now examine the processes that occur within hydrothermal systems.

Evolution of Hydrothermal Fluids

In a simplistic fashion, hydrothermal systems can be divided into 3 zones (Alt, 1995). The principal processes and reactions involved are shown schematically in Figure 15.32. The first is the *recharge zone*, where seawater enters the oceanic crust and is heated as it penetrates downward. In this zone, which may be several kilometers or more off the axis of the ridge, flow is diffuse, so water rock ratios are low, and temperature of reaction is relatively low (<200° C). The second is the *reaction zone*, which is located near the base of the sheeted dikes and in the upper gabbros (see Chapter 11). In this zone, water reacts with hot rock at high temperature (\geq 350° C), and the primary chemical characteristics of the fluid are determined. Because of reduced permeability of the crust, wa-

ter/rock ratios are lower here than in the overlying recharge zone. Phase separation, in which the fluid separates into a high density brine and lower density fluid, may occur in this zone. The third zone, in which the final composition of the fluid is achieved, is the *upflow* zone. Here water rises rapidly and cools somewhat as it does so. Precipitation of some sulfides may occur. Boiling may also occur if pressure is sufficiently low and temperature sufficiently high. Since upflow appears to be quite focused, water/rock ratios are low. Mixing with low temperature seawater may also occur, which leads to extensive sulfide and anhydrite precipitation.

Recharge Zone

As water enters the oceanic crust, low temperature reactions occur in which volcanic



Figure 15.32. Cartoon illustrating some of the important reactions occurring in mid-ocean ridge hydrothermal systems.

Geochemistry

CHAPTER 15: OCEANS

glass and minerals are transformed to clays such as celadonite, nontronite, ferric micas, and smectite, and oxyhydroxides. In this process, the rock takes up alkalis Li, K, Rb, Cs, B, and U from seawater. Calcite may precipitate in veins. Isotopic exchange also occurs, leading to higher Sr and O and lower B and Li isotope ratios in the basalt.

As seawater warms to temperatures around 150° C, anhydrite (CaSO₄) precipitates. By 200° C, essentially all the Ca²⁺ and two-thirds of the SO₄²⁻ (as well as a significant fraction of the Sr) are lost in this way. Addition of Ca²⁺ to the fluid by reaction with basalt will result in further removal of sulfate. Thus the fluid entering the reaction zone is severely depleted in calcium and sulfate. Although anhydrite is found in altered oceanic crust, it is rare. Thus it is likely that much of the anhydrite precipitated in this way later dissolves when the crust cools.

The third major reaction in the recharge zone is loss of Mg²⁺ from seawater to the oceanic crust. This occurs through replacement of primary igneous minerals and glass by clay minerals such as sapolite and smectites, for example, by replacing plagioclase:

 $2Mg^{2+} + 4H_2O + 2SiO_2 + 2CaAl_2Si_2O_8 \rightleftharpoons Mg_2Ca_2Al_4Si_6O_{20}(OH)_4 + 4H^+$ At higher temperatures, chlorite forms, for example, by replacing pyroxene:

 $2Mg^{2+} + 6H_2O + 2MgFeSi_2O_6 \rightleftharpoons Mg_4Fe_2Si_4O_{10}(OH)_8 + 4H^+$

The significance of these reactions in not only loss of Mg from the solution, but also the production of H^+ (or, equivalently, the consumption of OH^-). It is these reactions that account, in part, for the low *pH of hydrothermal vent solutions*. This greatly increases the fluid's capacity to leach and transport metals.

Reaction Zone

Seismic studies of the East Pacific Rise reveal the existence of a "melt lens" at shallow depth (2 km or so) beneath the rise axis. The depth of this magma lens is an upper limit to the depth to which hydrothermal solutions may circulate. This implies pressures in the reaction zone are less than 50 MPa. Geochemical observations are consistent with these geophysical constraints. For instance, Von Damm and Bischoff (1987) used measured SiO₂ concentrations in Juan de Fuca vent fluids together with thermodynamic data to estimate that the fluids equilibrated with quartz at pressures of 46-48 MPa and temperatures of 390-410° C. Geothermometry performed on minerals in altered oceanic crust indicate temperatures as high as 400-500° C. By comparing a thermodynamic model of hydrothermal interactions and assuming fluids are in equilibrium with the assemblage anhydrite-plagioclase-epidote-pyrite-magnetite, Seyfried and Ding (1995) estimated temperatures of 370° to 385° C and 30 to 40 MPa for equilibration of fluids from the 21° N on the EPR and the MARK area of the Mid-Atlantic Ridge.

Under these conditions, reactions would include the formation of amphiboles, talc, actinolite, and other hydrous silicates from reactions involving ferromagnesian silicates (olivines and pyroxenes), the formation of epidote from plagioclase:

$$Ca^{2+} + 2H_2O + 3CaAl_2Si_2O_8 \rightleftharpoons 2Ca_2Al_3Si_3O_{12}(OH) + 2H^+$$

as well as the exchange of Na^+ for Ca^{2+} in plagioclase, a process termed *albitization*, and precipitation of quartz:

 $2Na^{+} + 2CaAl_2Si_2O_8 \rightleftharpoons 2NaAlSi_3O_8 + SiO_2 + 2Ca^{2+}$

The evidence for albitization comes not only from the identification of albitized plagioclase in hydrothermally altered rocks, but also the inverse correlation between Na/Cl and Ca/Cl in hydrothermal fluids (Figure 15.33). In addition, the fluid will be reduced by oxidation of ferrous iron in the rock, e.g.:

$$2SO_4^{2-} + 4H^+ + 11Fe_2SiO_4 \rightleftharpoons FeS_2 + 7Fe_3O_4 + 11SiO_2 + 2H_2O$$

The solubility of transition metals and S increase substantially at temperatures above 350°C, so sulfides in the rock are dissolved, e.g.:

 $Cl^{2-} + 2H^+ + FeS \rightleftharpoons H_2S_{aq} + FeCl_{aq}^0$

The metals released will be essentially completely complexed by chloride, which is by far the dominant anion in the solution, as most sulfate has been removed or reduced and sulfide and CO_2 will be

Geochemistry

CHAPTER 15: OCEANS



Figure 15.33. Na/Cl vs Ca/Cl in hydrothermal vent fluids. Most fluids define an inverse correlation that results from albitization of plagioclase. Only the fluids from vents in the 9-10°N region of the EPR that developed after the 1991 eruption (open symbols) deviate from this trend. Composition of seawater indicated by the star. After Von Damm (1995).

largely protonated at the prevailing pH. Isotopic compositions of H_2S in vent fluids indicates most is derived from dissolution of sulfides in the rock, with a smaller contribution from reduction of seawater sulfate.

Whereas the alkalis Li, K, Rb, Cs, B, and U are taken up by the rock at low temperature, they are released at high temperature. Loss of K, Rb, and Cs begins around 150° C, but loss of Li probably does not begin until higher temperatures are reached (Na, though, is actively taken up, even at high temperatures, by albitization).

Fluids make their closest approach to the magma chamber in the reaction zone, and magmatic volatiles may be added to fluids within this zone. Hydrothermal vent fluids with seawater chlorinities have CO_2 concentrations as high as 18 mmol/kg, which is substantially more than seawater (~2 mmol/kg). The isotopic composition of this carbon ($\delta^{13}C \approx -4\%$ to -10‰) is similar to that of mantle carbon (see Chapter 9) and distinct from that of seawater bicarbonate ($\delta^{13}C \approx 0$). Thus the

excess CO_2 is probably of magmatic origin. Other magmatic volatiles present in the fluid may include He, CH_4 , H_2 , and even H_2O . In most cases, any contribution of magmatic H_2O will be insignificant compared to seawater-derived H_2O . However, fluids from vents at 9-10°N on the EPR have negative δD values (Figure 15.34), which suggest a small but significant contribution of magmatic water. These vents developed and were sampled shortly after an eruption in 1991. Shanks et al. (1995) calculated that the observed δD values could be explained by addition of 3% magmatic water and that this water could be supplied by degassing of a dike 20 km long, 1.5 km deep and 1 m wide. The magmatic water would be exhausted in about 3 years.

The reaction zone is also the region where phase separation is most likely to occur. Unlike pure water, which cannot boil at pressures above its critical point (see Chapter 2), seawater will undergo phase separation above its critical point, but the two phases produced are different from those produced below the critical point. Below its critical point, at 29.8 MPa and 407° C, seawater boils to produce a low salinity vapor phase and a liquid whose salinity initially approximates that of the original liquid. In the case of hydrothermal fluids, the vapor produced would be strongly enriched in H_2S and CO_2 as well as other volatiles. As boiling continues, the liquid becomes increasingly saline. Above its critical point, seawater separates into a dense brine and a fluid whose salinity initially approximates that of the original liquid. As phase separation continues, the fluid becomes increasingly dilute while the brine becomes more concentrated. The phase diagram (P-X) for the system H_2O -NaCl shown in Figure 15.35 illustrates this. Seawater behaves approximately as 3.5% NaCl solution. At a pressure of 36.4 MPa and 400° C, a 3.5% NaCl solution would be above the two-phase curve, so only one phase exists. At the same pressure and 430° C, it lies just on the two-phase curve, and a brine containing ~10% NaCl begins to separate. If temperature is increased to 450°, NaCl in the brine increases to ~20%, and decreases in the other phase to 0.4%.

Cl behaves nearly conservatively in hydrothermal solutions. Phase separation, and mixing between the fluids produced by it, provides the best explanation for the large variations in Cl content

Geochemistry

CHAPTER 15: OCEANS

observed in hydrothermal fluids. Brines produced by phase separation may be too dense to rise to the seafloor. Instead, they may reside at depth for prolonged periods of time, slowly mixing with less saline fluids. The high chlorinity of the North Cleft fluid results from mixing of a fluid with seawater chlorinity and a high salinity brine (Von Damm, 1988). Indeed, most hydrothermal vent fluids appear to be mixtures of 3 components: hydrothermally altered seawater, a high chlorinity brine produced by supercritical phase separation, and a low chlorinity, H₂S-rich vapor produced by subcritical phase separation (Edmonds and Edmond, 1995). The chlorinity of fluids also influences other compositional factors. The solubility of H₂S decreases with increasing chlorinity. However, concentrations of metals such as Fe and Cu increase with increasing chlorinity because of the formation of metal chloride complexes. Thus a fluid produced by mixing will likely be out of equilibrium with the rock and thus undergo further reaction after mixing (Seyfried and Ding, 1995).

Two other important controls on solution chemistry are pH and f_{O_2} . The Fe/Cu ratio of the fluid is sensitive to both of these, high f_{O_2} and low pH favoring a high Fe/Cu ratio. Fe/Cu ratios of hydrothermal fluids indicate pH values in the reaction zones of 4.8 to 5.2 and f_{O_2} buffered by the assemblage anhydrite-magnetite-pyrite (Seyfried and Ding, 1995).

Upflow Zone

The density decrease caused by heating eventually forces the hydrothermal fluid to rise to the seafloor. The concentrated flow out of vents indicates that the upwelling zone is narrow and flow is strongly focused through fractures. Upflow zones in exposed sections of oceanic crust (ophiolites) are altered to the assemblage epidote-quartz-titanite or actinolite-albite-titanite-chlorite assemblages. This is consistent with thermodynamic calculations, which show hydrothermal fluids in equilibrium with similar assemblages (e.g., Bowers et al., 1988). The fluid experiences decompression as it rises, and may experience phase separation at this point. The low chlorinity Virgin Mound fluid (Table 15.5) is an example of a fluid whose chlorinity has been reduced by mixing with a low salinity vapor phase produced by boiling. The fluid is also very strongly enriched in CO_2 , as CO_2 partitions into the vapor during boiling, and this accounts for the high CO_2 of this fluid. The shallow depth of this vent

would mean that hydrothermal fluids would reach the critical point nearly 1500 m beneath the seafloor, providing ample opportunity for subcritical phase separation during ascent.

Most fluids appear to have undergone some conductive cooling during ascent, as calculated equilibrium temperatures generally somewhat exceed measured vent temperatures. Cooling of fluids will induce precipitation of sulfides and quartz. In ophiolites, upflow zones are marked by mineralized alteration pipes, or stockworks. The solubility of Cu shows the strongest temperature dependence, followed by Fe, thus the concentrations of these two elements may drop significantly during upflow. The solubility of Mn and Zn are temperature dependent.

In the Guaymas Basin of the Gulf of California and the Escanaba Trough of the southern Gorda Ridge, vent fluids exit through sediment cover. Hydrothermal fluids must traverse up to 500 m of sediment before exiting to seawater. Vent temperatures are somewhat cooler, 100-315°C. The compositions of these fluids is distinct. The fluid alkaline because of the presence of ammonia produced by break-



Figure 15.34. Oxygen and hydrogen isotope ratios in hydrothermal vent fluids. Red arrows show predicted paths for various subseafloor processes (phase separation refers to the vapor path during adiabatic phase separation). Only the fluids from the 9-10° N vents (open circles) and Guaymas Basin fluids do not conform to the predicted path for fluid-basalt interaction. Negative values in the 0-10° N fluids indicate mixing with magmatic water. Modified from Shanks et al. (1995).

Geochemistry

CHAPTER 15: OCEANS



Figure 15.35. Pressure-composition phase diagram for the system H_2O —NaCl. A seawater-like 3.5% NaCl solution at 36.4 MPa (cross) will lie above the 2 phase region at 400° C, and will thus consist of a single phase. At 430° C, it lies just on the two-phase curve and a brine of composition A begins to separate. At 450°, it lies within the 2-phase field and has separated into a fluid of composition C and a brine of composition B. Adapted from Bischoff and Pitzer (1989).

down of organic matter in the sediment. Guaymas Basin fluids are somewhat richer in some alkalis and alkaline earths, due to dissolution of carbonate and leaching of sediment. Both the Guaymas and Escanaba fluids are transition metal-poor, as a result of sulfide precipitation in the sediment. The Guaymas fluid is also rich in hydrocarbons, which are produced by thermal degradation of organic matter in the sediment.

Finally, hydrothermal fluids eventually mix with seawater, either in the shallow subsurface or as they exit the seafloor. This induces additional cooling and precipitation. Along with sulfide precipitation, mixing causes the seawater-derived sulfate to precipitate as anhydrite. Precipitation of anhydrite accounts for the white "smoke" of white smokers (of which the Virgin Mound vent is an example). Precipitation at the sea surface quickly builds chimneys, which can reach more than 10 meters above the seafloor. Chimneys consist primarily of Fe and Cu sulfides such as pyrite (FeS₂), marcasite (FeS₂), pyrrohotite (FeS), chalcopyrite (CuFeS₂),

bornite (Cu_5FeS_4), cubanite ($CuFe_2S_3$) with lesser amounts of sphalerite (ZnS), wurtzite (ZnS), galena (PbS), silica, silicates, anhydrite, and barite ($BaSO_4$). For the most part, they are rather fragile structures subject to weathering in which anhydrite redissolves and sulfides oxidize to oxyhydroxides once venting terminates.

Hydrothermal Plumes

As vent fluid is diluted with seawater, a hydrothermal "plume" is created, which can rise hundreds of meters above the vent site because of its slightly warmer temperature and therefore lower density than surrounding water. Precipitation of sulfide "smoke" immediately above the vent removes up to half the dissolved Fe. The remained is oxidized to Fe^{II} and precipitated as oxyhydroxides in the plume. The half life for Fe^{II} oxidation in seawater is anywhere from a few minutes to a day or more, depending on O_2 concentrations and pH. During Fe precipitation, a number of elements may be coprecipitated, including Mn, P, V, Cr, and As. The kinetics of Mn oxidation are considerably slower, and Mn precipitation is generally delayed until the plume reaches neutral buoyancy and begins to spread out horizontally. Mn oxidation appears to be bacterially mediated. The Fe-Mn particles produced within the plume strongly scavenge particle-reactive elements, such as Th, Be, and the rare earths, from seawater.

Hydrothermal Fluxes

The importance of mid-ocean ridge hydrothermal systems in controlling the composition of seawater was immediately realized upon the discovery of the first hydrothermal vents. Initially, it ap-
Geochemistry

W. M. White

CHAPTER 15: OCEANS

peared that estimating the flux of elements into and out of the oceanic crust was straightforward (e.g., Edmond et al., 1989; Von Damm et al., 1985). Unfortunately, the problem has proven to be not so simple. A particularly important problem is the differences between high temperature vents and diffuse vents. In diffuse vents, mixing between vent fluids and seawater leads to extensive precipitation of the dissolved metals in the subseafloor; for such elements, the global flux depends strongly on the ratio of diffuse to high temperature venting (Kadko, et al., 1995; Elderfield and Schultz, 1996). Despite the uncertainties, several approaches converge on estimates of heat flux and water flux of ridge crest hydrothermal activity of $2-4 \times 10^{12}$ W and 2- 4×10^{13} kg/yr respectively (Elderfield and Schultz, 1996). Based on these values, it is possible to make rough estimates of the global ridge crest hydrothermal flux to the oceans. The most recent estimates of these fluxes are given in Table 15.6. Estimates of the riverine flux are shown for comparison. These fluxes are lower than those estimated by Edmond et al. (1979) and Von Damm et al. (1985), in many cases by an order of magnitude. Even these recent estimates of fluxes remain substantially uncertain and should be used with caution. Nevertheless, it appears reasonably well established that hydrothermal activity appears to represent a substantial flux for many elements, including the alkalis (Li, K, Rb, Cs), Be, Mn, Fe, and Cu, and an important sink for others (e.g., Mg, U).

Low temperature basalt-seawater interaction, i.e., weathering, must also be considered in assessing global fluxes. Estimates of these fluxes are included in Table 15.6. For CO_2 Si and Ca, the low temperature flux into the crust may exceed the high temperature flux out of it. The oceanic crust is a sink for U at both high

TABLE 15.6. GLOBAL FLUXES TO SEAWATE	r from
Ridge Crest Hydrothermal Activity	

	Hydrothermal	Low-T	Riverine
	Flux	Alteration Flux	Flux
H_2	$0.3 \rightarrow 1.5 \times 10^{10}$		
H_2O		-1.3×10^{10}	2.1×10^{15}
Li	$1.2 \rightarrow 3.9 \times 10^{10}$	$0.2 \rightarrow 1.1 \times 10^{10}$	$1.4 imes 10^{10}$
Na		$-0.3 imes 10^{12}$	7.3×10^{12}
Κ	$2.3 \rightarrow 6.9 \times 10^{11}$	$-1 \rightarrow -7 \times 10^{12}$	1.9×10^{12}
Rb	$2.6 \rightarrow 9.5 \times 10^8$	$-1.9 \rightarrow -3.7 \times 10^8$	3.7×10^{8}
Cs	$2.6 \rightarrow 6.0 \times 10^{6}$	$-2.0 \rightarrow -3.8 \times 10^{6}$	$4.8 imes10^6$
Be	$3.0 \rightarrow 12 \times 10^5$		3.7×10^{6}
Mg	-1.6×10^{12}	0.2×10^{12}	5.3×10^{12}
Ca	$9 \rightarrow 1300 \times 10^9$	-1.3×10^{12}	1.2×10^{13}
Sr	0	9×10^{8}	2.2×10^{10}
Ва	$2.4 \rightarrow 13 \times 10^8$		1×10^{10}
CH_4	$0.67 \rightarrow 2.4 \times 10^{10}$		
CO_2	$1.0 \rightarrow 12 \times 10^{11}$	-2.3×10^{12}	
SiO_2	$4.3 \rightarrow 6.6 \times 10^{11}$	-7×10^{11}	6.4×10^{12}
Al	$1.2 \rightarrow 6.0 \times 10^8$		6×10^{10}
SO_4	$-8.4 imes10^{11}$		3.7×10^{12}
H_2S	$0.85 \rightarrow 9.6 \times 10^{11}$		
Mn	$1.1 \rightarrow 3.4 \times 10^{10}$	$2.8 imes10^6$	4.9×10^{9}
Fe	$2.3 \rightarrow 19 \times 10^{10}$	3.9×10^{9}	2.3×10^{10}
Со	$6.6 \rightarrow 68 \times 10^5$		$1.1 imes 10^8$
Cu	$3 \rightarrow 13 \times 10^9$		5.0×10^{9}
Zn	$1.2 \rightarrow 3.2 \times 10^9$		$1.4 imes 10^{10}$
As	$0.9 \rightarrow 140 \times 10^5$		7.2×10^{8}
Se	$3 \rightarrow 220 \times 10^4$		$7.9 imes 10^7$
Ag	$7.8 \rightarrow 11 \times 10^{6}$		$8.8 imes 10^7$
La	4.1×10^{5}		11.9×10^{6}
Ce	9.1×10^{5}		$18.8 imes 10^6$
Nd	5.3×10^{5}		9.2×10^{6}
Sm	$1.0 imes 10^5$		$1.8 imes 10^6$
Eu	3.4×10^{5}		2.2×10^{5}
Gd	$9.0 imes 10^4$		1.7×10^{6}
Dy	$6.4 imes 10^4$		
Er	$2.6 imes 10^4$		7.9×10^{5}
Yb	$1.7 imes 10^4$		7.6×10^{5}
Lu	2.1×10^{3}		1.9×10^{5}
Pb	$2.7 \rightarrow 110 \times 10^5$		$1.5 imes 10^8$
U	$-0.18 \rightarrow -1.6 \times 10^{\circ}$	7 -3.8 × 10 ⁶	$3 \rightarrow 6 \times 10^7$

All numbers are in moles/year. From Chen et al. (1986), Rudnicki and Elderfield (1993), Kadko et al. (1994), Lilley et al. (1995), Elderfield and Schultz (1996), and Staudigel et al. (1996).

and low temperature. For transition metals, the low temperature flux appears to be insignificant. For K, Rb, and Cs, the high-temperature loss from the oceanic crust appears to slightly exceed the low temperature gain; for Li, the high temperature loss significantly exceeds the low temperature gain (Elderfield and Schultz, 1996). Sr is also interesting. As Table 15.6 shows, there is no net flux of Sr to seawater from hydrothermal vents. Studies of basalt similarly show the concentration of Sr in basalt

CHAPTER 15: OCEANS

TABLE 15.7. Ridge Flank Hydrothermal Fluxes

	Flank Flux	Rivers
В	$-0.19 \rightarrow -1.9 \times 10^{10}$	5.4×10^{10}
CO_2	$2.2 \rightarrow 2.9 \times 10^{12}$	
Mg	$-0.7 \rightarrow -1.1 \times 10^{11}$	5.3×10^{12}
SiO ₂	$1.3 \rightarrow 1.8 \times 10^{12}$	$6.4 imes 10^{12}$
Р	-3.2×10^{9}	3.3×10^{10}
S	8×10^{12}	3.7×10^{12}
Ca	$2.0{\rightarrow}~5.5{\times}10^{11}$	1.2×10^{13}
Ва	2×10^{8}	1×10^{10}
U	-9.7×10^{6}	$3 \rightarrow 6 \times 10^7$

All numbers are in moles/year. From Kadko et al. (1994) quoted in Elderfield and Schultz (1996).

Table 15.8. Hydrothermal Plume Removal Fluxes

TADLE	E 17.0. HYURUTHERWALT LUVIE REMOVALT LUXES				
	Plume	Hydrothermal	Rivers		
	Removal Flux	Flux			
Be	$1.8 imes10^6$	$3.0 \rightarrow 12 \times 10^5$	3.7×10^{6}		
Р	$1.1 imes 10^{10}$		3.3×10^{10}		
V	$4.3 imes 10^{8}$		5.9×10^{8}		
Cr	$4.8 imes 10^7$		6.3×10^{8}		
Мо	$1.9 imes10^6$		2×10^{8}		
As	$1.8 imes 10^8$	$0.9 \rightarrow 140 \times 10^5$	7.2×10^{8}		
Y	2.1×10^{7}				
La	$7.5 imes10^6$	4.1×10^{5}	$11.9 imes10^6$		
Ce	0.9×10^{6}	9.1×10^{5}	$18.8 imes10^6$		
Nd	$8.8 imes10^6$	5.3×10^{5}	9.2×10^{6}		
Sm	2.1×10^{6}	1.0×10^{5}	$1.8 imes10^6$		
Eu	$5.0 imes 10^5$	3.4×10^{5}	2.2×10^{5}		
Gd	$1.9 imes 10^{6}$	$9.0 imes 10^4$	$1.7 imes10^6$		
Dy	$1.7 imes10^6$	$6.4 imes10^4$			
Er	$1.0 imes10^6$	$2.6 imes 10^4$	$7.9 imes 10^5$		
Yb	$8.1 imes 10^5$	$1.7 imes 10^4$	$7.6 imes 10^5$		
Lu	$1.0 imes 10^5$	2.1×10^{3}	1.9×10^5		
U	$4.3 imes 10^4$	-8.1×10^{-6}	$3 \rightarrow 6 \times 10^7$		

All numbers are in moles/year. From compilation by Lilley et al. (1995).

As we noted earlier, a large fraction of the transition metals dissolved in hydrothermal fluids quickly precipitate or are scavenged by precipitation of Fe sulfides and oxyhydroxides. Additional scavenging of particle-reactive elements occurs through later precipitation of Mn oxides. Estimates of the removal flux in hydrothermal plumes are listed in Table 15.8. For Be, As and the rare earths, the plume removal flux exceeds the primary hydrothermal flux, in most cases by an order of magnitude. Thus hydrothermal plumes are a net sink such elements, even though they are enriched in vent fluids.

Effect on the Oceanic Crust

Though basalt gains Mg, loses silica, etc., such changes in major element concentrations have an insignificant effect on the oceanic crust because of the high concentrations of these elements in the basalt. Minor elements may be more seriously affected. Since the basalt is subducted, these changes will ultimately affect the composition of the mantle. Quite likely, most of the water and CO_2 gained by the oceanic crust during low temperature alteration are lost during subduction. However, there

does not change significantly during alteration. However, vent waters have ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ ratios between 0.703 and 0.706, generally more similar to basalt (0.7025) than seawater. Also studies of basalts show their ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ ratios are increased by interaction with seawater. Thus hydrothermal activity serves to buffer the Sr isotopic composition of seawater: the average ${}^{87}\text{Sr}/{}^{86}\text{Sr}$ of river water is 0.7119 (Palmer and Edmond, 1989), that of seawater is 0.71018. This difference reflects the effect of hydrothermal activity. We could say that the oceanic crust is a sink for ${}^{87}\text{Sr}$ in seawater, but not for the other isotopes of Sr.

Most of the heat lost by the oceanic crust occurs not through high temperature hydrothermal sys-

> tems at the ridge flanks, but through lower temperature, more diffuse hydrothermal systems operating on ridge flanks. Because the flow is diffuse and of low temperature, these fluids have been more difficult to characterize and these systems are not well understood. Nevertheless, they are undoubtedly important in global fluxes to and from seawater. Elderfield and Schultz (1996) estimate the water flux through ridge flank systems at 3.7-11 \times 10¹⁵ kg/yr, more than 2 orders of magnitude greater than the ridge crest flow. The most recent estimates of these fluxes are listed in Table 15.7. Some of these fluxes are substantial. The amount of Mg and U removed from seawater in this way is comparable to that removed by ridge crest hydrothermal systems; the Si flux to seawater exceeds that or ridge crest systems, the S flux exceeds the riverine flux.

Geochemistry

CHAPTER 15: OCEANS

may be important effects are on those elements involved in radioactive decay schemes. We have noted the Sr isotope ratio of basalt changes as a result of reaction with hydrothermal fluids. The Rb/Sr ratio also changes and there appears to be a net decrease in Rb/Sr. REE are enriched in fluids relative to seawater, but the absolute levels are nevertheless low and the effect on the basalt is small. Thus the Sm-Nd system is little affected. The U-Th-Pb system may be profoundly affected since U, like Mg, seems to be quantitatively removed from seawater in high temperature hydrothermal reactions. Pb, on the other hand, is leached from the basalt. The net effect is an increase in the U/Pb ratio of the oceanic crust. The magnitude of the effect is, however, unclear.

Oxygen Isotope Ratios

Oxygen isotopes also behave differently at high and low temperature. At low temperature, there is a true fractionation, with the heavy isotope, ¹⁸O, partitioning preferentially into the solid phase. This tends to increase the δ^{18} O of the basalt (which starts at +6) and decrease that of seawater (which starts at 0). However, because isotopic fractionation depends inversely on the square of temperature, there is a much smaller fractionation at high temperature, and the principal effect is simple isotopic exchange (mixing). As a result, the δ^{18} O of the basalt decreases and the δ^{18} O of the water increases. Muchlenbachs and Clayton (1976) suggested that these opposing reactions actually buffer the isotopic composition of seawater, maintaining a δ^{18} O of about 0. According to them, the net of low and high temperature fractionations was about +6, just the observed difference between the oceanic crust and the oceans. Thus, the oceanic crust ends up with an average δ^{18} O value about the same as it started with, and the net effect on seawater must also be close to zero.

Studies of ophiolites have confirmed that the O isotope change at high temperature just about offsets the O isotope change at low temperature, in accordance with the theory of Muehlenbachs and Clayton (1976). Based on their study of the Samail ophiolite, Gregory and Taylor (1981) estimated that the net δ^{18} O fractionation between oceanic crust and seawater is about $\Delta = 6.1$. Thus, the oceanic crust and seawater end up with a average δ^{18} O values about the same as they start with, just as Muehlenbachs and Clayton (1976) predicted. The O isotopic composition of the Samail ophiolite is shown in Figure 15.36, illustrating how the δ^{18} O ratio of the upper crust increases while that of the lower crust decreases. Ocean Drilling Project (ODP) results show much the same pattern as the Samail ophiolite, though no complete section of the oceanic crust has yet been drilled.



Let's consider this in somewhat more detail. What are the controls on the O isotope composition of

Figure 15.36. δ^{18} O variations through a cross section of the Samail ophiolite. From Gregory and Taylor (1981).

seawater? Suppose seawater had a much lower δ^{18} O, say -5. In this case, since the Δ for net seawater—oceanic crust interaction is 6.1, but the initial difference between oceanic crust and seawater would be 11.7, seawateroceanic crust interaction would lower the δ^{18} O of the oceanic crust by 2.6 and increase the $\delta^{18}O$ of the water with which it reacted by 2.6 (assuming a molar water-rock ratio of 1). Over time, the $\delta^{18}O$ of seawater would gradually increase as seawater is cycled through the crust until it reached a steady-state value equal to that which results from seawater-oceanic crust interaction.

Since the average δ^{18} O of MORB is about +5.7 and that of seawater 0, whereas the Δ for the entire process is 6.1, the equilibrium seawater value,

Geochemistry

CHAPTER 15: OCEANS

according to Gregory and Taylor (1981), should be -0.4 per mil. This disequilibrium reflects the effect of the Antarctic and Greenland ice sheets. Ice has δ^{18} O of about -33. Storage of continental ice would tend to increase the δ^{18} O of seawater (though equilibrium crystallization of ice from seawater would result in the ice having a higher δ^{18} O, this is not the process that stores ice on continents; ice is produced by evaporation and subsequent precipitation after 'Rayleigh distillation'; see Chapter 9).

The "half-time" for this process, has been estimated to be about 46 Ma. The half-time is defined as the time required for the disequilibrium to decrease by half. For example, if the equilibrium value of the ocean is 0 ‰ and the actual value is -2 ‰, the δ^{18} O of the ocean should increase to -1 ‰ in 46 Ma. It would then require another 46 Ma to bring the oceans to a δ^{18} O of -0.5‰, etc.

The Atmospheric Source

The atmosphere is, of course, the principal source and sink of dissolved gases in the ocean, but it is also a surprisingly important source of other dissolved constituents, as well as particulate matter, in the oceans. These other constituents are derived from particles in the atmosphere called *aerosols*. Aerosols have several sources: sea spray, mineral dust derived from soils and desert sands, volcanic eruptions, condensation reactions in the atmosphere, the biosphere (including fires), and anthropogenic activity such as combustion of fossil fuels, mining and mineral processing, agriculture, and the production and consumption of various chemicals. Of these sources, sea spray is the most important;

natural gas-to-particle conversions, mineral dust, and anthropogenic sources are roughly equal in magnitude; biogenic sources are least important. However, sea spray does not represent a true flux to the oceans as it is derived directly from them.

Interestingly, sea spray does not have the same composition as seawater. Sea spray is enriched in trace metals and other substances. This reflects the enrichment of these elements found in the surface microlayer at the ocean-atmosphere interface. Within the microlayer, metals are adsorbed or complexed by organic substances that form a thin film on the sea surface.

The material flux from the atmosphere to the ocean may occur through *dry deposition*, which includes both settling of particles from the atmosphere and gas adsorption, and *wet deposition*. Wet deposition includes all matter, both particulate and gaseous, first scavenged from the atmosphere by precipitation (i.e., rain and snow) before being delivered to the oceans.

Though marine chemists generally now agree that the atmosphere is an important source for a variety of species in the ocean, quantifying the atmosphere-to-ocean flux is difficult. There are few actual measurements of dry deposition rates, and while there is a fair body of data on wet deposition, both dry and wet deposition rates are very heterogeneous in both space and time, due to variations in climate, wind patterns, and aerosol

TADLE 15.9.	ATMOS	pheric	Flux	то	THE	OCEANS
-------------	-------	--------	------	----	-----	--------

Element	Atmospheric	Riverine	Atmosphere/
	Dissolved Flux	Flux	River
Ν	2.14×10^{12}	2.50×10^{12}	0.86
Al	2.49×10^{11}	6.94×10^{10}	3.59
Si	$9.98 imes 10^{11}$	6.67×10^{12}	0.15
Р	9.69×10^{09}	1.39×10^{11}	0.07
Min	5.94×10^{08}	5.58×10^{09}	0.11
Fe	3.44×10^{10}	1.97×10^{10}	1.75
Ni	1.62×10^{08}	1.90×10^{08}	0.85
Cu	$1.78 imes10^{08}$	1.60×10^{08}	1.11
Zn	1.55×10^{09}	$9.18 imes 10^{07}$	16.92
As	4.45×10^{07}	1.33×10^{08}	0.33
Cd	4.89×10^{06}	2.70×10^{06}	1.81
La	4.10×10^{06}	5.00×10^{06}	0.82
Ce	6.50×10^{06}	6.30×10^{06}	1.03
Nd	4.40×10^{06}	4.60×10^{06}	0.96
Sm	9.10×10^{05}	1.10×10^{06}	0.83
Eu	2.00×10^{05}	3.00×10^{05}	0.67
Gd	8.50×10^{05}	1.40×10^{06}	0.61
Dy	7.20×10^{05}	1.20×10^{06}	0.60
Er	4.50×10^{05}	8.00×10^{05}	0.56
Yb	3.20×10^{05}	1.10×10^{06}	0.29
Lu	$5.80 imes10^{04}$	2.00×10^{05}	0.29
Pb	2.27×10^{08}	$1.00\times10^{\scriptscriptstyle 07}$	22.74

All fluxes are in moles/yr. Data from Duce et al. (1991) and Greaves et al. (1993). Fluxes for Cu, Pb, Al, and Fe modified based on revised solubilities given by Chester et al. (1993). Mn flux estimated from the crustal Mn/Fe ratio, the Fe atmospheric flux of Duce, and solubilities given in Chester et al. (1993).

Geochemistry

CHAPTER 15: OCEANS

source distribution. Particle concentrations can be extremely high, up to 700 μ g per cubic meter of air, over the Atlantic between 30° and 5° N, where the Northeast trades carry mineral dust from Saharan dust storms. There is also a significant flux of dust from the Asian deserts to the North Pacific, though it is smaller than the Saharan dust plume. In contrast, particle concentrations re less than 0.01 μ g/m³ over remote areas of the South Pacific. Aerosol composition also varies widely, due both to difference in sources and fractionation that occurs as the aerosol is transported from the source area. This variability, together with the vastness of the ocean, makes it difficult to derive global fluxes.

There is considerably more data on the composition of aerosols than on actual deposition rates. Thus most estimates of the atmosphere-to-ocean flux are calculated by multiplying the mean atmospheric concentration times a deposition velocity (e.g., Chester, 1990; Duce et al., 1991):

F = CV 15.41

Estimates of deposition velocity are based on models that incorporate such factors as meteorology and the nature of the sea-air interface. The reader is referred to Duce et al. (1991) for a fuller discussion.

The next question we must ask is what fraction of the particulate matter deposited on the ocean surface dissolves? This depends both on the element and the nature of the particle. Whereas the sea spray aerosol dissolves entirely, only a fraction of the mineral aerosol dissolves. Anthropogenic particles have intermediate solubilities. For example, Chester et al. (1993) found that 8% and <1% of the Al is leached by seawater from an anthropogenic aerosol collected over the U.K and a mineral dust aerosol from the Arabian Sea respectively. From these same two aerosols, 29% and 35%, respectively, of the Mn was leached.

Table 15.9 gives estimates of the atmospheric flux of dissolved matter from the oceans. The estimates were derived from the product of atmospheric concentrations, depositional velocities, and solubilities. These data suggest that the atmosphere is the principal source of dissolved Al, Fe, Cu, Zn, Cd, Ce, and Pb in the ocean. However, the fluxes of Cu, Zn, Cd, and Pb are primarily anthropogenic, as are the fluxes of Ni and As. Hence these fluxes do not necessarily represent the steady state for these elements. The fluxes of Al, Fe, and Ce are, however, primarily due to mineral dust, so that even in the absence of anthropogenic activity, the atmosphere may be the principal source of these elements.

Sedimentary Sinks and Sources

Sediments are, in one way or another, the major sink for dissolved matter in the oceans. There are several ways in which elements find their way into sediments: (1) biologic uptake, (2) scavenging by organic particles, (3) scavenging or adsorption by or reaction with clay and other particles, (4) precipitation of, coprecipitation with, or adsorption by hydroxides and oxides, and (5) precipitation as evaporite salts. In addition, diffusion of dissolved species into or out of sediments may occur. In the latter case, sediments may serve as a source, rather than a sink, for a particular element.

Biogenic Sediments and Evaporites

As we have seen, biological activity controls the distribution of not only the major nutrients but also the micronutrients. Biogenic particles incorporated in sediment are an important sink for such elements. About 60% of the ocean floor is covered with biogenic siliceous and calcareous oozes. However, the cycle of biologically utilized elements is not simple. For example, it is estimated that the residence time of SiO₂ in seawater before biological utilization is some 200-300 years. But the overall residence time is about 18,000 yrs. In other words, a Si atom will cycle between seawater and biogenic SiO₂ an average of 60 to 90 times before leaving the oceans for good. Recalling that Si is a refractory nutrient, we can expect that labile nutrients, such as P, Ni, and Cd, are recycled many more times before finally leaving the system. Biogenic opal and carbonate particles also scavenge other dissolved components from seawater as they fall through the water column, hence biogenic sediments serve as a sink even for elements not biologically utilized.

Bathymetric and oceanographic factors control the distribution of biogenic sediments. As we found earlier, calcareous oozes are found only above the carbonate compensation depth because of the increasing solubility of carbonate with depth. The distribution of siliceous oozes, on the other hand, is

Geochemistry

CHAPTER 15: OCEANS

largely independent of depth and reflects the biological productivity in overlying surface water. Siliceous oozes are found beneath the Pacific equatorial high productivity belt and beneath the productive waters at high latitudes, particularly in the Southern Ocean.

Burial of significant fractions (>0.5%) of organic matter in sediments is unusual and restricted to highly productive upwelling areas. In such circumstances, the downward flux of organic matter is greater than what can be consumed by scavenging organisms. Organic carbon that survives early diagenesis may eventually be converted to petroleum, as we found in Chapter 14. The coast of Southern California in one example of an area of petroleum generation. Beneath other upwelling areas, phosphorite actively precipitates from P-rich sediment pore waters. The P has been redissolved from sediment originally enriched in P because of the high flux of organic matter to the bottom.

Evaporites are an important sink for several of the major species in seawater, such as Na, Cl, K, and SO_4 . As their name implies, they form when seawater is evaporatively concentrated. When this occurs, salts precipitate in the order calcium carbonate, $CaSO_4 \bullet H_2O$ (gypsum), NaCl (halite), MgSO₄, NaBr, and KCl. However, evaporites rarely form in a closed system. Open system evaporation, where seawater is supplied and removed from a basin, but at rates such that saturation in the least soluble salts is maintained, but saturation of the more soluble salts not achieved. Such a system can produce extensive gypsum, or gypsum and halite, layers without precipitation of subsequent salts. Significant evaporite formation is episodic; there are only a few places where evaporites are currently form, and in all cases the volume of salt precipitating is inconsequential for the marine budget. Nevertheless, there are many examples of ancient massive evaporites in the sedimentary record. For example, much of the Mediterranean Sea is underlain by evaporites that formed some 5 million years



Figure 15.37. Elemental accumulation rate plotted against sedimentation rate for Nares Plain sediments. (After Thomson et al., 1984).

ago when that Sea's connection to the rest of the ocean was cut by tectonic processes. Hence over geologic time, evaporites do represent a major sink for major ions. It is estimated that evaporite deposits contain about as much Cl as do the oceans.

Red Clays, Metalliferous Sediments, and Mn Nodules

Lithogenous particles also scavenge dissolved components from seawater. The predominant sediment in the deep, remote areas of the world ocean is *red clay*. Red clays consist of lithogenous particles, derived from the continents and delivered by winds or rivers, that are strongly enriched in transition metals and other particlereactive elements. Indeed, their brownish red appearance and name reflect their iron-rich nature.

Why red clays are so much more enriched in these elements than rapidly accumulating sediments is best illustrated by a study of Atlantic sediments by Thomson et al. (1984). In the Nares Abyssal Plain (northeast of Puerto Rico), both 'red' and 'gray' clays occur. The gray clays are fine-grained distal ends of turbidity currents originating on the North American shelf. Red clays are restricted to areas of low sedimentation rate and have a lower ⁸⁷Sr/⁸⁶Sr ratio than gray clays (0.722 vs. 0.728). However, after leaching the ⁸⁷Sr/⁸⁶Sr of the two were identical.

Geochemistry

CHAPTER 15: OCEANS



Figure 15.38. Distribution of metalliferous sediments in the oceans. From Edmond et al. (1982).

Thomson et al. (1984) also found that when the sedimentation flux of a specific element was plotted against overall sedimentation rate, a straight line resulted whose intercept was greater than zero (Figure 15.37), implying there was a flux of the metals even when there was no sediment flux! How can this be?

The interpretation is as follows. The red and gray clay both had the same origin on the North American continent, but the red clay, which accumulates slowly, contained an adsorbed component of seawater Sr (⁸⁷Sr/⁸⁶Sr=0.709), which could be removed by leaching. Metals were also adsorbed on the gray clays, but less so because they had spent a shorter had spent a shorter period exposed to seawater than the red clay particles. The non-zero intercept indicates authigenic flux of elements such as Fe and Mn, the value of the authigenic flux being independent of the flux of sediment to the bottom. When sedimentation rates are high, the authigenic component is simply highly diluted. Lack of this dilution at low sedimentation rates results in high Fe and Mn concentrations.

Another means of removal of elements from seawater is precipitation of oxides and hydroxides, principally of Mn and Fe, and coprecipitation or adsorption of particle-reactive elements by them. This occurs in two principle ways. The first is in hydrothermal plumes. As we found in above, hy-drothermal fluids are enriched in Fe²⁺ and Mn²⁺. Fe quickly oxidizes and precipitates. Oxidation of Mn is somewhat slower, so that most precipitation is delayed until the plume becomes neutrally buoyant and begins to spread laterally. As they settle out of the water column, precipitated particles then scavenge other particle-reactive elements from seawater. When these hydrothermal particles are abundant, they produce so-called "metalliferous sediment". Al is not enriched in hydrothermal fluids and Al in marine sediments is derived entirely from continental sources. Thus sediments enriched in hydrothermally derived particles will have low ratios of Al/(Al+Fe+Mn). Figure 15.38 shows the distribution of these metal-rich sediments. Highest concentrations of metalliferous sediments are found near the mid-ocean ridges, particularly adjacent fast spreading regions of the EPR, but the influence of hydrothermal plumes can still be seen thousands of kilometers from the ridge crest.

Geochemistry

CHAPTER 15: OCEANS

Mn and Fe oxides and hydroxides may also precipitate directly on the seafloor, with a previously existing surface acting as a nucleation site. In sediment-covered areas, shards of volcanic glass, shark's teeth and other such particles may serve as a nucleation site, with the Mn-Fe precipitates eventually forming a coating of up to 10 or more centimeter diameter. Typically, the form flattened spheres with botryoidal, smooth, or rough surfaces. These are known as *manganese nodules*. Solid surfaces, such as the surface of a lava flow, may also provide a nucleation site. In this case, the Mn-Fe precipitates will form a coating on the surface up to several cm thick. Such coatings are known as *manganese crusts*. These nodules and crusts grow extremely slowly and occur only in areas of low sedimentation rate. They are most common in the deep basins of the Central Pacific, as low sedimentation rates are most common there, but they also occur in the other oceans. Nodules and crusts also occur on seamounts, mid-ocean ridges, and some areas of continental margins.

Mn nodules and crusts consist principally of mixtures of δMnO_2 , birnessite (also called 7Å manganite), toderokite (also called 10Å manganite), and amorphous iron hydroxide (FeOOH• nH_2O). δMnO_2 , birnessite, and manganite all consist of primarily of sheets of MnO_2 , but differ in their structure and the amount of water and other metals they contain. Their average composition is given in Table 15.10. As Mn and Fe particles of hydrothermal plumes, nodules and crusts scavenge other particle-reactive elements from solution so that they are usually strongly enriched in several transition metals as well as other particle reactive elements such as the rare earths, as may be seen from Table 15.10. This enrichment has generated interest in the possibility of mining nodules for Ni, Cu, and Co, but while mining companies have invested in exploration and research on Mn nodules, no large scale mining operations have been undertaken yet.

Nodules may grow by precipitation* from seawater or by precipitation from sediment pore waters. Nodules growing from seawater are called "hydrogenous"; those growing from sediment pore waters

Flement	Average Conc	Enrichment	Floment	Average Conc	Enrichment	-
Liement	Average Colic.	Entern	Liement	Average Colic.	Entern	
	wi. percent	Factor		ppin	Factor	_
Na	1.94	0.82	В	277	27.7	
Mg	1.82	0.78	Sr	830	2.2	
AĪ	3.06	0.34	Y	310	9.39	
Si	8.62	0.31	Zr	648	3.92	
Р	0.22	2.13	Мо	412	274.7	
Κ	0.64	0.31	Pd	0.0055	0.83	
Ca	2.47	0.56	Ag	6	85.7	
Sc	0.00097	0.44	Cď	7.9	39.5	
Ti	0.65	1.14	Sn	2.7	13.5	
V	0.056	4.13	Te	50		
Cr	0.0035	0.35	La	160	5.33	
Mn	16.02	168.6	Yb	31	10.33	
Fe	15.55	2.76	W	60	40	
Со	0.28	113.6	Ir	0.00935	70.83	
Ni	0.48	64.0	Au	0.00248	0.62	
Cu	0.26	47.0	Hg	0.50	6.25	
Zn	0.078	11.2	ΤĬ	129	286.66	
Ga	0.001	0.67	Pb	900	72.72	
Ва	0.20	4.73	Вi	8	47.05	

TAble 15.10. Average Composition of Manganese Nodules and Crusts

From Cronan, 1980. Enrichment Factor is the enrichment over average continental crust.

^{*} Hydrogenous nodules and crusts may grow primarily from colloidal, rather than dissolved, Mn and Fe in seawater. Strictly speaking, then, the process is not one of precipitation and accumulation may better term.

Geochemistry

W. M. White

CHAPTER 15: OCEANS

and are called "diagenetic". Hydrogenous nodules tend to be dominantly composed of δMnO_2 , while toderokite is more common in diagenetic nodules. Dymond et al. (1984) distinguished two kinds of diagenetic growth: oxic and suboxic. Suboxic diagenetic nodules occur beneath high productivity regions where sufficient organic matter reaches the sediment that they become reducing at depth, thereby mobilizing Mn as Mn²⁺. Most nodules are of mixed diagenetic and hydrogenetic origin, growing both upward by precipitation from seawater and downward, by precipitation from porewater. This produces a compositional difference between the tops and bottoms of nodules. In contrast to nodules, crusts, which form on impermeable surfaces, grow by accretion from seawater and are thus strictly "hydrogenous". Near mid-ocean ridges and other submarine volcanos, crusts may grow primarily from metals provided by hydrothermal vents. Such crusts are termed "hydrothermal".

As mentioned, most Mn nodules grow extremely slowly, rates of 3-5 mm/Ma are typical. Growth rates of suboxic diagenetic nodules may be substantially higher, up to 200 mm/Ma. The slow growth explains why nodules are found only in areas of low sedimentation rate: in other areas, the nuclei are buried before having a chance to grow. Nevertheless, many nodules appear to grow are rates lower than the local sedimentation rate. Even more puzzling, they are concentrated at the sediment sur-



Figure 15.39. Fe, Mn, and Cu, Ni, and Zn concentrations of nodules and crusts from three sites in the eastern equatorial Pacific analyzed by Dymond et al. (1984). Nodules from the red clay site (Site R) have combined hydrogenous and oxic diagenetic growth mechanisms. Nodules from the siliceous ooze site (Site S) grow predominantly through oxic diagenesis. Nodules from the hemipelagic site (Site H) grow primarily by suboxic diagenesis, which results from high organic matter content of the sediment. Crusts from site H, which are hydrogenous, are compositionally similar to site R nodules. From Dymond et al. (1984).

face: there are typically twice as many nodules at the sediment surface as buried in the upper meter of sediment. This appears to be due to the action of burrowing organisms that keep the nodules at the sediment surface. Nodules vary greatly in abundance, both regionally and locally. This variation is largely related to sedimentation rate. On a local scale, nodules are more abundance on topographic highs, apparently because currents can keep such areas relatively sediment-free.

There are significant variations in the composition of nodules and crusts, as is shown in Table 15.11. These variations are related to the growth mode and variations in the supply of metals (Figure 15.39). Hydrogenous crusts and nodules tend to be richest in Fe and Co. Nodules growing principally through oxic diagenesis are richest in Cu, Ni, and Zn, while suboxic diagenetic nodules are rich in Mn and depleted in Co. Hydrothermal crusts are generally poor in Fe as well as Co, Ni, and Cu. Even within a single genetic type, however, significant compositional variations occur. In a study of Mn crusts from the Line Islands in the central equatorial Pacific, Aplin and Cronan (1984) found that the concentrations of most trace metals, including Co, Ni, Cd, Mo, Zn, and Pb were highest in crusts from depths of less than 2000 m. These crusts also had the highest Mn concentrations. Concentrations of Fe, Cu, and Be were highest in those crusts from >2000 m depth. Aplin and Cronan (1984) suggested these variations reflected higher concentrations of dissolved and colloidal Mn at depths of 1000 to 2000 m in the Pacific. This Mn enrichment results from both reductive dissolution of particulate Mn in the oxygen minimum and horizontal advective transport of Mn diffusing out of sediments of the continental margins. They argued that variation in concentrations of other metals primarily reflects the differing affinities of these metals for Fe and Mn oxide surfaces, with elements such as Co, Ni, Cd, and Mo being particularly strongly scavenged by Mn oxides compared to Fe oxides.

Geochemistry

CHAPTER 15: OCEANS

	sou or composi			JOJN J.
Hydrogenous Crust	Oxic Diagenetic Nodule	Suboxic Diagenetic Nodule	Hydrothermal Crust	Hydrothermal Crust
22.2	31.65	48.0	41.0	55.0
19.0	4.45	0.49	0.8	0.2
1300	280	35	33	39
5500	10100	4400	310	180
1480	4400	2000	120	50
750	2500	2200	400	2020
1.2	7.1	98	51	275
	Hydrogenous Crust 22.2 19.0 1300 5500 1480 750 1.2	Hydrogenous Crust Oxic Diagenetic Nodule 22.2 31.65 19.0 4.45 1300 280 5500 10100 1480 4400 750 2500 1.2 7.1	Hydrogenous Crust Oxic Diagenetic Nodule Suboxic Diagenetic Nodule 22.2 31.65 48.0 19.0 4.45 0.49 1300 280 35 5500 10100 4400 1480 4400 2000 750 2500 2200 1.2 7.1 98	Hydrogenous Crust Oxic Diagenetic Nodule Suboxic Diagenetic Nodule Hydrothermal Crust 22.2 31.65 48.0 41.0 19.0 4.45 0.49 0.8 1300 280 35 33 5500 10100 4400 310 1480 4400 2000 120 750 2500 2200 400 1.2 7.1 98 51

TABLE 15.11. COMPARISON OF COMPOSITION OF DIFFERENT MARINE FE-MN DEPOSITS.

From Chester (1990).

Diffusion Into and Out of Sediments

Sediments can serve as a sink for dissolved matter in yet another way: through diffusion of dissolved components into sediments. Dissolved components may also diffuse out of sediment poor waters into seawater, or pore water may be expelled by compaction. In these latter cases, sediments serve as a source of dissolved matter in seawater. As we found in Chapter 5, diffusion occurs only when a compositional gradient exists. Sediment pore water originates simply as seawater trapped between sedimentary particles. Its composition is thus initially identical to seawater. Reactions occurring within the sediment, however, produce changes in pore water composition, establishing chemical gradients that drive diffusion into or out of the sedimentary column. Furthermore, as sediment is buried beneath subsequently accumulating material, it is compacted, driving pore water back into the overlying seawater, producing a flux of porewater-enriched components to seawater. In this section we consider the examples of two elements: U and Li. Diffusion into sediments is an important sink for dissolved U, while diffusion and porewater expulsion is an important source of dissolved Li in the oceans.

U is present in seawater in the VI state, generally as the soluble uranyl tricarbonate species $(UO_2(CO_3)_3^{-4})$. The reduced species, however, U (IV) is relatively insoluble. While seawater only rarely becomes reducing (examples are the deep or bottom waters of the Black Sea, some fjords, and the Cariaco Trench), suboxic or anoxic conditions are more frequently achieved at depth in marine sediments. This occurs in regions where there is a flux of high organic carbon to the seafloor as a result of high biological productivity in the overlying surface water. Such areas occur most often on or near continental margins and cover roughly 8% of the total area of the sea floor.

Figure 15.40a shows an example of the U profiles determined by Klinkhammer and Palmer (1991) in cores taken from the California continental shelf just south of the Monterey Fan. Between 1 and 2 cm depth, U pore water concentrations slightly exceed the seawater concentration (13.4 nmol/l). This results from release of U from labile organic phases in the upper part of the core. At great depths, however, U pore water concentrations to concentrations around 5 nmol/l above seawater values occur. U concentrations in the coexisting solid phase increase sharply from 2.7 ppm to 5.7 ppm at depths of 6 to 9 cm. Consumption of the 2.5% organic carbon in this core lead to suboxic conditions and reduction of U^{6+} to U^{4+} , as well as reduction of Mn and Fe. Judging from the sharp increase in U concentrations in the solid phase, it appears that pe values appropriate for reduction of U first occur at depths of around 6 cm. Increases in pore water concentrations of Mn and Fe are highly insoluble in their oxidized states. Once reduced, U is immobilized in the solid phase, reducing it concentration in the pore waters. This produces a concentration gradient that causes U to diffuse downward from seawater into the sediment.

Geochemistry

CHAPTER 15: OCEANS



Figure 15.40. a. U concentration in sediment (black squares, in ppm) and pore waters (red circles, in nmol/kg), from the California Shelf sediments. b. Mn (red squares) and Fe (black diamonds) in the same pore waters as in (a). Data from Klinkhammer and Palmer (1991).

Organic-rich sediments tend to be rich in U generally. Whether this reflects biological uptake of U or adsorption of U on dead organic particles falling through the water column remains unclear. This U may be released when the organic matter is remineralized in the sediment, producing high concentrations of U in sediment pore water and, consequently, diffusion of U from sediments to seawater. This appears to be occurring, for example, in sediments on the Brazilian continental shelf near the Amazon River.

On the whole, diffusion into sediments appears to be dominant. Table 15.12 shows the marine U

budget according to Barnes and Cochran (1990). They estimate diffusion into suboxic sediments, as exemplified by the Figure 15.39, removes 0.25 to 0.32×10^{10} g/yr. Klinkhammer and Palmer (1991) estimate a higher flux of 0.67 $\times 10^{10}$ g/yr for suboxic diagenesis. Both sets of authors find that this is the single largest sink for dissolved U in the oceans.

Diffusion out of sediments appears to be an important source of dissolved Li in the oceans. Li is readily incorporated into interlayer sites of clays at low temperature, with some clays mineral containing as much as 500 ppm Li. Because of it high hydration energy, however, Li is easily removed from these sites. Thus as sediments are subjected to increasing temperatures during burial, Li is lost from the sediment solid phases and its concentration in pore waters increases. This sets up a concentration gradient that drives diffusion of Li out of the sediment. Perhaps more importantly, Li-rich pore waters are expelled from the sediment as it compacts, producing a flux of dis- From Barnes and Cochran (1990).

TABLE 15.12. THE MARINE	U Budget
Sources (10 ¹⁰ g U/yr)	
Riverine Input	1
Amazon Shelf Sediments	0.14
Total	1.14
Sinks (10 ¹⁰ g U/yr)	
Sediments	
Oxic, deep sea	0.08
Metalliferous	0.14
Underlying anoxic water	0.13
Suboxic	0.25-0.32
Corals and Molluscs	0.08
Oceanic Crust	
Low temperature	0.23
High temperature	0.04
Total	0.95-1.02

CHAPTER 15: OCEANS

solved Li to seawater. Li concentrations in pore waters and coexisting sediments from ODP Site 688 on the Peru margin analyzed by Martin et al. (1991) shown in Figure 15.41 provide an example. Concentrations increase from near seawater values (25.6 μ M) at the top of the core to over 500 μ M at a depth of 450 m. At the same time, Li concentrations in the sediment decrease. This decrease in Li concentration in sediments does not seem sufficient to account for the increase in pore water Li concentrations and Martin et al. (1991) speculated that additional Li was being released from underlying basement rocks.

Martin et al. (1991) estimated that expulsion of pore waters of continental margin sediments supplied 1 to 3×10^{10} moles/yr to the ocean. The high end of this estimate exceeds both the riverine flux (1.1–1.7 × 10¹⁰ moles/yr) and the hydrothermal flux (0.4 – 1.3 × 10¹⁰ moles/yr). Thus sediment pore water may be the dominant source of dissolved Li in the oceans. Even the minimum value represents a substantial flux, supplying 20-25% of the marine Li.

References and Suggestions for Further Reading

Alt, J. C. 1995. Subseafloor processes in mid-ocean ridge hydrothermal systems. In Seafloor Hydrothermal Systems: Physical, Chemical, Biological and Geological Interactions, Geophysical Monograph Vol. 91, S. E. Humphris, R. A. Zierenberg, L. S. Mullineaux and R. E. Thomson. ed., pp. 85-114. Washington: AGU.



Figure 15.41. Li concentrations in pore waters (red circles, μ M) and sediments (black squares) from ODP Site 688 on the Peru continental shelf. Data from Martin et al. (1991).

- Aplin, A. C. and D. S. Cronan. 1985. Ferromanganese oxide depos- Martin et al. (1991). its from the Central Pacific Ocean, I. Encrustations from the Line Islands Archipelago. *Geochim. Cosmochim. Acta*. 49: 427-436.
- Barnes, C. E. and J. K. Cochran, 1990. Uranium removal in oceanic sediments and the oceanic U balance. *Earth. Planet. Sci. Lett.*, 97: 94-101.
- Barth, T. W. 1952. Theoretical Petrology. New York: John Wiley.
- Berner, R. A. 1975. The role of magnesium in the crystal growth of aclite and aragonite from seawater. *Geochim. Cosmochim. Acta.* 39: 489-504.
- Bischoff, J. L. and K. S. Pitzer. 1989. Liquid-vapor relations for the system NaCl-H2O: summary of the P-T-x surface from 300° to 500°C. *Am. J. Sci.* 289: 217-248.
- Boyle, E. A., J. M. Edmond and E. R. Sholkovitz. 1977. The mechanism of iron removal in estuaries. *Geochim. Cosmochim. Acta.* 41: 1313-1324.
- Boyle, E. A., F. R. Sclater and J. M. Edmond. 1977. The distribution of dissolved copper in the Pacific. *Earth Planet. Sci. Lett.* 37: 38-54.
- Bowers, T. S., A. C. Campbell, C. I. Measures, A. J. Spivack and J. M. Edmond. 1988. Chemical controls on the composition of vent fluids at 13°-11°N and 21°N, East Pacific Rise. *J. Geophys. Res.* 93: 4522-4536.

Broecker, W. S. and T.-H. Peng. 1982. Tracers in the Sea. New York: Eldigio Press.

- Bruland, K. W. 1983. Chapter 45: Trace elements in seawater. In *Chemical Oceanography, Vol. 8,* J. P. Riley and R. Chester. ed., pp. 157-220. London: Academic Press.
- Busenberg, E. and L. N. Plummer. 1985. Kinetic and thermodynamic factors controlling the distribution of SO₄²⁻ and Na⁺ in calcites and selected aragonites. *Geochim. Cosmochim. Acta*. 49: 713-726.
- Chen, J. H., G. J. Wasserburg, K. L. von Damm and J. M. Edmond. 1986. The U-Th-Pb systematics in hot springs on the East Pacific Rise at 21°N and Guaymas Basin. *Geochim. Cosmochim. Acta*. 50: 2467-2479.
- Chester, R., Marine Geochemistry, Unwin Hyman, London, 698 p., 1990.

Geochemistry

CHAPTER 15: OCEANS

- Chester, R., K. J. T. Murphy, F. J. Lin, A. S. Berry, G. A. Bradshaw and P. A. Corcoran. 1993. Factors controlling the solubilities of trace metals from the non-remote aerosols deposited to the sea surface by the 'dry' deposition mode. *Mar. Chem.* 42: 107-126.
- Chou, L., R. M. Garrels and R. Wolast. 1989. Comparative study of the kinetics and mechanisms of dissolution of carbonate minerals. *Chem. Geol.* 78: 269-282.
- Craig, H. 1974. A scavenging model for trace elements in the deep sea. *Earth Planet. Sci. Lett.* 23: 149-159.
- Cronan, D. S. 1980. Underwater Minerals. London: Academic Press.
- Duce, R. A., P. S. Liss, J. T. Merrill, E. L. Atlas, P. Buat-Menard, B. B. Hicks, J. M. Miller, et al. 1991. The atmospheric input of trace species to the world ocean. *Global Biogeochem. Cycles.* 5: 193-259.
- Dymond, J., M. Lyle, B. Finney, D. Z. Piper, K. Murphy, R. Conard and N. Pisias. 1984. Ferromanganese nodules from MANOP Sites H, S, and R — Control of mineralogical and composition by multiple accretionary processes. *Geochim. Cosmochim. Acta*. 48: 931-949.
- Dzombak, D. A. and F. M. M. Morel. 1990. Surface Complexation Modelling: Hydrous Ferric Oxide. New York: Wiley Interscience.
- Eckert, J. M. and E. R. Sholkovitz. 1976. The flocculation of iron, aluminum, and humates from river water by electrolytes. *Geochim. Cosmochim. Acta*. 40: 847-848.
- Edmond, J. M., A. Spivack, B. C. Grant, H. Ming-Hui, C. Zexiam, et al., Chemical dynamics of the Changjiang (Yangtze) Estuary, *Cont. Shelf Res.*, 4, 17-36, 1985.
- Edmond, J. M., C. Measures, R. E. McDuff, L. H. Chan, R. Collier, et al., Ridge crest hydrothermal activity and the balances of the major and minor elements in the ocean: the Galapagous data, *Earth and Planetary Science Letters*, 46, 1979.
- Edmond, J.M., Von Damm, K.L., McDuff, R.E., and Measures, C.I., (1982) Chemistry of hot springs on the East Pacific Rise and their effluent dispersal. *Nature*, **297**: 187-191.
- Edmonds, H. N. and J. M. Edmond. 1995. A multicomponent mixing model for ridge-crest hydrothermal fluids. *Earth Planet. Sci. Lett.* 134: 53-67.
- Elderfield, H., N. Luedtke, R. J. McCaffrey and M. Bender. 1981. Benthic flux studies in Narragansett Bay. *Am. J. Sci.* 281: 768-787.
- Elderfield, H., R. Upstill-Goddard and E. R. Sholkovitz. 1990. The rare earth elements in rivers, estuaries, and coastal seas. *Geochim. Cosmochim. Acta*. 54: 971-992.
- Elderfield, H. and A. Schultz. 1996. Mid-ocean ridge hyddrothermal fluxes and the chemical composition of the ocean. *Annu. Rev. Earth Planet. Sci.* 24: 191-224.
- Elderfield, H. and E. R. Sholkovitz. 1987. Rare earth elements in the pore waters of reducing nearshore sediments. *Earth Planet. Sci. Lett.* 82: 280-288.
- Erel, Y. and J. J. Morgan. 1991. The effect of surface reactions on the relative abundances of trace metals in deep-ocean water. *Geochim. Cosmochim. Acta*. 55: 1807-1813.
- Froelich, P. N., G. A. Hambrick, M. O. Andreae, R. A. Mortlock and J. M. Edmond. 1985. The geochemistry of inorganic germanium in natural waters. J. Geophys. Res. 90: 1131-1141.
- Goldberg, E. D. and G. O. S. Arrhenius. 1958. Chemistry of Pacific pelagic sediments. *Geochim. Cosmochim. Acta.* 13: 153-212.
- Greaves, M. J., P. J. Statham and H. Elderfield. 1994. Rare earth element mobilization from marine atmospheric dust into seawater. *Mar. Chem.* 46: 255-260.
- Gregory, R. T. and H. P. Taylor, An oxygen isotope profile in a section of Cretaceous oceanic crust, Samail Opiolite, Oman: Evidence for δ^{18} O buffering of the oceans by deep (>5 km) seawater-hy-drothermal circulation at mid-ocean ridges, *J. Geophys. Res.*, 86, 2737-2755, 1981.
- Hoyle, J., H. Elderfield, A. Gledhill and M. Greives, The behavior of the rare earth elements during mixing of river and seawaters, *Geochim. Cosmochim. Acta*, 48, 143-149, 1984.
- Kadko, D., J. Baross and J. Alt. 1995. The magnitude and global implications of hydrothermal flux. In Seafloor Hydrothermal Systems: Physical, Chemical, Biological and Geological Interactions, Geophysical Monograph Vol. 91, S. E. Humphris, R. A. Zierenberg, L. S. Mullineaux and R. E. Thomson. ed., pp. 446-466. Washington: AGU.

Geochemistry

CHAPTER 15: OCEANS

- Kadko, D., E. Baker, J. Alt and J. Baross. 1994. Global impact of submarine hydrothermal processes. *RIDGE Vent Workshop*. 55p.
- Keir, R. S. 1980. The dissolution kinetics of biogenic calcium carbonates in seawater. *Geochim. Cosmo-chim. Acta.* 44: 241-252.
- Klinkhammer, G. P. and M. R. Palmer. 1991. Uranium in the oceans: where it goes and why. *Geochim. Cosmochim. Acta.* 55: 1799-1896.
- Kraepiel, A. M. L., J.-F. Chiffoleau, J.-M. Martin and F. M. M. Morel. 1997. Geochemistry of trace metals in the Gironde estuary. *Geochim. Cosmochim. Acta*. 61: 1421-1436.
- Lilley, M. D., F. R. A. and J. H. Trefry. 1995. Chemical and biochemical transformations in hydrothermal plumes. In Seafloor Hydrothermal Systems: Physical, Chemical, Biological and Geological Interactions, Geophysical Monograph Vol. 91, S. E. Humphris, R. A. Zierenberg, L. S. Mullineaux and R. E. Thomson. ed., pp. 369-391. Washington: AGU.
- Mantoura, R. F. C., A. Dickson and J. P. Riely. 1978. The complexation of metals with humic materials in natural waters. *Estuar. Coastal Shelf Sci.* 6: 367-408.
- Martin, J. B., M. Kastner and H. Elderfield. 1991. Lithium: sources in pore fluids of Peru slope sediments and implications for oceanic fluxes. *Mar. Geol.* 102: 281-292.
- Martin, J. H. and R. M. Gordon. 1988. Northeast Pacific iron distributions in relation to phytoplankton productivity. *Deep Sea Res.* 35: 177-196.
- Martin, J. H., K. H. Coale, K. S. Johnson, S. E. Fitzwater, R. M. Gordon, S. J. Tanner and e. al. 1994. Testing the iron phyothesis in ecosystems of the equatorial Pacific. *Nature*. 371: 123-129.
- Measures, C. I., B. Grant, M. Khadem, D. S. Lee and J. M. Edmond. 1984. Distribution of Be, Al, Se and Bi in the surface waters of the western North Atlantic and Carribean. *Earth Planet. Sci. Lett.* 71: 1-12.
- Millero, F. J. 1995. Thermodynamics of the carbon dioxide system in seawater. *Geochim. Cosmochim. Acta*. 59: 661-678.
- Mucci, A. 1983. The solubility of calcite and aragonite in seawater of various salinities, temperatures, and one atmosphere total pressure. *Am. J. Sci.* 283: 780-799.
- Muehlenbachs, K. and R. Clayton, Oxygen isotope geochemistry of submarine greenstones, *Can. J. Earth. Sci.*, 9, 471-478, 1972.
- Muehlenbachs, K., Oxygen isotope composition of the oceanic crust and its bearing on seawater, J. *Geophys. Res.*, 81, 4365-4369, 1976.
- Nozaki, Y. 1997. A fresh look at element distribution in the North Pacific Ocean. EOS. 78: 221. (Electronic supplement may be foun at: "http://www.agu.org/eos_elec/97025e.html").
- Orians, K. J., E. A. Boyle and K. W. Bruland. 1990. Dissolved titanium in the open ocean. *Nature*. 348: 322-325.
- Peterson, M. N. A. 1966. Calcite: rates of dissolution in a vertical profiel in the central Pacific. *Science*. 154: 1542-1544.
- Quinby-Hunt, M. S. and K. K. Turekian. 1983. Distribution of elements in sea water. EOS. 64: 130-132.
- Redfield, A. C. 1958. The biological control of chemical factors in the environment. *Am. J. Sci.* 46: 205-221.
- Rudnicki, M. D. and H. Elderfield. 1993. A chemical model of the buoyant and neutrally buoyant plume above the TAG vent field. *Geochim. Cosmochim. Acta*. 57: 2939-2957.
- Schaule, B. K. and C. C. Patterson. 1983. Perturbations of the natural lead depth profile in the Saragasso Sea by industrial lead. In *Trace metals in Seawater*, Vol. C. S. Wong, E. Boyle, K. W. Bruland, J. D. Burton and E. D. Goldberg. ed., pp. 487-503. New York: Plenum Press.
- Seyfried, W. E. J. and K. Ding.1995. Phase equilibria in subseafloor hydrothermal systems: a review of the role of redox, temperature, pH and dissolved Cl on the chemistry of hot spring fluids at midocean ridges. In Seafloor Hydrothermal Systems: Physical, Chemical, Biological and Geological Interactions, Geophysical Monograph Vol. 91, S. E. Humphris, R. A. Zierenberg, L. S. Mullineaux and R. E. Thomson. ed., pp. 248-272. Washington: AGU.
- Shanks, W. C. I., J. K. Böhlke and S. R. R. II. 1995. Stable isotopes in mid-ocean ridge hydrothermal systems: interactions between fluids, minerals, and organisms. In *Seafloor Hydrothermal Systems*:

Geochemistry

CHAPTER 15: OCEANS

Physical, Chemical, Biological and Geological Interactions, Geophysical Monograph Vol. 91, S. E. Humphris, R. A. Zierenberg, L. S. Mullineaux and R. E. Thomson. ed., pp. 194-221. Washington: AGU.

- Sholkovitz, E. R. 1976. Flocculation of dissolved organic and inorganic matter during the mixing of river water and seawater. *Geochim. Cosmochim. Acta.* 40: 831-845.
- Sholkovitz, E. R. 1978. The flocculation of dissolved Fe, Mn, Al, Cu, Ni, Co and Cd during estuarine mixing. *Earth Planet. Sci. Lett.* 41: 77-86.
- Sholkovitz, E. R. 1992. Chemical evolution of rare earth elements: fractionation between colloidal and solution phases of filtered river water. *Earth Planet. Sci. Lett.* 114: 77-84.
- Sholkovitz, E. R., E. A. Boyle and N. B. Price. 1978. The removal of dissolved humic acids and iron during estuarine mixing. *Earth Planet. Sci. Lett.* 40: 130-136.
- Sholkovitz, E. and D. Copland. 1981. The coagulation, solubility and adsorption properties of Fe, Mn, Cu, Ni, Cd, and Co and humic acids in river water. *Geochim. Cosmochim. Acta.* 45: 181-189.
- Stanley, J. K. and R. H. Byrne. 1990. Inorganic speciation of zinc(II) in seawater. *Geochim. Cosmo-chim. Acta*. 54: 753-760.
- Staudigel, H., T. Plank, W. White and H.-U. Schmincke. 1996. Geochemical fluxes during seafloor alteration of the basaltic upper oceanic crust: DSDP sites 417 and 418. In *Subduction: Top to Bottom*, *Geophysical Monograph Vol. 96*, G. Bebout, D. W. Scholl, S. H. Kirby and J. P. Platt. ed., pp. 19-37. Washington: AGU.
- Stumm, W. and J. J. Morgan. 1996. Aquatic Chemistry. New York: Wiley Interscience.
- Stuvier, M., P. Quay and H. G. Ostland. 1983. Abyssal water carbon-14 distribution and the age of the world oceans. *Science*. 219: 849-851.
- Takahashi, T. 1989. The carbon dioxide puzzle. Oceanus. 32 (2): 22-29.
- Taylor, S. R. and S. M. McLennan. 1985. *The Continental Crust: Its Composition and Evolution*. Oxford: Blackwell Scientific Publications.
- Thompson, G. A., Hydrothermal fluxes in the ocean, in *Chemical Oceanography*, *vol8*, edited by J. P. Riley and C. R., p. 270-337, Academic Press, London, 1983.
- Thomson J., M. S. N. Carpenter, S. Colley, T. R. S. Wilson, H. Elderfield, and H. Kennedy. 1984. *Geochim. Cosmochim. Acta*, 48:1935-1948.
- Turner, D. R., M. Whitfield and A. G. Dickson. 1981. The equilibrium speciation of dissolved components in freshwater and seawater at 25°C and 1 atm pressure. *Geochim. Cosmochim. Acta*. 45: 855-882.
- Von Damm, K. L. 1988. Systematics of and postulated controls on submarine hydrothermal solution chemistry. J. Geophys. Res. 93: 4551-4562.
- Von Damm, K. L. 1995. Controls on the chemistry and temporal variablility of seafloor hydrothermal systems. In Seafloor Hydrothermal Systems: Physical, Chemical, Biological and Geological Interactions, Geophysical Monograph Vol. 91, S. E. Humphris, R. A. Zierenberg, L. S. Mullineaux and R. E. Thomson. ed., pp. 222-247. Washington: AGU.
- Von Damm, K. L. and J. L. Bischoff. 1987. Chemistry of hydrothermal solutions from the southern Juan de Fuca ridge. *J. of Geophys. Res.* 92: 11334-11346.
- Von Damm, K. L., J. M. Edmond, B. Grant and C. K. Measures. 1985. Chemistry of submarine hydrothermal solutions at 21° N, East Pacific Rise *Geochimica et Cosmochimica Acta*. 49: 2197-2220.
- Wong, C. S., E. Boyle, K. W. Bruland, J. D. Burton and E. D. Goldberg (ed.), 1983, *Trace Metals in Seawater*, New York: Plenum Press.
- Zhong, S. and A. Mucci. 1993. Calcite precipitation in seawater using a constant addition technique: a new overall reaction kinetic expression. *Geochim. Cosmochim. Acta*. 57: 1409-1418.
- Zuddas, P. and A. Mucci. 1994. Kinetics of calcite precipitation from seawater: I. a classical chemical description for strong electrolyte solutions. *Geochim. Cosmochim. Acta*. 58: 4353-4352.

Geochemistry

CHAPTER 15: OCEANS

Problems

1. In Figure 15.7, the δ^{13} C profile shows a pronounced maximum whereas the ΣCO_2 profile does not. Why? (HINT: Organic matter is decomposed rapidly in the ocean, calcium carbonate dissolves much more slowly).

2. A number of reaction mechanisms have been proposed for the precipitation of calcite from seawater. These include

 $Ca^{2+} + HCO_{3} \rightleftharpoons H^{+} + CaCO_{3}$ $Ca^{2+} + CO_3^{2-} \rightleftharpoons CaCO_3$ $Ca^{2+} + 2HCO_3 \rightleftharpoons H_2CO_3 + CaCO_3$ $Ca^{2+} + HCO_{3} + OH^{-} \rightleftharpoons H_{2}O + CaCO_{3}$

Suggest a series of laboratory experiments that you could performed that would enable you to distinguish which of these mechanisms actually occurs. Assume you have a well equiped laboratory in which you can measure all macroscopic properties (i.e., concentrations, partial pressure, pH, carbonate alkalinity). Describe what properties you would measure and how you would use the data you obtained to descriminate between these mechanisms.

3. How does calcite solubility vary with temperature and pressure in the ocean? Assume that temperature can be represented by a simple function of pressure:

$$T = 24 e^{-P/5} + 1$$

where pressure is in MPa and temperature in °C. Make a plot of the solubility product as a function of depth between 0 and 5000 m using the equations in Examples 15.2 and 15.3. (Hint: remember to use thermodynamic temperature).

- 4. Composition of seawater (a.) Calculate the molar concentrations of the major ions in seawater listed in Table 15.2.
 - (b.) Calculate the ionic strength of this solution.

(c.) Using the equilibrium constants in Table 12.1, calculate the concentration of carbonate ion in equilibrium with this solution at 25°C.

(d.) Calculate the total alkalinity of this solution assuming a pH of 8.1.

5. Use the one-dimensional advection diffusion model in the depth interval of 595m to 4875 m and the chemical data from the North Pacific in the adjacent table to answer the following questions. For this locality, the ratio K was determined to be 2300 and ω to be 4. Is salinity conservative and the one-dimensional model applicable? Make a plot of S vs. f(z) (equation 15.33). Are Cu, Ni, and Al conservative? Are they being produced or scavenged? Concentrations in nmol/kg; salinity in ppt.

			-	
Depth	Salinity	Cu	Ni	Al
0		0.54	2.49	90
75	33.98	0.69	2.9	88
185	33.92	0.91	3.79	84
375	34.05	1.45	5.26	70
595	34	1.9	7.49	60
780	34.19	2.15	9.07	52
985	34.37	2.38	9.64	48
1505	34.55	2.8	9.79	45
2025	34.61	3.18	10.6	47
2570	34.65	3.46	10.8	50
3055	34.66	3.9	10.9	54
3533	34.66	4.26	10.7	63
4000	34.67	4.57	10.8	66
4635	34.68	5.03	10.3	74
4875	34.68	5.34	10.4	79

CHEMICAL DATA FROM THE NORTH PACIFIC

For each of these elements, find a value of ψ or \tilde{J} to fit the one-dimensional advection-diffusion model to the data.

6. Stanley and Byrne (1990) give the following stability constants for Zn complexes in seawater:

$$Zn^{2+} + Cl^{-} \rightleftharpoons ZnCl^{-} \qquad \beta^{*} = -0.40$$
$$Zn^{2+} + HCO_{\frac{-}{3}} \rightleftharpoons ZnHCO_{\frac{+}{3}} \qquad \beta^{*} = 0.83$$

Geochemistry

CHAPTER 15: OCEANS

$\operatorname{Zn}^{2+} + \operatorname{CO}_{3}^{2-} \rightleftharpoons \operatorname{ZnCO}_{3}$	$\beta^* = 2.87$
$\operatorname{Zn}^{2+} + 2\operatorname{CO}_{3}^{2-} \rightleftharpoons \operatorname{Zn}(\operatorname{CO}_{3})_{2}^{2-}$	$\beta^* = 4.41$
$\operatorname{Zn}^{2+} + \operatorname{H}_2\operatorname{O} \rightleftharpoons \operatorname{Zn}(\operatorname{OH})^+ + \operatorname{H}^+$	$\beta^{*} = -9.25$
$Zn^{2+} + SO_4^{2-} \rightleftharpoons ZnCO_4$	$\beta^* = 0.90$

Using these stability constants, a pH of 8.1, the ligand concentrations given in Table 15.2 (and the equilibrium concentration of carbonate ion calculated with the equilibrium constant in Table 6.1 for a temperature of 25°C), calculate the fraction of Zn present as each of these species plus free Zn^{2+} .

7. Using the flows of water through mid-ocean ridge crest and flank hydrothermal systems and the mass of the oceans given in Appendix I, how long does it take to cycle the entire ocean through these systems?

8. In the San Clemente Basin, off the southern California coast, Barnes and Cochran (1990) found that U concentrations in sediment pore waters decreased to 3.35 nM/l in the top 7 cm of sediment. Assuming an effective diffusion coefficient (corrected for porosity and tortuosity in the sediment) of 68.1 cm²/yr, calculate the flux (in nM/cm²) of U from seawater to sediment in this locality.

9. Using the fluxes of U to the ocean in Table 15.11, estimate the residence time of U in the ocean.

Geochemistry

APPENDIX II: CONSTANTS., UNITS & CONVERSIONS

Physical and Chemical Constants

Speed of Light (c)	$2.998 \times 10^8 \mathrm{m/s}$
Planck's Constant (h)	6.626 ×10 ⁻³⁴ J/Hz
Boltzmann's Constant (k)	$1.380 \times 10^{-23} \text{ J/K}$
Graviational Constant (G)	$6.672 \times 10^{-11} \mathrm{N} \cdot \mathrm{m}^2 / \mathrm{kg}$
Avagadro's Number (N _A)	$6.022 \times 10^{23} \text{ mol}^{-1}$
Gas Constant (R)	8.314 J/mol-K
	(1.987 cal/mol-K)
Faraday Constant (F)	96.487 kJ/V-eq.
Electron Charge (e)	1.602×10^{-19} coulombs (C)
Permittivity in vacuum (ε)	$8.85 \times 10^{-12} \text{ C}^2/\text{J-m}$
Dielectric constant of water	78.54

THE EARTH

Mass of the Earth (M_{\oplus})	$5.97 \times 10^{24} \mathrm{kg}$
Mantle	4.0×10^{24} kg
Core	$1.94 \times 10^{24} \text{ kg}$
Continental Crust	$2.2 \times 10^{22} \text{ kg}$
Oceans	$1.4 \times 10^{21} \text{ kg}$
Atmosphere	$5.1 \times 10^{18} \text{ kg}$
Mean Radius	$6.37 \times 10^{6} \mathrm{m}$
Radius of Core	$3.47 \times 10^{6} \text{ m}$
Radius of Orbit	$1.49 \times 10^{11} \mathrm{m}$
The Sun	
Mass (M_{\odot})	$1.99 \times 10^{30} \text{ kg}$
Radius	$6.96 \times 10^8 \text{ m}^{\circ}$

SI Units and Conversions

Mass

Pound

amu (atomic mass unit)

Distance

inch ångstrom mile (US) antronomical unit (AU) parsec

light-year Force

Kilogram (kg) 1 lb = 0.4535 kg(1 kg = 2.205 lb) $1 \text{ amu} \equiv \text{mass of } ^{12}\text{C atom}$ $1 \text{ amu} = 1.66 \times 10^{-27} \text{ kg}$ Meter (m) 1 in = .0254 m $1\text{\AA} \equiv 10^{-10} \text{ m}$ 1 mi = 1609 m 1 AU = 1.609 m $1 \text{ AU} = 1.49 \times 10^{11} \text{ m}$ $1 \text{ parsec} = 3.084 \times 10^{16} \text{ m}$ $= 2.07 \times 10^5 \text{ AU}$ $1 ly = 6.35 \times 10^4 AU$ Newton (N) $1 \text{ N} \equiv 1 \text{ kg-m/s}^2$ $1 \text{ dyne} = 10^5 \text{ N}$ $1 \text{ dyne} \equiv 1 \text{ gm-cm}/\text{sec}^2$

Geochemistry

APPENDIX II: CONSTANTS., UNITS & CONVERSIONS

Energy

erg

calorie liter-atmosphere liter-Pascal electron volt Dalton Volt kilowatt-hour Pressure pascal bar atmosphere Volume

U.S. gallon Concentration molarity molality

Joule (J) $1 \text{ J} \equiv 1 \text{ kg-m}^2/\text{s}^2$ $1 \text{ erg} = 10^{-7} \text{ J}$ $1 \text{ erg} = 1 \text{ gm-cm}^2/\text{sec}^2$ 1 calorie = 4.184 J1 l-atm = 101.29 J $1 \text{ l-Pa} = 99.98 \times 10^{-5} \text{ J}$ $1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}$ 1 dalton = 9.315×10^2 MeV 1 Volt-coloumb = 1 J $1 \text{ kWh} = 3.6 \times 10^6 \text{ J}$ Pascal (Pa) $1 \text{ Pa} \equiv 1 \text{ N}/\text{m}^2$ $1 \text{ bar} = 10^5 \text{ Pa} (= 0.1 \text{ MPa})$ $1 \text{ atm} = 1.013 \times 10^5 \text{ Pa}$ Liter (l) $1 l \equiv 10^3 \text{ cm}$ $1 l = 10^{-6} m^3$ 1 gal = 3.785 lmoles/l(M)moles/kg (m) $1 \,\mu M$ (micromole) = $10^{-6} M$ $1 \text{ nM} (\text{nanomole}) = 10^{-9} \text{ M}$ $1 \text{ pM} (\text{picomole}) = 10^{-12} \text{ M}$

 $1 \text{ fM (femtomole)} = 10^{-15} \text{ M}$

Geochemistry

APPENDIX III SUMMARY OF IMPORTANT EQUATIONS

EQUATIONS OF STATE:

Ideal GasLaw:

$$\mathbf{PV} = \mathbf{NRT}$$

Coefficient of Thermal Expansion:

$$\boldsymbol{\alpha} \equiv \frac{1}{\mathbf{V}} \left(\frac{\partial \mathbf{V}}{\partial \mathbf{T}} \right)$$

Compressibility:

$$\beta \equiv -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)$$

Van der Waals Equation:

$$\mathbf{P} = \frac{\mathbf{RT}}{\overline{\mathbf{V}} - b} - \frac{a}{\overline{\mathbf{V}}^2}$$

The Laws of Thermdynamics:

First Law:

$$\Delta U = Q + W$$
written in differential form: $\mathbf{d} \mathbf{U} = \mathbf{d} \mathbf{Q} + \mathbf{d} \mathbf{W}$
2

work done on the system and heat added to the system are positive. The first law states the equivalence of heat and work and the conservation of energy.

Second Law:

$$\mathbf{d}\mathbf{Q}_{\mathbf{rev}} = \mathbf{T}\mathbf{d}\mathbf{S}$$

3

Two ways of stating the second law are *Every system left to itself will, on average, change to a condition of maximum probability* and *Heat cannot be extracted from a body and turned entirely into work.*

Third Law:

$$\lim_{\mathbf{T}\to 0} \mathbf{S} = 0$$

This follows from the facts that $S = R \ln \Omega$ and $\Omega = 1$ at T = 0 for a perfectly crystalline pure substance.

Primary Variables of Thermodynamics

The leading thermodynamic properties of a fluid are determined by the relations which exist between the volume, pressure, termperature, energy and entropy of a given mass of fluid in a state of thermodynamic equilibrium - J. W. Gibbs

The primary variables of thermodynamics are P, V, T, U, and S. Other thermodynamic functions can be stated in terms of these variables. For various combination of these variables there are

Geochemistry

EQUATION SUMMARY

characteristics functions. The characteristic function for S and V is one of the primary variables: U. Thus

$$\mathbf{dU} = \mathbf{T}\mathbf{dS} + \mathbf{P}\mathbf{dV}$$
 5

Other Important Thermodynamic Functions

What then is the use of thermodynamic equations? They are useful precisely because some quantities are easier to measure than others. - M. L. McGlashan

Enthalpy:
$$\mathbf{H} \equiv \mathbf{U} + \mathbf{P}\mathbf{V}$$
6In differential form in terms of its characteristic variables: $d\mathbf{H} = \mathrm{TdS} + \mathrm{VdP}$ 7Helmholtz Free Energy: $\mathbf{A} \equiv \mathbf{U} - \mathrm{TS}$ 8and: $d\mathbf{A} = -\mathrm{PdV} - \mathrm{SdT}$ 9Gibbs Free Energy: $\mathbf{G} \equiv \mathbf{H} - \mathrm{TS}$ 10The Gibbs Free Energy change of a reaction at constant temperature and pressure is:10aand: $\Delta \mathbf{G}_{\mathbf{r}} = \Delta \mathbf{H}_{\mathbf{r}} - \mathrm{T}\Delta \mathbf{S}_{\mathbf{r}}$ 10a

and:

Your choice of which of these functions to use should depend on what the independent variables in your system are. In geochemistry, P and T are the most common independent variables, so the Gibbs Free Energy is often the function of choice.

()

Exact Differentials and the Maxwell Relations

Any expression that may be written:

$$M(x,y)dx+N(x,y)dy$$
 12

is an exact differential if there exists a function z = f(x,y) such that

$$f(x,y) = M(x,y)dx + N(x,y)dy$$
13

The total differential of the function z(x,y) is written:

$$dz = \left(\frac{\partial z}{\partial x}\right)_{y} dx + \left(\frac{\partial z}{\partial y}\right)_{x} dy = Mdx + Ndy$$
 14

If dz is an exact differential, then

$$\frac{\partial \mathbf{M}}{\partial \mathbf{y}} = \frac{\partial \mathbf{N}}{\partial \mathbf{x}}$$
 15

which is equivalent to:

$$\left(\frac{\partial \mathbf{M}}{\partial \mathbf{y}}\right)_{\mathbf{y}} = \left(\frac{\partial \mathbf{N}}{\partial \mathbf{x}}\right)_{\mathbf{x}}$$
 16

APPENDIX

Geochemistry

EQUATION SUMMARY

Appendix 💵

All thermodynamic <u>variables of state</u> are exact differentials. Thus the practical application of the properties of exact differentials can be illustrated as follows. Equation 11 (dG = VdP - SdT) has the form dz = M(x,y)dx+N(x,y)dy since V and S are functions of temperature and pressure. Equation 11 may also be written as

$$dG = \left(\frac{\partial G}{\partial P}\right)_{T} dP + \left(\frac{\partial G}{\partial T}\right)_{P} dT$$
 17

and comparing equations 11 and 16, we conclude that

$$\left(\frac{\partial \mathbf{G}}{\partial \mathbf{P}}\right)_{\mathbf{T}} = \mathbf{V} \qquad \mathbf{and} \left(\frac{\partial \mathbf{G}}{\partial \mathbf{T}}\right)_{\mathbf{P}} = -\mathbf{S}$$
 18, 19

Applying the rule embodied in Equation 15, we can conclude that:

$$\left(\frac{\partial \mathbf{V}}{\partial \mathbf{T}}\right)_{\mathbf{P}} = -\left(\frac{\partial \mathbf{S}}{\partial \mathbf{P}}\right)_{\mathbf{T}}$$
 20

Playing similar games with Equations 5 through 9, we can develop a series of relationships:

from dE
$$\left(\frac{\partial \mathbf{T}}{\partial \mathbf{V}}\right)_{\mathbf{S}} = -\left(\frac{\partial \mathbf{P}}{\partial \mathbf{S}}\right)_{\mathbf{V}}$$
 21

from dH
$$\left(\frac{\partial \mathbf{T}}{\partial \mathbf{P}}\right)_{\mathbf{V}} = \left(\frac{\partial \mathbf{V}}{\partial \mathbf{S}}\right)_{\mathbf{P}}$$
 22

from dA
$$\left(\frac{\partial P}{\partial T}\right)_{V} = \left(\frac{\partial S}{\partial V}\right)_{T}$$
 23

Equations 20 - 23 are known as the Maxwell Relations.

DERIVATIVES OF ENTROPY

pressure:
$$\left(\frac{\partial S}{\partial P}\right)_{T} = -\alpha V$$
 24

temperature:
$$\left(\frac{\partial S}{\partial T}\right)_{V} = \frac{C_{V}}{T}$$
 and $\left(\frac{\partial S}{\partial T}\right)_{P} = \frac{C_{P}}{T}$ 25, 26

volume
$$\left(\frac{\partial S}{\partial V}\right)_{T} = \frac{\alpha}{\beta}$$
 27

DERIVATIVES OF ENTHALPY

pressure
$$\left(\frac{\partial H}{\partial P}\right)_{T} = V(1 - \alpha T)$$
 28

temperature
$$\left(\frac{\partial H}{\partial T}\right)_{P} = C_{P}$$
 29

DERIVATIVES OF ENERGY

temperature:
$$\left(\frac{\partial U}{\partial T}\right)_{V} = C_{V}$$
 and $\left(\frac{\partial U}{\partial T}\right)_{P} = C_{P} - P\alpha V$ 30, 31

Geochemistry

APPENDIX

EQUATION SUMMARY

volume: $\left(\frac{\partial U}{\partial V}\right)_{T} = \frac{T\alpha}{\beta} - P$

32

Difference between C_P and C_V

$$C_{\rm P} - C_{\rm V} = \frac{TV\alpha^2}{\beta}$$
 33

THE GIDDS PHASE RULE:

The Gibbs Phase Rule is a rule for determining the *degrees of freedom* of a system.

$$\mathbf{f} = \mathbf{c} - \mathbf{p} + \mathbf{2}$$
 34

f is the number of degrees of freedom, c is the number of components, and p is the number of phases. The minimum number of components needed to describe a system is:

$$c = N - R$$

where N is the number of species, and R is the number of reactions possible between these species.

The Clapeyron Equation

The slope of a phase boundary in P-T space is:

$$\frac{\mathbf{dT}}{\mathbf{dP}} = \frac{\Delta \mathbf{V}_{\mathbf{r}}}{\Delta \mathbf{S}_{\mathbf{r}}}$$
35

Solutions

Raoult's Law: applies to ideal solutions:

$$\mathbf{P}_{i} = \mathbf{X}_{i} \mathbf{P}_{total}$$
 36

Henry's Law: applies to very dilute solutions, and state that the partial pressure of a component in solution is proportional to it mole fraction:

$$\mathbf{P}_{i} = \mathbf{h}\mathbf{X}_{i} \quad \text{for } \mathbf{X}_{i} << 1$$

Chemical Potential

Chemical potential is defined as:

$$\mu_{i} = \left(\frac{\partial \mathbf{G}}{\partial \mathbf{n}_{i}}\right)_{\mathbf{P},\mathbf{T},\mathbf{n}}$$
38

where n_i is the number of moles of the *i*th component.

In multicomponent systems, the full expression for the Gibbs Free Energy is:

$$dG = VdP - SdT + \sum_{i} \mu_{i} dn_{i}$$
³⁹

Geochemistry

APPENDIX

EQUATION SUMMARY

THE GIDDS-DUHEM RELATION

At equilibrium and at constant pressure and temperature:

$$\sum_{i} n_{i} d\mu_{i} = 0$$
 40

THERMODYNAMIC VARIABLES IN IDEAL SOLUTIONS

$$\mu_{i, ideal} = \mu_{i}^{0} + \mathbf{RT} \ln \mathbf{X}_{i}$$

$$41$$

$$\Delta V_{ideal \ mixing} = 0$$
 and therefore: $\overline{V}_{ideal} = \sum_{i} X_{i} \overline{V}_{i} = \sum_{i} X_{i} \overline{V}_{i}$

$$\Delta H_{ideal \ mixing} = 0$$
 and therefore: $\overline{H}_{ideal} = \sum_{i} X_{i}h_{i} = \sum_{i} X_{i}\overline{H}_{i}$

$$\Delta S_{ideal \ mixing} = -R \sum_{i} X_{i} ln \ X_{i}$$

$$S_{ideal \ solution} = \sum_{i} X_{i} \overline{S}_{i} - R \sum_{i} X_{i} ln \ X$$
42

$$\Delta G_{\text{ideal mixing}} = RT \sum_{i} X_{i} \ln X_{i}$$
43

$$\overline{G}_{ideal \ solution} = \sum_{i} X_{i} \mu_{i}^{o} + RT \sum_{i} X_{i} ln \ X$$

$$44$$

THERMODYNAMIC VARIABLES IN NON-IDEAL SOLUTIONS

Fugacity: Fugacity can be thought of as the escaping tendency of a gas in non-ideal solutions. Because systems tend toward ideal at low pressure, it has the property:

$$\underset{P \to 0}{\lim} \frac{f_{i}}{P_{i}} = 1$$

$$45$$

$$\mu_{i} = \mu_{i}^{o} + RT \ln \frac{f_{i}}{f_{i}^{o}}$$

$$46$$

Activity: Activity is defined as:

and

$$\mathbf{a}_{i} \equiv \frac{\mathbf{f}_{i}}{\mathbf{f}_{i}^{0}}$$

$$47$$

hence:

$$\mu_{i} = \mu_{i}^{o} + \mathbf{RT} \ln \mathbf{a}_{i}$$

$$48$$

M. White

Geochemistry

EQUATION SUMMARY

APPENDIX

The activity in an ideal solution is:

When Henry's Law law holds:

$$a_{i,ideal} = X_i$$
 49

The activity coefficient, λ , is defined as:

$$a_i = X_i \lambda_i$$
 50

$$\lambda_i = h_i$$
 51

The **Debye-Hückel equation** is used to calculate activity coefficients in aqueous solutions. It is:

$$\log_{10} \gamma_{i} = \frac{-Az_{i}^{2}\sqrt{I}}{1 + Ba\sqrt{I}}$$
52

where z is charge, I is ionic strength, å is the hydrated ionic radius (significantly larger than ionic radius), and A and B are solvent parameters. I is calculated as:

$$I = \frac{1}{2} \sum_{i} m_i z_i^2$$
53

Excess Free Energy and activity coefficients:

$$\overline{G}_{excess} = RT \sum_{i} X_{i} ln\lambda_{i}$$
54

Excess Free Energy and Margules Parameters of a Regular Solution:

$$\overline{\mathbf{G}}_{\mathrm{ex}} = \mathbf{X}_1 \mathbf{X}_2 \mathbf{W}_{\mathrm{G}}$$
55

Excess Free Energy and Margules Parameters of an Asymmetric Solution:

$$G_{\text{excess}} = \left(W_{G_1} X_2 + W_{G_2} X_1 \right) X_1 X_2$$
 56

Equilibrium Constant

The equilibrium constant is defined as:

$$\mathbf{K} = \prod_{i} \mathbf{a}_{i}^{\mathbf{v}_{i}}$$
 57

It is related to the Gibbs Free Energy change of the reaction by:

$$\mathbf{K} = \mathbf{e}^{-\Delta \mathbf{G}^{*}/\mathbf{RT}}$$
58

It is related to enthalphy and entropy changes of the reaction by:

$$\ln K = -\frac{\Delta H_r^0}{RT} + \frac{\Delta S_r^0}{R}$$
 59

Pressure and temperature dependencies of the equilibrium constant are:

$$\left(\frac{\partial \ln K}{\partial P}\right)_{T} = -\frac{\Delta V_{r}^{o}}{RT}$$
60

Oxidation and Reduction:

The redox potential is related to the Gibbs Free Energy change of reaction as:

$$\Delta \mathbf{G} = -\mathbf{n} \mathbf{F} \mathbf{E}$$
 61

Geochemistry

APPENDIX II

EQUATION SUMMARY

where E is the redox potential, n is the number of electrons exchanged and F is the Faraday constant. The Nernst Equation is:

$$E = E^{\circ} - \frac{RT}{nF} \ln \Pi a_i^{\nu_i}$$
⁶²

The pɛ is defined as:

$$\mathbf{p}\boldsymbol{\varepsilon} = -\log \mathbf{a}_{\mathbf{e}^-}$$
 63

and is related to hydrogen scale redox potential, E_H, as:

$$\mathbf{p}\boldsymbol{\varepsilon} = \frac{\mathbf{F}\mathbf{E}_{\mathbf{H}}}{\mathbf{2.303RT}}$$

Kinetics

Reaction Rates: For a reaction such as:

$$aA + bB \rightleftharpoons cC + dD$$

A general form for the rate of a reaction is:

$$\frac{1}{a}\frac{dA}{dt} = \frac{1}{b}\frac{dB}{dt} = -\frac{1}{c}\frac{dC}{dt} = -\frac{1}{d}\frac{dD}{dt} = k A^{n_A}B^{n_B}C^{n_c}D^{n_D}$$
65

where n_A, etc. can be any number. For an elementary reaction, this reduces to:

$$\frac{1}{a}\frac{dA}{dt} = \frac{1}{b}\frac{dB}{dt} = -\frac{1}{c}\frac{dC}{dt} = -\frac{1}{d}\frac{dD}{dt} = kA^{a}B^{b}$$
⁶⁶

The temperature dependence of the rate constant is given by the Arrhenius Relation:

$$\left| \boldsymbol{k} = \mathbf{A} \, \exp\!\!\left(- \frac{\mathbf{E}_{\mathbf{B}}}{\mathbf{RT}} \right) \right| \tag{67}$$

Rate constants of elementary reactions are related to the equilibrium constant as:

$$\frac{k_{+}}{k_{-}} = \frac{[\mathbf{B}]_{\mathbf{eq}}}{[\mathbf{A}]_{\mathbf{eq}}} = \mathbf{K}^{\mathbf{app}}$$
68

69

Diffusion: Fick's First Law is:

J = -D

дc

$$\left(\frac{\partial \mathbf{c}}{\partial \mathbf{t}}\right)_{\mathbf{x}} = \mathbf{D} \left(\frac{\partial^2 \mathbf{c}}{\partial x^2}\right)_{\mathbf{t}}$$
 70

Geochemistry

APPENDIX

EQUATION SUMMARY

74

The temperature dependence of the diffusion coefficient is:

$$\mathbf{D} = \mathbf{D}_{\mathbf{o}} \exp\left(-\frac{\mathbf{E}_{\mathbf{A}}}{\mathbf{RT}}\right)$$
71

$$\left(\frac{\partial \mathbf{C}_{\mathbf{i}}}{\partial \mathbf{t}}\right)_{\mathbf{x}} = -\left(\frac{\partial \mathbf{F}_{\mathbf{i}}}{\partial \mathbf{x}}\right)_{\mathbf{t}} + \sum \mathbf{R}_{\mathbf{i}}$$
72

TRACE ELEMENTS

Diagentic Equation:

Equilibrium or Batch Partial Melting:

Fractional Partial Melting: $\frac{C_i^{\ell}}{C_i^{\circ}} = \frac{1}{D} \left(1 - F\right)^{1/D - 1}$

Zone Refining:

$$\frac{C_{i}^{\ell}}{C_{i}^{o}} = \frac{1}{D} - (\frac{1}{D} - 1)e^{-DR}$$
75

Equilibrium Crystallization:

Fractional Crystallization:

 $\frac{\mathbf{C}_{i}^{l}}{\mathbf{C}_{i}^{e}} = \frac{1}{\mathbf{D}\mathbf{X} + (\mathbf{1} - \mathbf{X})}$ 76

$$\boxed{\frac{C_i}{C_i^{\circ}} = (1 - X)^{D-1}}$$
77

ISOTOPE GEOCHEMISTRY

Binding Energy per Nucleon: E	$_{\rm b} = \left[\frac{{\rm W}-{\rm M}}{{\rm A}}\right]{\rm c}^2$	78
Basic Equation of Radioactive Decay:	$\frac{\mathbf{dN}}{\mathbf{dt}} = -\lambda \mathbf{N}$	79
Isotope Growth (or Isochron) Equation:		
R = F	$R_0 + R_{P/D} \left(e^{\lambda t} - 1\right)$	80

Geochemistry

APPENDIX IIII: SOME /MATHEMATICS USEFUL IN GEOCHEMISTRY

LINEAR REGRESSION

Fitting a line to a series of data is generally done with a statistical technique called *least squares regression*. Real data are not likely to fall exactly on a straight line; each point will deviate from the line somewhat. The idea of least squares regression is to find the best line fitting the data by minimizing the *squares of the deviations* from the regression line. The quantity to be minimized is:

$$\sum_{i=1}^{n} e^2 = \sum_{i=1}^{n} (y - a - bx)^2$$
(1)

This is know as the *sum of the squares of the deviations* from the line y = a + bx. The use of the squares of the deviations means that large deviations will affect the calculated slope more than small deviations. By differentiating equation (1), it can be shown that the minimum value for the left side occurs when the slope is:

$$b = \frac{\sum (x_{i} - \bar{x})(y_{i} - \bar{y})}{\sum (x_{i} - \bar{x})^{2}}$$
(2)

where \bar{x} and \bar{y} are the means of x and y respectively and x_i and y_i are the *i*th pair of observations of x and y respectively. We can see from 7.23 that the regression slope is the product of the deviations of x and y from the mean divided by the square of the deviations of x from the mean. A more convenient computational form of (2) is:

$$b = \frac{\sum (x_i y_i) - \bar{y}\bar{x}n}{\sum x_i^2 - \bar{x}^2 n}$$
(3)

The intercept is then given by:

 $a = \bar{y} - b\bar{x} \tag{4}$

Because real data never fit a line exactly, it is of interest to know the error on the estimate of slope and intercept.. The error on the slope is given by:

$$\sigma_{\rm b} = \sqrt{\left[\sum y_i^2 - \bar{y}^2 n - \frac{\left(\sum (x_i \, y_i) - \bar{y} \bar{x} n\right)^2}{\sum x_i^2 - \bar{x}^2 n}\right]} \left[\frac{1}{(n-2)(\sum x_i^2 - \bar{x}^2 n)}\right]$$
(5)

The error on the intercept is:

$$\sigma_{a} = \sqrt{\left[\sum y_{i}^{2} - \bar{y}^{2}n - \frac{\left(\sum (x_{i} y_{i}) - \bar{y}\bar{x}n\right)^{2}}{\sum x_{i}^{2} - \bar{x}^{2}n}\right] \left[\frac{1}{n} + \frac{\bar{x}^{2}}{\left(\sum x_{i}^{2} - \bar{x}^{2}n\right)}\right] \left[\frac{1}{n-2}\right]}$$
(6)

Statistics books generally give an equation for linear least squares regression in terms of one dependent and one independent variable. The independent variable is assumed to be known absolutely. With geochemical data, both x and y are often measured parameters and have some error associated with them. These must be taken into account for a proper estimate of the slope and the errors associated with it. In some cases, the errors in measurement of *x* and *y* can be correlated, and this must also be taken into account. The so-called *two-error regression* algorithm does this. This is, however, considerably less straight-forward than the above. The approach is to weight each observation according to the measurement error (the weighting factor will be inversely proportional to the analytical error). A solution, written in the context of geochronology, has been published by York (1969).

Geochemistry

ERRATA FOR 2001 VERSIONS

Errata for 2001 version

(CORRECTED IN ZOOS VERSIONS)

Chapter 1

Page 10, Reference to Fig. 1.6 should be Fig 1.4.

Page 19: Pyroxenes have two bridging oxygens per tetrahedron; amphiboles have 2.5 bridging oxygens between tetrahedron

Chapter 2

p. 30 equation 2.26: delete extra x

p. 31: equation 2.32 denominator on left hand side of equation should be ∂T , not ∂P

p. 51: Figure number 2.13 should be numbered as 2.11.

p. 33. The word "permutation" should be replaced by "combination" on this page and elsewhere

- p. 44 equation 2.89 should read $U = nkT^2 \left(\frac{\partial \ln Q}{\partial T}\right)_V$
- p. 44, equation 2.90 should read $U = RT^2 \left(\frac{\partial \ln Q}{\partial T}\right)_V$

p. 48. equation 2.102 should read: $C_p = a + bT - \frac{c}{T^2}$

p. 47 last paragraph, replace "equation 2.125" with "equation 2.93"

p. 53 equation 2.120 should read $S_{298} = \int_0^{298} \frac{C_P dT}{T} + S_0 + \Delta S_{\phi}$

p. 53 second equation in example 2.5 should read

$$\Delta H = \left[aT + \frac{b}{2}T^2 + \frac{c}{T}\right]_{298}^{573} = \left[46.94 \times T + \frac{0.0343}{2}T^2 + \frac{1129680}{T}\right]_{298}^{573}$$

p. 57 Example 2.7 third paragraph should read "...we use equations 2.127 and 2.128 to see how ΔG will change"

Chapter 3

p. 64. Since the reaction: $K - Na = KNa_{.1}$ may also be written, this system have 4 species, 2 reactions and hence a minimum of 2 components.

p. 71. Equation 3.19 should read:

$$G = \sum_{i} n_{i} \left(\frac{\partial G}{\partial n_{i}} \right)_{P,T,n_{j,j\neq i}} = \sum_{i} n_{i} \mu_{i}$$

p. 75 "defined as" sign in equation 2.41 should be replaced by an equal sign.

p. 78 Equation 3.49 should read:

Geochemistry

ERRATA FOR ZOO1 VERSIONS (CORRECTED IN $\mathbb{ZOO}^{\mathbb{Z}}$ VERSIONS) $\int_{P_{ref}}^{P'} \Delta V_r dP = \Delta V^s (P' - P_{ref}) + \int_{f_{ref}}^{f'} V_{H_2O} df = \Delta V^s (P' - P_{ref}) + RT \ln \frac{f'}{f^{ref}}$

p 78.

Reference to equation 2.139 in Example 3.2 should be to equation 2.133.

equation 3.70 – should read: p. 83

 $\lim_{m \to 0} \frac{a_i}{m_i} = 1$

p. 92 Equation 3.84 should read

$$\sum \mathbf{v}_i \mu_i^o + RT \ln \prod_i a_i^{\mathbf{v}_i} = 0$$

Equation 3.85 should read $K = \prod_{i} a_{i}^{v_{i}}$ p. 107 Equation 3.120 should read: $\log K = -2pH - 2p\varepsilon$

Chapter 4

p. 129. Equations 4.29 and 4.30 should read:

$$\mu_{Ab}^{os} - \mu_{Ab}^{o\ell} = RT \ln \left(\frac{X_{Ab}^{\ell}}{X_{Ab}^{s}}\right)$$
$$\mu_{An}^{os} - \mu_{An}^{o\ell} = RT \ln \left(\frac{X_{An}^{\ell}}{X_{An}^{s}}\right)$$

and