Chapter 1

INTRODUCTION TO NMR SPECTROSCOPY

1.1 Introduction

Nuclear magnetic resonance (NMR) is a spectroscopic technique that detects the energy absorbed by changes in the nuclear spin state. The application of NMR spectroscopy to the study of proteins and nucleic acids has provided unique information on the dynamics and chemical kinetics of these systems. One important feature of NMR is that it provides information, at the atomic level, on the dynamics of proteins and nucleic acids over an exceptionally wide range of time scales, ranging from seconds to pico-seconds. In addition, NMR can also provide atomic level structural information of proteins and nucleic acids *in solution* (see Fig. 1.1), i.e. there is no need to crystallize the sample for NMR studies. Thus NMR provides a method of obtaining structural information if the molecule cannot be crystallized or there is some question regarding a structure obtained by X-ray crystallography. Lastly, it is relatively easy to study protein-ligand interactions under physiological conditions by simply adding ligand to the NMR sample of the unliganded protein.

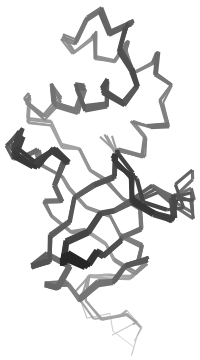


Figure 1.1. Protein structure determined by NMR spectroscopy. Four structures of a 130 residue protein, derived from NMR constraints, are overlaid to highlight the accuracy of structure determination by NMR spectroscopy.

Although NMR is a powerful technique, it does have its limitations. First, almost all experiments require that the observed NMR absorption peaks are assigned to a particular atom in the protein. Although resonance assignment methods are well characterized, they do require considerable time for data acquisition and analysis. Secondly, the size of the protein or nucleic acid that can be studied by NMR is limited. Assemblies with rotational correlation time of greater than 25 ns (corresponding to a protein with a molecular weight of 60 kDa) may be difficult to study at the detailed atomic level. However, more limited NMR studies can be performed on much larger proteins and biological assemblies. Generally, it is necessary to label larger proteins with ^{13}C, ^{15}N, and perhaps ^{2}H, to successfully apply NMR techniques to such large systems.

Labeling of this type is most easily accomplished biosynthetically in either *E. coli* or in tissue culture (at a much higher expense). A rough indication of the isotopic labeling requirements as a function of protein size is given in Table 1.1. Lastly, due to the small energy difference between the ground and excited state of the nuclear spins, NMR is a particularly insensitive technique. Protein concentrations on the order of 0.5 to 1 mM are typical, thus a single 0.4 ml NMR sample of a 20 kDa protein would require between 4 and 8 mg of protein. Fortunately, the techniques are not destructive and the sample can be used for other purposes.

For most of this text we will employ a semi-classical model of the nuclear spins to obtain an intuitive understanding of many of the fundamental aspects of modern NMR spectroscopy. In this chapter we will highlight a number of important features of NMR spectroscopy, including:

1. How energy states are created by the magnetic field,
2. The relationship between the environment and the absorption energy,
3. Coupling between nuclear spins.

1.2 Classical Description of NMR Spectroscopy

The basic phenomenon of nuclear magnetic resonance NMR spectroscopy is similar to other forms of spectroscopy, such as visible spectroscopy. A photon of light causes a transition from the ground state to the excited state. For example, in the case of visible spectroscopy the absorption of a photon by an electron causes the electron to move from its ground state orbital to an orbital of higher energy, the excited state. In the case of NMR, the absorption of a radio-frequency photon promotes a nuclear spin from its ground state to its excited state.

NMR spectroscopy differs in a number of important aspects from other forms of spectroscopy. First, the generation of the ground and excited NMR states requires the existence of an external magnetic field. This requirement is a very important distinction of NMR spectroscopy in that it allows one to change the characteristic frequencies of the transitions by simply changing the applied magnetic field strength. Second, the NMR excited state has a lifetime that is on the order of 10^9 times longer than the lifetime of the excited electronic states. This difference in lifetimes follows directly from Einstein's law for spontaneous emission that relates the lifetime of the excited state, τ, to the frequency of the

Table 1.1. Molecular Weight Limitations for Chemical Shift Assignments

Isotopic Labeling	*Mol. Weight*
None	$\leq$10 kDa
^{15}N	10-15 kDa
^{15}N, ^{13}C	15-30 kDa
^{15}N, ^{13}C, ^{2}H	30-60 kDa

transition, ω:

$$\tau \propto \frac{1}{\omega^3} \tag{1.1}$$

The long lifetime of the excited state implies extremely narrow spectral lines since the ability to define the energy of a transition is proportional to the lifetime of the excited state [1]. In the case of small organic molecules, linewidths less than 1 Hz are easily attainable. Thus it is possible to detect small changes in absorption energies that arise from subtle differences in the environment of a nuclear spin. The persistence of the excited state also facilitates multidimensional spectroscopy, by allowing the resonance frequency information associated with one spin to be passed to another. Finally, the long lifetime of the excited state permits the measurement of molecular dynamics over a wide range of time scales.

1.2.1 Nuclear Spin Transitions

In all forms of spectroscopy it is necessary to have two or more different states of the system that differ in energy. In a system with two energy levels, the one of lower energy if often referred to as the ground state and the higher energy state is the excited state. In the case of nuclear magnetic resonance spectroscopy, the energies of the states arise from the interaction of a *nuclear magnetic dipole moment* with an intense external magnetic field. Excitation of transitions between these states is stimulated using radio-frequency (RF) electromagnetic radiation.

1.2.1.1 Magnetic Dipole

The nuclear magnetic dipole moment arises from the *spin angular momentum* of the nucleus. All nuclei with an odd mass number (e.g. ^{1}H, ^{13}C, ^{15}N) have spin angular momentum because they have an unpaired proton. All nuclei with an even mass number and an odd charge (e.g. ^{2}H, ^{14}N) also have spin angular momentum.

The spin angular momentum, $\vec{S}$, is quantized (as is all angular momentum) and the different quantum states are indexed with the spin quantum number I. The total angular momentum of a nuclear spin is: $\vec{S} = \hbar\sqrt{I(I+1)}$. We will generally be interested in the z-component of the angular moment, S_z, which is restricted to integral steps of $\hbar$ ranging from $-I$ to $+I$. For example, a spin one-half nuclei would have two possible values of S_z: $+\frac{1}{2}\hbar$, and $-\frac{1}{2}\hbar$, corresponding to spin quantum numbers $m_z = +\frac{1}{2}$ and $m_z = -\frac{1}{2}$, respectively. The magnetic moment of a nuclear spin, $\vec{\mu}$, is proportional to its spin angular momentum, $\hbar\vec{I}$ by a factor, γ, which has units of $radians\ sec^{-1}\ gauss^{-1}$.

$$\vec{\mu}_n = \gamma_n \hbar \vec{I} \tag{1.2}$$

The magnitude of γ depends on the type of nuclei. NMR properties of various

[1] This is one form of Heisenberg's uncertainty principle: $\Delta E \Delta t \geq \hbar/2$.

Table 1.2. Properties of NMR Active Nuclei.

Nuclei[1]	$\gamma(rad \cdot sec^{-1} \cdot gauss^{-1})^{\dagger}$	*I*	*Natural Abundance (%)*
^{1}H	26,753	1/2	99.980
^{2}H	4,106	1	0.016
^{19}F	25,179	1/2	100.000[2]
^{13}C	6,728	1/2	1.108[3]
^{15}N	-2,712	1/2	0.37[3]
^{31}P	10,841	1/2	100.00

[1]The term "Protons" is used interchangeably with ^{1}H in the text.
[2]Fluorine is not normally found in biopolymers, therefore it has to be introduced by chemical or biosynthetic labeling.
[3]These isotopes of carbon and nitrogen are normally found in low levels in biopolymers, therefore the levels of these two spins are generally enriched, often to 100%, by biosynthetic labeling.
[†]CGS units.

nuclear spins, including values of γ, are shown in Table 1.2. NMR active isotopes of hydrogen, carbon, nitrogen, and phosphorus exist, thus it is possible to observe NMR signals from virtually **every** atom in biopolymers. Protons (^{1}H) and phosphorus are highly abundant in natural biopolymers, while in the case of carbon and nitrogen it is usually necessary to introduce the appropriate isotope into the sample (see footnote 4 in Table 1.2). Also note, that with the exception of deuterium (^{2}H), all of these nuclei have a z-component of the spin angular momentum of $\hbar/2$. Consequently, the material presented in this text applies to all of the above atomic nuclei, except for deuterium. Deuterium with a spin quantum number $I = 1$ is a quadrapolar nuclei and in certain instances needs to be treated differently than spin-1/2 nuclei.

1.2.1.2 Transition Energies - Nuclear Dipole-Magnetic Field Interaction

When the orientation of a collection of nuclear spins is observed in the absence of a magnetic field, all possible orientations of the magnetic dipole are possible (see Fig. 1.3). However, once the spins are placed in a magnetic field, the direction of z-axis becomes defined by the direction of the field, and the magnetic moments of spin-1/2 nuclei assume two orientations, either along or opposed to the magnetic field, as illustrated in Fig. 1.3. Note that the magnetic moments cannot orient parallel to the magnetic field because of the restrictions placed on the value of μ_z by the quantum mechanical properties of the system.

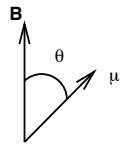


Figure 1.2. Interaction of magnetic dipole with an applied field.

The energy of a state depends on the interaction of the aligned magnetic dipole with an externally applied magnetic field. The size of this interaction

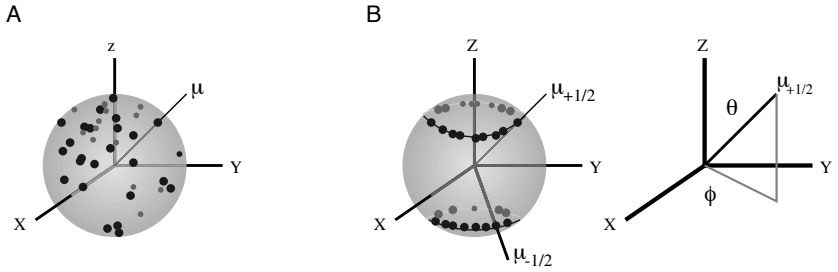


Figure 1.3. Orientation of Magnetic Dipoles.
A. Orientation of nuclear magnetic dipoles in the absence of a magnetic field. A unit sphere is shown and the dots on the surface illustrate the various orientations of the dipoles in space. The orientation of one dipole is indicated by a line drawn from the center of the sphere.
B. The orientation of the nuclear spin dipoles in a static magnetic field along Z. Note that ≈ one-half of the spins are pointed up and the other half are pointed down. Also note that they can assume any value of ϕ, but only two values of θ. ϕ and θ represent the orientation of the magnetic dipole in spherical coordinates, as shown on the right part of this figure.

can be found from classical electromagnetic theory. For example, consider a magnetic dipole in a static magnetic field, $\vec{B}$, with the magnetic field along the z-axis, as shown in panel B of Fig. 1.3. The energy required to change the angle, θ, is

$$\begin{aligned} E &= \int \Gamma d\theta = \int (\vec{\mu} \times \vec{B}) d\theta = |\mu||B| \int \sin(\theta) d\theta = -|\mu||B| \cos(\theta) \\ &= -\vec{\mu} \cdot \vec{B} \end{aligned} \tag{1.3}$$

The torque, Γ arises as the static field attempts to align the magnetic dipole of the nucleus. The energy of interaction, or Hamiltonian [2], between the field and the dipole is therefore given by the dot product of the two vectors: $\mathcal{H} = -\vec{\mu} \cdot \vec{B}$.

The actual magnetic field that is present at the nucleus is usually attenuated, or shielded, by the presence of electrons that surround the nucleus, giving a modified field at the nucleus, B:

$$B = (1 - \sigma) B_o \tag{1.4}$$

where σ represents the degree of shielding. An extensive discussion of shielding effects is found in Section 1.3.

Assuming that the magnetic field is along the z-axis, the energy of each state is:

$$\mathcal{H} = -u_z B_z \tag{1.5}$$

[2]The Hamiltonian is a quantum mechanical operator that gives the energy of the system when applied to the wavefunction of the system.

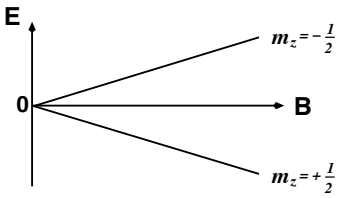


Figure 1.4. Energy of the two quantum states of a spin $\frac{1}{2}$ particle, with $\gamma > 0$, as a function of the magnetic field strength. The diagram for a spin with $\gamma < 0$ would be identical, except that the quantum number would be interchanged.

At this point, we have a result that is entirely classical, it is applicable to *any* magnetic dipole that is placed into a magnetic field. To relate the energy of a nuclear spin to its quantum state, we make use of the relationship between the magnetic dipole and the z-component of the spin angular momentum, $\mu_z = \gamma\hbar m_z$, giving the energy for a spin in a quantum state, m_z:

$$\mathcal{H} = -\gamma\hbar B m_z \tag{1.6}$$

This is often referred to as the Zeeman Hamiltonian.

Using Eq. 1.6 it is possible to draw an energy diagram for the system as a function of magnetic field. For a spin-1/2 particle ($m_z = \pm 1/2$) the energy as a function of the magnetic field is shown in Fig. 1.4.

The ground, or lower energy state, referred to as α, corresponds to $m_z = +1/2$, and the excited, or higher energy state, is referred to as β, corresponds to $m_z = -1/2$. The energies of the two states can be calculated using Eq. 1.6,

$$E_\alpha = -\frac{\gamma\hbar B}{2} \qquad E_\beta = +\frac{\gamma\hbar B}{2} \tag{1.7}$$

The energy difference between the two states is easily computed,

$$\Delta E = E_\beta - E_\alpha = \gamma\hbar B \tag{1.8}$$

and using the relationship, $E = \hbar\omega$, gives the well known Larmor equation:

$$\omega_\mathbf{s} = \gamma\mathbf{B} \tag{1.9}$$

In the above equation, ω_s refers to the absorption, or resonance, frequency of the shielded nucleus, i.e. its observed resonance frequency. The Larmor equation is one of the *key* equations in NMR spectroscopy, it states that the absorption frequency of a transition is equal to γ multiplied by the strength of the magnetic field *at the nucleus*.

The energy of an NMR transition is quite low, requiring radiowaves to excite the spins. The small value of ΔE has two important consequences:

1. The population difference between the two energy levels is very small, on the order of 1 part in 10^6. The actual population difference can be easily calculated from Boltzmann's relationship:

$$\frac{N_\beta}{N_\alpha} = e^{\frac{-\gamma\hbar B}{kt}} \approx 1 - \frac{\gamma\hbar B}{kt} \tag{1.10}$$

   The consequence of having a small population differences is that NMR spectroscopy is a relatively insensitive experimental technique because of the small excess of spins in the ground state. In any form of spectroscopy the presence of electro-magnetic radiation induces transitions from the ground to the excited state and vice versa, consequently, the *net* absorption depends on the population difference between the two states. Due to the low sensitivity, it is common to increase the signal-to-noise of the spectrum by signal averaging. In addition, typical NMR experiments require protein concentrations on the order of 1 mM. However, in some cases concentrations in the range of 50 μM have be used.

2. The lifetime of the excited state can be quite long, on the order of *m*sec to sec. As discussed above, a long lifetime provides three benefits: narrow resonance lines, experimental manipulation of the excited state in multi-dimensional experiments, and sensitivity to molecular motion over a wide time scale.

1.3 Chemical Shielding

The magnetic field strength at the nucleus differs slightly from the applied field, B_o, because of shielded by the electron density surrounding the nucleus. This shielding is due to precession of electrons under the influence of the applied magnetic field, which generates an additional magnetic field that usually opposes the externally applied magnetic field.

The magnetic field from the electrons shields the nucleus from B_o, and results in a local magnetic field strength at the nucleus that is given by:

$$B = (1 - \sigma) \cdot B_o \tag{1.11}$$

where *sigma* represents the shielding of the nuclear spin. For an isotropic electron distribution the shielding is given by the Lamb formula [**?**]:

$$\sigma = \frac{e^2}{3mc^2} \int \frac{\rho(r)}{r} dr \tag{1.12}$$

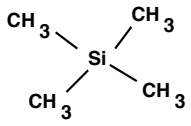


Figure 1.5. Tetramethyl silane (TMS) is often used as a reference compound. The Si atom is electropositive with respect to carbon, therefore the electron density on the methyl groups is higher than would be found on the equivalent hydrocarbon. The high electron density shields the methyl carbon and protons, leading to a lower effective field at the nucleus and a lower resonance frequency.

As the electron density around the nucleus increases, the effective field decreases, leading to lower resonance frequencies. Since the resonance frequency is due to the chemical environment of the nuclear spin, the observed frequency is referred to as a *chemical shift*. Due to differences in shielding, different spins will experience different values of local magnetic field, giving rise to shifts in their frequencies. Chemical shift reference standards, such as tetra-methyl silane were chosen because their protons are highly shielded (see Fig. **??**).

For anisotropic electron distributions the shielding is described by a tensor. A tensor is a concise mathematical expression of the anisotropic properties of a physical system in three-dimensional space and has the following form[3]

$$\sigma = \begin{bmatrix} \sigma_{xx} & 0 & 0 \\ 0 & \sigma_{yy} & 0 \\ 0 & 0 & \sigma_{zz} \end{bmatrix} \tag{1.13}$$

Equation 1.13 describes the chemical shielding if the magnetic field were along the x-axis (σ_{xx}), y-axis (σ_{yy}), or the z-axis (σ_{zz}). Under conditions of rapid tumbling, which is generally the case in solution, an averaged shielding is observed:

$$\bar{\sigma} = \frac{1}{3}\left[\sigma_{xx} + \sigma_{yy} + \sigma_{zz}\right] \tag{1.14}$$

1.3.1 Chemical Shift Scale - ppm

Equation 1.9 indicates that the observed absorption frequency depends on the magnetic field strength. Commercial NMR spectrometers can be purchased with different magnetic field strengths. A field strength that gives a proton absorption frequency of 500 MHz (11.7 Tesla) is fairly common. However, spectrometers with proton frequencies ranges as high as 900 MHz are becoming more common. To faciliate the comparison of spectra obtained with different field strengths, the effect of the field strength is removed by converting all frequencies to a dimensionless scale, the chemical shift scale. This scale is defined as:

$$\delta = \frac{\nu - \nu_o}{\nu_o} \times 10^6 \tag{1.15}$$

with units of ppm, or parts-per-million. The conversion from frequency to chemical shift makes the position of the spectral line independent of the magnetic field strength (by dividing by ν_o).

The constant, ν_o, is a reference frequency, in units of Hertz (Hz). It is often the frequency of the line from a reference compound whose resonance is at one end of the spectrum. For example, tetra-methyl silane is used to reference organic samples and the proton and carbon frequencies of its spectral line are set to zero ppm. In the case of protein solutions the water line can be used as

[3]This simple form, with all off-diagonal elements having the value of zero, is only found for one particular orientation of the molecule with respect to the magnetic field, called the principle axis system (PAS).

Table 1.4. Nitrogen chemical shifts for side-chain atoms. The amide nitrogen chemical shifts are ≈ 120 ppm, with the exception of Glycine, which is found at 109.9 ppm. Data from BioMagResBank [**?**].

Residue	*Shifts*
Arg	89.8 (ϵ), 74.8 NH1, 75.8 NH2
Asn	112.8 (δ)
Gln	111.8 (ϵ)
His	190.7 (δ1), 179.8 (ϵ2)
Lys	71.86 (ζ)
Trp	129.5 (ϵ)

an approximate proton chemical shift reference point, with a chemical shift of about 4.70 ppm.

1.4 Characteristic ^{1}H, ^{13}C and ^{15}N Chemical Shifts

The proton, carbon, and nitrogen chemical shifts found for amino-acids in proteins are presented in tables 1.3, 1.5, 1.4.

Table 1.3. Average proton chemical shifts in proteins. Data from BioMagResBank

Residue	*NH*	H_α	H_β	*Others*
Gly	8.34	3.94		
Ala	8.20	4.26	1.38	
Val	8.29	4.16	1.99	0.84, 0.83(CH3)
Ile	8.26	4.20	1.80	1.30, 1.24 (CH2), 0.80 (γCH3), 0.70 (δCH3)
Leu	8.22	4.32	1.63,1.57	1.54 (γCH), 0.77, 0.76(δCH3)
Pro	-	4.41	2.05,2.05	1.93 (γCH2), 3.64, 3.63 (δCH2)
Ser	8.29	4.51	3.88	5.33 Hγ (OH)
Thr	8.27	4.48	4.17	1.16 (γCH3), 4.40 Hγ1 (OH)
Asp	8.33	4.61	2.74,2.70	
Glu	8.34	4.26	2.04	2.31 (γCH2)
Lys	8.22	4.28	1.79,1.78	1.38 (γCH2), 1.61 (δCH2), 2.93 (ϵCH2), 7.52 (ζNH3)
Arg	8.24	4.27	1.79	1.58 (γCH2), 3.13 (δCH2), 7.32, 6.74, 6.72 (NH)
Asn	8.37	4.70	2.80,2.78	7.27, 7.20 (δNH2)
Gln	8.22	4.28	2.05,2.04	2.32 (γCH2), 7.17, 7.07 (γNH2)
Met	8.26	4.39	2.03,2.01	2.44 (γCH2), 1.86 (ϵCH3)
Cys	8.42	4.73	2.95,2.98	1.66 -SH
Trp	8.35	4.74	3.32,3.18	6.68-7.17 (aromatic), 10.13 (NH)
Phe	8.42	4.62	2.97,2.99	6.89-6.91 (aromatic)
Tyr	8.37	4.63	1.91	6.86 (Hδ), 6.64 (Hϵ), 9.25 (-OH)
His	8.25	4.62	3.11,3.12	Hδ1 10.14(NH), Hδ2 7.08, Hϵ1 8.08, Hϵ2 10.43(NH)

Table 1.5. Carbon chemical shifts from the BioMagResBank [**?**]. Carbonyl shifts have been omitted since they are quite uniform at approximately 175 ppm.

Residue	C_α	C_β	*Others*
Gly	45.3		
Ala	53.1	18.9	
Val	62.5	32.6	21.3 (CH3)
Ile	61.6	38.6	27.6 (γ1), 17.3 (γCH3), 13.4 (δCH3)
Leu	55.7	42.3	26.8 (γ), 24.5 (δCH3)
Pro	63.3	31.8	27.1 (γ), 50.3 (δ)
Ser	58.6	63.8	
Thr	62.1	69.6	21.4 (γCH3)
Asp	54.5	40.7	178.41 (γ) sidechain
Glu	57.4	30.0	36.0 (γ), 181.9 (δ) sidechain
Lys	56.8	32.8	24.9 (γ), 28.8 (δ), 40 (ϵ)
Arg	56.9	30.7	27.3 (γ), 43.1 (δ), 159.0 (ζ)
Asn	54.5	40.7	178.41 (γ) sidechain
Gln	56.6	29.1	33.7 (γ), 179.7 (δ) sidechain
Met	56.1	32.9	32.1 (γ), 17.2 (ϵCH3)
Cys	57.4	34.1	
Trp	57.7	30.1	110-137 (aromatic)
Phe	58.2	40.0	129-138 (aromatic)
Tyr	58.0	39.1	117 (ϵC), 132 (δC), 156 (ζ)
His	56.4	30.0	119.8 (δ2), 136 (ϵ1)

1.4.1 Effect of Electronic Structure on Chemical Shifts

The chemical shifts presented in tables 1.3 and 1.5 are clearly different from atom to atom. For example, amide protons resonate at $\approx$ 8 ppm, H_α protons at $\approx$ 4 ppm and methyl protons at $\approx$ 1 ppm. A similar trend in carbon shifts is observed for α- and β-carbons. These trends in chemical shifts can be explained by the electronegativity of the atoms that are chemically bonded to the atom of interest. The amide proton has a high chemical shift because the nitrogen atom is more electron withdrawing than carbon. The reduced electron density at the amide proton decreases the shielding and therefore increases the effective field and consequently the resonance frequency. Similarly the H_α shifts are higher than the methyl-H shifts because of the proximity of the α-protons to the electronegative nitrogen.

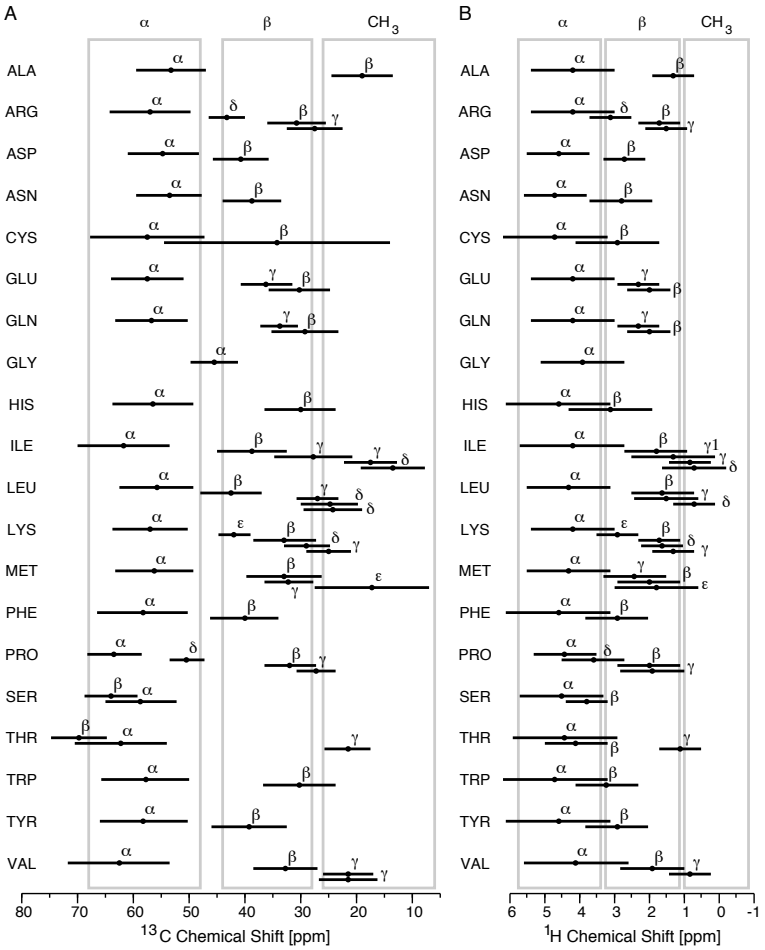


Figure 1.6. The distribution of observed carbon (A, left) and proton (B, right) chemical shifts in proteins. The solid circles (•) mark the average chemical shift. The solid lines indicate $\pm 3\sigma$; 95% of the observed chemical shifts fall within this range. The gray boxes indicate nominal chemical shift ranges for α, β, and methyl atoms. In the case of carbon shifts, these range separate the atom types quite well. Note that there are a few exceptions, for example, the β-carbons of Ser and Thr fall in the α-region and the α-carbon of Gly can fall in the β-carbon region. The large range of β-carbon shifts for Cys is due to the fact that both free and disulfide bonded residues are included in this figure. In the case of proton shifts, the separation is not as clean due to the extensive chemical shift overlap between the various atom types. Data from the BioMagResBank database of chemical shifts [**?**].

Note that within a residue, the relationship between atom type and chemical shift is similar for both carbon and proton shifts. For example, in the case of Arginine the following ordering is found: $\alpha > \delta > \beta > \gamma$ for both carbon and

proton shifts (see Fig. 1.6. The identical ordering reflects a similar electronic environment for both the carbon and its attached proton.

The conformation, or secondary structure, of the polypeptide backbone alters the chemical shift of chemically equivalent atoms in a predictable fashion. For example, the C_α shift of Alanine decreases by 1.3 ppm when this residue is found in a β-strand, and increases by 2.3 ppm when found in an α-helical configuration. Although the secondary structure of the backbone is only one of the factors that alter chemical shifts, it is possible to predict secondary structure by correlating the deviation of the chemical shift from random-coil values for a number of different atoms (e.g. C_α, H_α, etc).

In addition to electronegativity effects, a formal positive change (e.g. Lys ϵN) also withdraws electrons from adjacent atoms, decreasing the shielding and therefore increasing the chemical shift. A formal negative change, will have the opposite effect.

1.4.2 Ring Current Effects

Chemical shifts are also perturbed by additional magnetic fields that arise from the precession of delocalized electrons in conjugated systems, such as aromatic rings or carboxylic acid groups. These *ring current effects* can be substantial. Several methods have been devised to calculate the effect of ring-currents on chemical shifts. The simplest method, introduced by Pople [**?**], is to consider the induced magnetic field as a point dipole at the center of the aromatic ring. The magnetic field that is generated by this dipole is given by

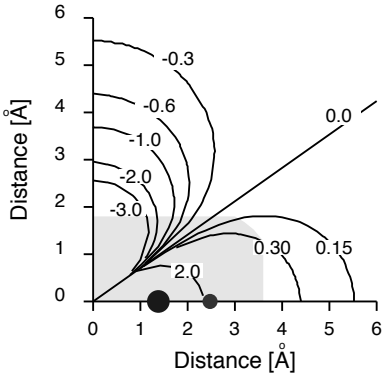


Figure 1.7. Calculated ring current shifts for a Phenylalanine ring. The x-axis lies in the plane of the ring and the y-axis is perpendicular to the plane of the ring. The location of the carbon and its attached hydrogen are indicated by the large and small spheres, respectively. The large gray area represents the approximate Van der Waals radius of the phenyl group. The lines represent contours of iso-chemical shift changes.

the standard dipole equation:

$$\sigma = iB\frac{1 - 3cos^2\theta}{r^3} \tag{1.16}$$

where i is a ring-current factor that depends on the geometry of the ring. For the phenyl group $i = 1$. B is a constant of proportionally, on the order of 25 ppm. θ is the angle between the normal to the aromatic ring and the vector that joins the atom to the origin.

This function is plotted in Fig. 1.7. The ring-current effect depends on both distance and orientation. The weak distance dependence, of $1/r^3$, implies that ring current effects can significantly perturb the chemical shifts of near-by residues. Chemical shift perturbations also depend on the orientation of the atom with respect to the plane of the aromatic ring. For example, placement of an atom 4 directly above the center of the ring will decrease the chemical shift by approximately 0.75 ppm. In contrast, if the atom is placed in the plane of the ring at the same distance from the center, its chemical shift will be increased by approximately 0.33 ppm. Aromatic protons, because of their close proximity to the center of the ring, experience an increase of $\approx$ 2 ppm due to ring current effects.

1.4.3 Effects of Local Environment on Chemical Shifts

The above discussion suggest that chemically equivalent atoms will have identical chemical shifts. For example, one might expect that all of the Alanine methyl protons in a protein to resonance at 1.39 ppm. Fortunately, the tertiary structure in folded proteins generates sufficient diversity in the local environment such that different resonance frequencies are observed for otherwise identical groups. The environmental differences that cause diversity in chemical shifts include electro-negativity effects (e.g. in hydrogen bonded amide protons), as well as the electrostatic and magnetic (ring current) effects that were discussed previously. Typical ranges of carbon and proton chemical shifts for aliphatic atoms in a large number of proteins are shown in Fig. 1.6. The carbon chemical shifts are more disperse, covering a range of approximately 70 ppm, while the proton chemical shift range is about 6 ppm.

1.4.3.1 Degeneracy and Equivalent Chemical shifts

In a number of cases, two or more spins (e.g. protons) will have identical chemical shifts. These spins are said to have degenerate chemical shifts.

Chemical shift degeneracy occurs when the two protons are magnetically equivalent. A simple test for magnetic equivalence is to replace each of the two protons with a test atom, e.g. F, generating two new compounds, and determine whether the two protons are magnetically equivalent using the following steps:

1. If two different compounds are generated, then the protons are non-equivalent and will show two separate peaks.

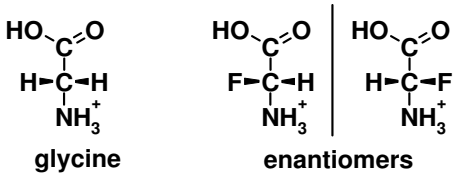


Figure 1.8. Testing for magnetic equivalence of glycine α-protons. Replacement of each α-proton with fluorine generates the two compounds on the right. These are enantiomers and therefore will show identical chemical shifts in an achiral environment. spectroscopy.

2. If the two molecules can be superimposed, then they are magnetically equivalent.
3. If the two molecules are mirror images of each other (enantiomers), then the two protons will have the same chemical shift in an achiral environment. However, two different shifts can be observed in a chiral environment, such as in biological polymers.
4. If the two molecules are diastereoamers, then the two protons are in different magnetic environments and will show different chemical shifts.

As an example of the second item in the above list, consider the methyl protons of alanine. Replacement of each methyl proton with F will generate three different derivatives. However, these can all be superimposed on each other by rotation about the alpha-beta bond, so all three protons are equivalent. A similar analysis shows that the delta and epsilon protons on tyrosine and phenylalanine are also equivalent due to rotation of the ring.

As an example of the third item on the above list, consider the two alpha protons on glycine. When each of these is replaced by F, a chiral center on the alpha carbon is generated, and the two compounds are mirror images of each other. Consequently, the chemical shifts of the two alpha protons in glycine will be identical in water, an achiral solvent. However, if glycine is incorporated into a protein, it is likely that the environment is no longer achiral and separate chemical shifts will be observed for each proton.

Although two protons may be magnetically equivalent in a formal sense, restriction of rotation can produce non-equivalency if the rotation was required to generate equivalency. For example, if the rotation of methyl groups on Ala, Val, Leu, Ile, or aromatics on Phe or Tyr sidechains, is somehow restricted, such that the methyl or aromatic protons are now in different environments, then distinct resonances will be seen for each proton. In order for this to occur the rate of rotation must be significantly slower than the difference in resonance frequencies of the protons in the different environments. In the case of aromatic residues, the rate of ring-flipping is often sufficiently slow that individual resonances are seen for each of the δ and ϵ protons. In the case of methyls, the activation barrier for rotation is usually so low that rapid rotation does occur, giving rise to degenerate shifts for all three methyl protons.

1.4.4 Use of Chemical Shifts in Resonance Assignments

A critical problem in applying NMR techniques to proteins is that of resonance assignment, or determining an atom's chemical shift. Once this has been accomplished the information obtained from the NMR resonance signal, such as relaxation rates, can be associated with a particular atom in the protein. During the assignment process resonance signals that belong to the same amino acid are identified by a number of different NMR experiments that are discussed in subsequent chapters. This collection of resonance signals is called a *spin-system*. Identification of the residue type of a spin-system is helpful in the assignment process. In particular, if a residue only appears once in the primary sequence of a protein, then identification of the spin-system leads directly to assignment. The residue type of a spin-system can ofter be discerned from the number of resonances and their chemical shifts. Inspection of Fig. 1.6 shows that there is a cleaner distinction between the type of atom and its carbon chemical shift versus the corresponding proton shift. For example, the chemical shift ranges for the α-carbons of Trp, Tyr, and Val do not overlap the range for their β-carbon shifts. In contrast, these ranges overlap in the case of proton shifts. Therefore, carbon chemical shifts are generally more reliable for predicting the atom type in a spin-system.

1.4.5 Chemical Shift Dispersion & Multi-dimensional NMR

Another problem that we will address is how to generate resolved NMR spectra from complex biopolymers, such as proteins. The chemical shift ranges shown in Fig. 1.6 indicate that the NMR spectra of a polypeptide of modest size, say 50 residues, will have a complex NMR spectrum that will contain many overlapping peaks. One solution to this problem has been to increase the dimensionality of the NMR experiment, such that the positions of peaks are defined by two or more resonance frequencies. Two-, three- and four-dimensional experiments are routinely performed on isotopically labeled proteins and nucleic acids. Multidimensional NMR experiments of this type are possible because of spin-spin coupling between spins.

Chapter 2

SPIN-SPIN COUPLING

Scalar couplings arise from spin-spin interactions that occur via bonding electrons. Consequently, they provide information on the chemical *connectivity* between atoms. Therefore, these couplings can be utilized to correlate NMR resonances of atoms that are chemically bonded to one another, providing chemical shift assignments if the molecular structure is known. In particular, the scalar coupling between the amide nitrogen and the carbonyl carbon of the preceeding residue permits the linkage of spins on one amino acid to those of its neighbor.

In addition to providing information on chemical connectivities, the sizes of three bond scalar couplings are sensitive to the electron distribution of the intervening bonds, consequently these couplings provide information on the conformation of chemical groups in proteins.

In this chapter we will first explore the origin of scalar couplings between nuclear spins, understanding the effect of this coupling on the resultant NMR spectrum from a classical perspective. The coupling will then be analyzed using quantum mechanics to fully evaluate the effect of the coupling on the frequency and intensity of resonance lines in the NMR spectrum of coupled spins. Finally, a density matrix treatment of coupled spins will be introduced in the subsequent Chapter as a prelude to analyzing the effect of scalar coupling in more complex multi-dimensional NMR spectra.

2.1 Scalar Coupling

Scalar, or J-coupling, occurs between nuclei which are connected by chemical bonds. The coupling causes splitting of the spectral lines for both coupled spins by an amount J, commonly referred to as the coupling constant (See Fig. 6.7). The nomenclature that is used to describe a coupling is as follows:

$$^{n}J_{AB}$$

where n refers to the number of intervening bonds, and A and B identify the two coupled spins. For example the coupling constant between the amide nitrogen and the C_β carbon would be written as: ${}^2J_{N\ C_\beta}$. The value of J is usually given in Hz and is the observed frequency separation between the resonance lines of the coupled spins (see below). A resonance line that is split due to J-coupling is generally referred to as a multiplet. The spectrum shown in Fig. 6.7 is an example of a doublet. If the resonance line was split into three signals, it would be called a triplet. Finally splitting into four lines generates a quartet.

The effect of J-coupling on the spectrum depends on the frequency separation of the coupled spins. If the two coupled spins differ greatly in their resonance frequencies ($\Delta\nu > J$), then the system is referred to as an AX system, where the X signifies the fact that the two chemical shifts are quite different. All coupling between different atom types, or heteronuclear spins, are AX couplings because of the large difference in the frequencies of coupled spins. Examples include, J_{NH}, J_{CH}, and J_{NC}. AX couplings can be analyzed using a classical analysis, similar to that depicted in Fig. 6.7. When two coupling spins have nearly equivalent resonance frequencies ($\Delta\nu \leq J$) then the system is referred to as an AB system. For example, the coupling between two H_β protons on an amino acid is an example of an AB system. Accurate analysis of AB systems require a detailed quantum mechanical treatment. Lastly, when the coupled spins have the identical resonance frequencies, the observed coupling disappears entirely. This is most often seen when multiple protons have equivalent environments, such as the three protons on a methyl group.

2.1.1 Origin of Scalar Coupling

Scalar coupling arises from the interaction of the nuclear magnetic moment with the electrons involved in the chemical bond. The nuclear spin polarization of one atom affects the polarization of the surrounding electrons. The electron polarization subsequently produces a change in the magnetic field that is sensed by the coupled spin. For example, consider a C-H group in a molecule, as illustrated in Fig. 6.7. The proton nuclear spin polarizes the electron in the σ bonding orbital. This polarization alters the magnetic field at the carbon nucleus. Since there are two possible spin states for the proton magnetic dipole, the effective field at the carbon nucleus is increased or decreased, depending on the spin state of the attached proton. Since the population difference between the two orientations of the proton spin are approximately equal, one-half of the attached carbons will experience an increase in the local magnetic field while the other half will experience a decrease. This difference in the local magnetic field at the carbon nucleus will lead to a shift of the carbon resonance frequency. Since there are two possible proton spin states, the carbon spectrum is split into two lines, with the separation between the lines equal to the J-coupling.

The change in the magnetic field induced by the proton spins is proportional to γ_H. The corresponding change in the carbon resonance frequency is proportional to the product of this field change and the gyromagnetic ratio of the

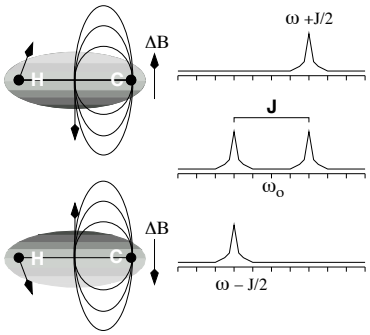


Figure 2.1. Nuclear Spin Coupling in a ^{13}C-H group. Two C-H groups in separate, but otherwise identical molecules, are shown. The two molecules differ only in the spin state of the proton. The bonding electrons in the σ orbital become polarized from the proton magnetic dipole, as indicated by the oval field lines. The polarization of the electrons can either increase or decrease the magnetic field at the ^{13}C nucleus, depending on the proton spin state. One orientation of the proton spin increases the magnetic field (upper molecule) while the other orientation causes a decrease in the apparent field (lower molecule), causing either an upfield or downfield shift in the resonance line. The observed carbon spectrum is the sum of these two resonances, giving two peaks that are separated by the J-coupling constant, as shown in the middle spectrum.

carbon spin, i.e.:

$$\Delta\omega \propto \pm\gamma_H\gamma_C \tag{2.1}$$

The effect of the carbon spin on the proton spin can be calculated in the same way. The change in the local magnetic field at the proton nucleus is proportional to $\pm\gamma_C$, giving rise to a frequency shift of the protons of $\pm\gamma_C\gamma_H$. Consequently, the proton spin will experience exactly the same shift in frequency as its coupled partner, giving rise to exactly the same splitting of the proton resonance line as the resonance line from the attached carbon.

Note also that the observed frequency shift only depends on the product of the gyromagnetic ratios of the coupled spins, the scalar coupling constant is independent of the applied magnetic field (B_o).

Values of J-coupling constants that are important in biomolecular NMR are shown in Table 2.1. The strength of the J-coupling depends on several factors, including the gyromagnetic ratio of the coupled spins, the number of bonds connecting the coupled spins, and the conformation of the intervening bonds in the case of multiple bond couplings. The series of single bond heteronuclear couplings (Table 2.1, right column) illustrates the effect of the gyromagnetic ratio on the coupling constant; the coupling constant increases with increasing γ. Scalar coupling through multiple bonds severely attenuates the coupling. For example, the strong single H-C coupling of 130 Hz is reduced to 5 Hz when an additional carbon-carbon bond is inserted between the two coupled spins (Table 2.1).

2.1.2 Coupling to Multiple Spins

The coupling between a carbon and a hydrogen in a ^{13}C-H group results in the splitting of both the proton and carbon spectral line by an amount J_{CH} Hz. If the carbon atom is coupled to more than one *equivalent* proton[1], such as in a $^{13}CH_2$ or $^{13}CH_3$ group, then a more complex splitting pattern is observed.

In the case of a $^{13}CH_2$ group a triplet of lines is observed in the carbon NMR spectrum, as illustrated in Fig. 2.2. This pattern arises because there are four possible combinations of the spin-states of the two coupled protons and the local magnetic field at the carbon nucleus will be the sum of the field changes induced by each proton. When the magnetic moment of both protons point upwards, in the direction of B_o, the frequency shift of the resonance line will be $2 \times J/2$ Hz, or J Hz. When the proton spins are both pointing in the opposite direction the shift is $-J$ Hz. When the direction of each proton spin are opposite to each other, the local field change at the carbon cancels, resulting in a zero frequency shift of the carbon spin. The intensity of the lines in the carbon spectrum is proportional to the number of molecules in the sample having one of the four possible proton spin-states. Since the state in which the proton spins point in opposing directions occurs twice as frequently than the other two states, the central line of the triplet will have twice the intensity of the outer lines, giving an observed intensity ratio of 1:2:1, as illustrated in Fig. 2.2.

The coupling between a carbon atom and three equivalent protons, such as in a methyl group ($^{13}CH_3$), can be analyzed in exactly the same way. The change in the local field that occurs when the magnetic dipoles from all of the protons are aligned in the same direction is $\pm\frac{3}{2}J$. When the magnetic dipoles of two protons are oriented in the same direction, while the third is pointing in the opposite direction, the frequency shift is $\pm\frac{1}{2}J$ since the opposing pair cancel

[1]Equivalent protons are generally considered to be a collection of protons that are attached to a single carbon atom and have the same chemical shift. Equivalency is most often a result of free rotation of the group, which averages the local environments of all of the protons.

Couplings Involving Heteronuclear (^{13}C or ^{15}N) Spins		*Proton-Proton Couplings*	
C-N	14 Hz	H-C-H	-12 to -15 Hz
C-C	35 Hz	H-C-C-H	2-14 Hz
H-N	92 Hz	H-C=C-H	10 (cis)/17 (trans)
H-C	130 Hz	H-N-C-H	1-10 Hz
H-C-C	5 Hz (two bond coupling)		(3 Hz α-helix)
			(10 Hz β-strand)

Table 2.1. Homonuclear (proton-proton) and heteronuclear coupling constants that are commonly found in biopolymers. The values in this table are approximate; the coupling constants will also be affected by the electronic environment of the associated spins.

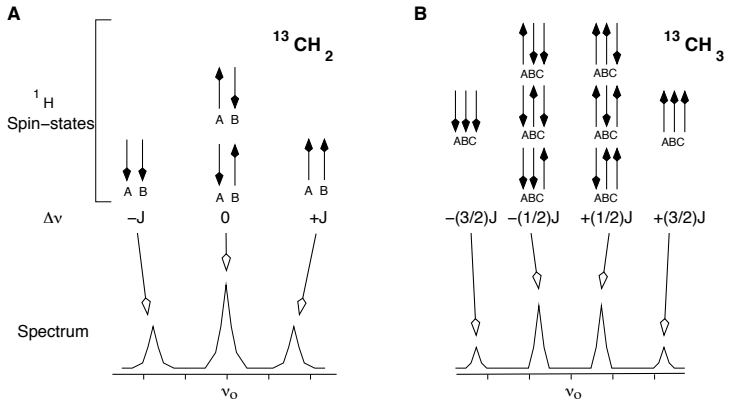


Figure 2.2. Effect of scalar coupling to two (A) or three (B) equivalent protons on the carbon spectrum. The possible arrangements of two (labeled A B) or three protons (labeled A B C) are shown in the upper part of each panel (^{1}H spin-states). The resultant shift in the frequency of the attached carbon is indicated by $\Delta\nu$. The final carbon spectrum is shown in the lower part of each panel. In both cases the splitting, or separation between the lines is equal to $^1J_{CH}$. The intensity of each line depends on the number of molecules in the sample with a particular spin state; a 1:2:1 ratio will be found for two coupled protons and a 1:3:3:1 ratio is found for three coupled protons.

each other's effect on the local magnetic field at the carbon nucleus. Again, the relative intensity of each line is proportional to the number of atoms that give a particular frequency shift, in this case the four lines in the quartet will have a relative intensities of 1:3:3:1.

The effect of coupling to multiple spins on an NMR spectral line can be easily obtained from Pascal's triangle, as illustrated in Fig. 2.3. Each row of the triangle indicates the location of each line in the multiplet as well as the relative intensity of each line.

In cases where an atom is coupled to two different, or non-equivalent, spins, then the couplings are treated independently. For example, the carbonyl carbon

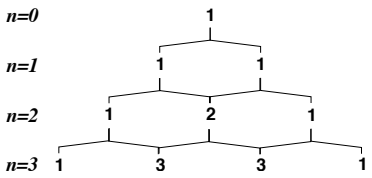


Figure 2.3. Pascal's triangle can be used to readily evaluate the effect of coupling to multiple equivalent spins on the appearance of a resonance line. The top of the triangle represents the resonance line from a spin with no coupling partner ($n = 0$) and each subsequent row represents the spectra that would be obtained as a result of coupling to one, or more, additional spins.

is coupled to both the amide nitrogen ($^1J_{N\ CO} \approx 12$ Hz) as well as the alpha carbon, ($^1J_{CO\ C_\alpha} \approx 55$ Hz), consequently the spectra line from the carbonyl will be a quartet, showing both couplings (see Fig. 2.4).

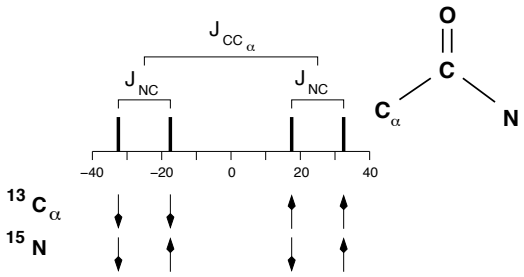


Figure 2.4. Scalar coupling to two non-equivalent spins. The spectrum of a carbonyl carbon is shown. A quartet is observed because the coupling to the alpha carbon (J_{CC_α}) is larger than the coupling to the nitrogen (J_{NC}). The spin-states of the alpha carbon and nitrogen are shown below the spectrum. Note that these four states are equally likely, therefore the intensity of the lines in the quartet is 1:1:1:1.

2.1.3 Strong Coupling (J≈ $\Delta\nu$)

When the chemical shift difference between the two coupled spins is of the same order as the coupling constant then it is necessary to perform a more complete quantum mechanical treatment. This analysis shows that in addition to simple splitting of the resonance lines, there are also intensity changes that occur for each member of the multiplet. As the frequency separation between the coupled spins decreases, the intensity of the outer lines of the multiple decrease in intensity. When the coupled spins have the same resonance frequency, the intensity of the outer lines becomes zero, and there is no observed effect of the coupling on the NMR spectrum, i.e. a single resonance line is observed (see Fig. 2.5).

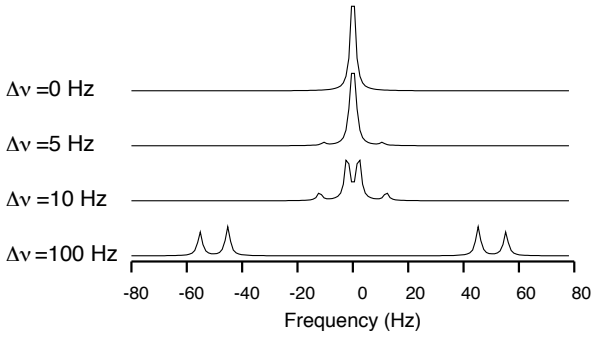


Figure 2.5. Simulated spectra are shown to illustrate the collapse of observed coupling as $\Delta\nu$ becomes smaller than J. The J-coupling constant is 10 Hz, and the separation between the lines is decreased from 100 Hz (bottom spectrum) to 0 Hz (top spectrum).

2.1.4 Conformational Effects on Scalar Coupling

In the case of multiple bond couplings, the conformation of the coupled atoms affects the coupling constant. For example, the three bond proton-proton coupling in the H-C-C-H group ranges from 2 to 14 Hz. The relationship between the coupling constant and the torsional angle is represented by the Karplus relationship [**?**]:

$$J = Acos^2\theta + Bcos\theta + C \tag{2.2}$$

where A, B, and C are empirical constants. For example, the ϕ angle in the peptide bond affects the strength of the coupling between the amide proton and the alpha proton, as illustrated in Fig. 2.6.

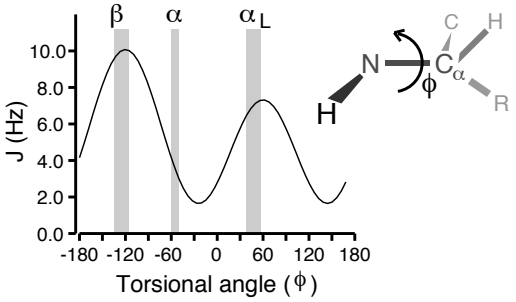


Figure 2.6. The Karplus curve for the ϕ torsional angle in polypeptides is shown on the left. The ϕ angles for regular secondary structures are indicated by the vertical gray bars. The ϕ torsional angle is defined by the relative orientation of the H-N bond vector to the C_α-CO bond vector. The molecular fragment to the right of the plot has a ϕ angle of $180^{/circ}$. The actual curve plotted is: $J = 6.98\ cos^2(\phi - 60) - 1.38\ cos(\phi - 60) + 1.72$ [**?**].

Chapter 3

ONE DIMENSIONAL NMR SPECTROSCOPY

3.1 Detection of Nuclear Spin Transitions

We have seen how placing a nuclear spin in a static magnetic field generates a ground and an excited state. Irradiation of a sample with radiofrequency (RF) waves of the appropriate frequency, $\omega_s = \gamma B$, will excite transitions from the ground to the excited state due to the interaction of the magnetic dipole with the oscillating magnetic field component of the electromagnetic radiation. This excitation field, called $\vec{B}_1$, must be orthogonal to the direction of the magnetic dipoles, such that $\vec{B}_1 \times \vec{\mu} \neq 0$, to generate transitions of the nuclear spin state. The relative orientation of the static B_0 field and the oscillatory B_1 field are shown in Fig. 3.1.

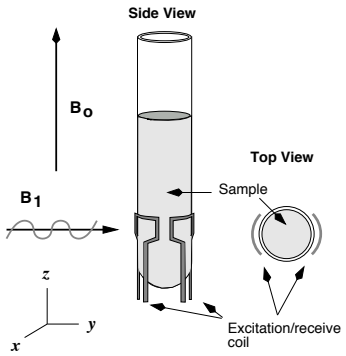


Figure 3.1. Geometry of magnetic fields around the sample. The sample is contained in a glass tube. The diameter of the sample tube is typically 5 mm and the sample volume is approximately 0.5 ml. The directions of the static B_0 field and the oscillating B_1 field are shown. The B_1 field is usually applied using two paddle shaped Helmholtz coils, shaded dark gray. These coils are also used to detect the magnetic field produced by the excited spins.

The B_1 field can be applied to the sample in one of two ways, either by scanning through multiple wavelengths (continuous wave NMR), or as a short burst of high power RF that excites a broad range of transitions (pulsed NMR). Each of these methods are discussed below.

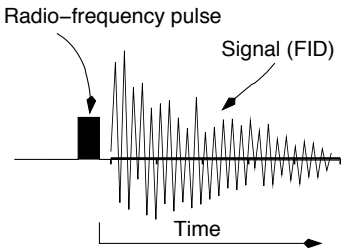


Figure 3.2. A simple one-pulse NMR experiment. The experiment begins with a short ($\approx$ 10 μsec) radio-frequency pulse. The induced signal (FID) is sampled as it evolves over time.

3.1.1 Continuous Wave NMR

In continuous wave (CW) spectroscopy the NMR spectrum is obtained by using a technique that is similar to traditional UV-visible spectroscopy, namely scanning the wavelength of the incident light and detecting the absorbance as a function of frequency. Prior to the introduction of pulsed methods in the early 1970's, all NMR spectra were acquired in this way. Low sensitivity and the general restriction to one-dimensional NMR experiments are the principal reasons why continuous wave spectroscopy is no longer used. Due to the low inherent sensitivity of NMR it is necessary to average signals. Consequently, the overall sensitivity depends on how fast each individual spectrum can be acquired. A continuous wave scan takes much longer than pulsed excitation because it is necessary to wait for the excited spins to return to the ground state while the spectrum is being scanned. Otherwise the signals from previously excited spins will interfere with the newly excited spins. In contrast, since all of the spins are excited at the same time with pulsed NMR, it is only necessary to allow the spins to relax for a single time period, between each excitation pulse.

3.1.2 Pulsed NMR

The simplest pulsed NMR experiment consists of a short RF-pulse followed by detection of the signal. This *pulse sequence* is shown in Fig. 3.2. In this experiment the nuclear spins are excited by a short burst of radiofrequency (RF) energy and the resultant excited states produce an oscillating magnetic field that induces a current in the receiver coil. In practice, the same coil that was used to excite the spins is also used to detect the signals. The induced current is measured as a function of time and is referred to the Free Induction Decay or FID. The subsequent Fourier transformation of the FID gives the normal NMR spectrum with absorption peaks at frequencies that represent the energy difference between the ground and excited states.

To understand how this procedure can generate an NMR spectrum, the motion of the spins during each segment of the experiment will be analyzed using classical mechanics. The one-pulse experiment will be divided into the

following three discrete time intervals, and the evolution of the spins within each of these periods will be discussed in detail.

Preparation: Prior to the excitation pulse the spins are at thermal equilibrium and are subject to only the static B_o field.

Excitation: During the excitation pulse the spins are subject to the static B_o field plus the oscillatory excitation field, B_1.

Detection: The excited spins precess under the static B_o field, generating the free induction decay or FID. The spectrum is obtained by Fourier transformation of the FID.

3.1.2.1 Rotating Frame of Reference

In the analysis of this experiment it is useful to consider two Cartesian coordinate systems. One that is stationary (the lab frame) and one that rotates about the z-axis (*the rotating frame*) at a frequency equal to the frequency of the applied B_1 field (ω), as shown in Fig. 3.3. The rotating frame serves two purposes in our analysis. First, it simplifies the analysis of the effect of the oscillating B_1 field on the spins. Secondly, the magnetization that is detected by the instrument is in the rotating frame.

The transformation of the coordinate system from the lab frame to the rotating frame induces a fictitious static field that is present in the rotating frame. This fictitious static field exactly cancels the applied external field. Consequently, there is no magnetic field along the z-axis in the rotating frame.

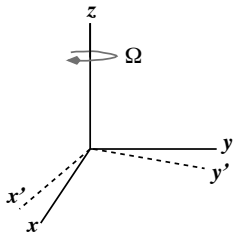


Figure 3.3. Rotating frame of reference. The coordinate system is rotating at a frequency $|\Omega|$ about the z-axis. A magnetic dipole, precessing about B_o at a frequency ω will appear stationary in the rotating frame.

3.1.2.2 Before the Pulse: Magnetization at Equilibrium

Since the nuclear spins possess angular momentum, the effect of applying *any* external field (B_o and/or B_1) to the spins is to generate a torque, Γ, on the spin. This torque will cause a change in angular momentum as described by the following classical equation:

$$\Gamma = \frac{dS}{dt} = \vec{\mu} \times \vec{B} \tag{3.1}$$

Using $\vec{\mu} = \gamma \vec{S}$ we can write

$$\frac{d\mu}{dt} = \gamma \vec{\mu} \times \ \vec{B} \tag{3.2}$$

This equation can be solved by standard methods. The result is that the magnetic dipole precesses around the static field at an angular frequency $\omega = \gamma B_{loc}$, that is, *at its resonance frequency.*

Bulk Magnetization Before progressing on to the next two segments of the pulse sequence, it is useful to consider the evolution of the bulk, or average,

magnetization during the experiment. The bulk magnetization of the sample is just the sum of the individual magnetic dipoles. The vector components of the bulk magnetization, $\vec{M}$, are defined as:

$$M_i = \sum_{i=1}^{All\ \ spins} \mu_i \tag{3.3}$$

In the presence of the static field, the sum of the z-components of each magnetic dipole will produce detectable bulk magnetization because there is a slight difference in the population of spins that are aligned in one direction versus spins aligned in the other direction (see Fig. 1.3). The net magnetization along z, referred to as the longitudinal magnetization, is therefore defined as:

$$M_z = M_o \tag{3.4}$$

In contrast, before the pulse, the distribution of the magnetic dipoles in the x-y plane is random. In other words, there is no relationship between the transverse (x-y) magnetization of one spin to another. The transverse magnetization is termed to be *incoherent*. Since the sum of a large collection of vectors aligned in random directions is zero, there is no bulk transverse magnetization at thermal equilibrium, i.e.:

$$M_x = M_y = 0 \tag{3.5}$$

3.1.2.3 Effect of the B_1 Pulse: Excitation of Nuclear Spins

The next step of the experiment, the application of the B_1 pulse, has to be considered. Assuming that the B_1 magnetic field oscillates in y-direction, it can be described as:

$$\vec{B_1} = |b_1| cos(\omega t)\hat{j} \tag{3.6}$$

where b_1 is the amplitude of the applied field, ω is its frequency, and $\hat{j}$ describes its direction. All three of these parameters; intensity, frequency, and direction, are under computer control in modern NMR instruments.

The total magnetic field in the lab frame is the sum of both the static field and the oscillating B_1 field (see Fig. 3.4). Since the rotation rate of the rotating frame is always set to the frequency of the B_1 pulse, the B_1 field is stationary in the rotating frame. Furthermore, the static magnetic field along the z-axis is canceled due to the coordinate change. Therefore, only a *static* B_1 field along the y-axis is present in the rotating frame.

The spins will precess about this field in exactly the same way they precessed about the static B_o field prior to the pulse. Recall that the precessional frequency of the spins depends on the magnetic field strength. During the pulse, the field strength is just B_1, the intensity of the applied field. Consequently, the precessional rate is:

$$\omega_1 = \gamma B_1 \tag{3.7}$$

Thus, $\vec{M}$ is tipped from the z-axis at a rate of ω_1 rad/sec. The direction of the change in $\vec{M}$, is usually given by the right-hand rule, the thumb is pointed

in the direction of the applied B_1 filed and the fingers curl in the direction of rotation.

The extent of precession about B_1 depends on both the strength of the B_1 field (ω_1) and the length of time, τ, the pulse was applied. A pulse of length τ, applied at a field strength of ω_1 rad/sec will rotate the magnetization through a "tip", or "flip", angle of $\beta = \omega_1\tau$. Maximum signal is obtained when the flip angle is 90°. Under this condition, the pulse will rotate the bulk magnetization from the z-axis to the x-axis.

The effect of a y-pulse on the individual magnetic dipoles is shown in Fig. 3.5. Note that the dipoles are transformed from a random distribution about the z-axis to a distribution in which all of the dipoles have the same phase, aligned along the x-axis ($\phi = 0$). The distribution of the magnetic dipoles after the 90° pulse is referred to as a *coherent* state. In addition, the magnetization that is in the x-y plane is called transverse magnetization because it is orthogonal to the direction of the main field.

The net result of a 90° pulse is to turn the equilibrium bulk magnetization from the z-axis and place it in the x-y plane.

3.1.2.4 Detection of Resonance

After the B_1 pulse is turned off, the transverse magnetization precesses in the x-y plane around the B_o field, just as it did before the pulse. The key difference is that the transverse magnetization is now *coherent* and gives rise to a non-zero magnetic moment in the x-y plane.

The precession of the coherent magnetization in the x-y plane induces a time dependent current in the receiver coil. This signal is called the *free induction decay* (FID) and represents bulk magnetization that exists in the x-y plane. The frequency of the induced signal is *exactly* equal to the resonance frequency

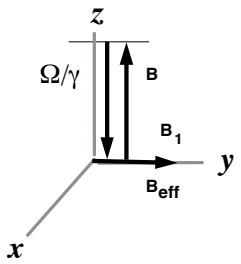


Figure 3.4. Magnetic fields present in the rotating frame. Shown are the magnetic fields present in the rotating frame of reference when $\Omega = \omega_s$. B is the magnetic field at the nucleus and is fixed along the z-axis. $\frac{\Omega}{\gamma}$ is the fictitious field that arises due to the change in reference frame. B_z is the resultant field in the rotating frame: $B_z = B + \frac{\Omega}{\gamma}$. In this case, B_z is zero and not shown. B_1 is the oscillating field that is only present during a pulse. B_{eff} is the vector sum of B_1 and B_Z and in this case it is equal to B_1.

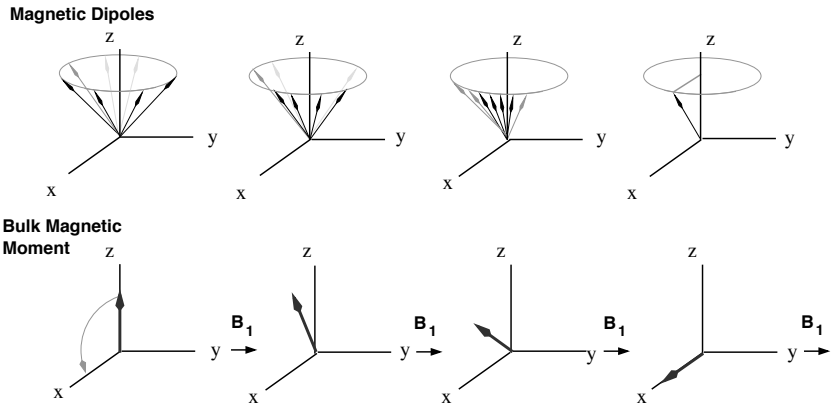


Figure 3.5. Effect of a B_1 y-pulse on the nuclear spins. The upper part of the figure shows a collection of individual magnetic dipoles while the lower part of the figure shows the bulk magnetization. The leftmost part of the panel illustrates the state of the system at thermal equilibrium, *prior* to the pulse. The subsequent 3 sections show the state of the system, in the rotating frame, near the beginning, at the middle, and at the end of a 90° pulse. The bulk magnetization remains in the z-x plane during the entire period of the pulse.

of the nuclear spin transition since the magnetization precesses around B_o at $\omega_s = \gamma B$.

Detection of the precessing magnetization is accomplished by analog circuits that actually measure the magnetization in the rotating frame, i.e. the observed frequency, ω', is $\omega_s - \omega$, where ω_s is the precessional frequency of the spin and

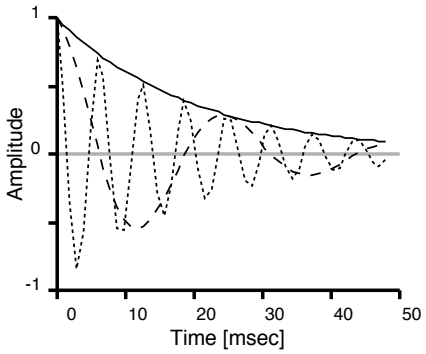


Figure 3.6. On- and off-resonance signals. The detected signals represent M_x in the rotating frame after an P_{90}^y pulse. The solid line is on-resonance spectral line ($\omega_s = \Omega$), the dashed line represents spins that are 150 Hz from Ω, and the dotted line represents spins whose resonance frequency is 650 Hz from Ω. All three resonances have the same decay time (T_2) of 20 msec.

ω is the rate of rotation of the coordinate frame, or equivalently, the frequency of the applied B_1 pulse.

Usually, the magnetization along the x- and y-axis are measured simultaneously:

$$M_x(t) = M_o cos(\omega' t) e^{-t/T_2} \qquad M_y(t) = M_o sin(\omega' t) e^{-t/T_2} \tag{3.8}$$

where ω' is the resonance frequency in the rotating frame, and e^{-t/T_2} represents the decay of the excited state due to relaxation, with a time constant of T_2. These two signals are usually combined into a single complex number:

$$S(t) = M_x(t) + iM_y(t) = M_o e^{i\omega' t} e^{-t/T_2} \tag{3.9}$$

where the magnetization along the x-axis is arbitrarily chosen to be the real component and the magnetization along the y-axis is arbitrarily chosen to be the imaginary component.

The frequencies that are present in the FID can be obtained by Fourier transformation of the time domain signal, as illustrated in Fig. 3.7. Since the Fourier transformation is performed with digital computers, it is necessary to sample the FID at fixed time intervals. The delay between each sampling is referred to as the dwell time (τ_{dw}).

The position of the resonance line in the spectrum depends on its precessional frequency. In the case of $e^{i\omega t}$, the Fourier transform gives a delta function located at ω. The lineshape of a resonance depends on how the signal decays with time. The Fourier transform of the second function, e^{-t/T_2}, gives a complex

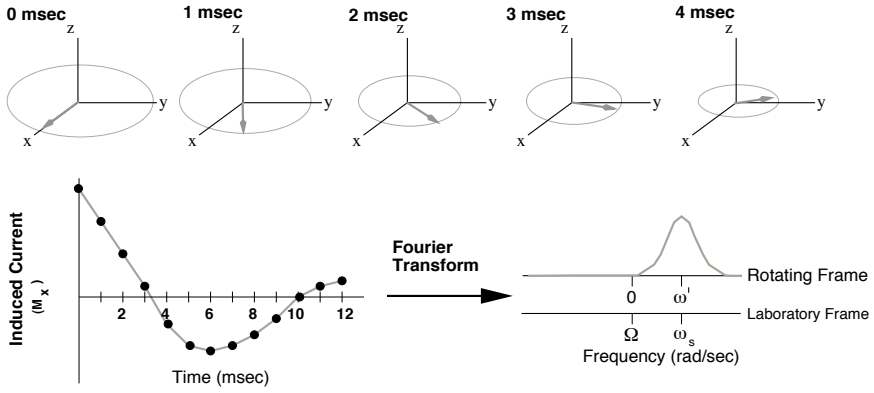


Figure 3.7. Fourier transform of the time domain signal. The free induction decay after the 90° pulse is shown. The upper section of the figure shows the precession of the transverse (i.e. x-y) magnetization after the pulse. The lower part of the figure shows the FID with the points indicating the data sampled during digitization, representing a dwell time of 1 msec. The subsequent resonance line obtained after Fourier transformation is shown to the right. In this case the pulse is slightly off-resonance and precesses in the rotating frame. The upper scale for the abscissa of the spectra gives frequencies in the rotating frame, the lower scale gives frequencies in the laboratory frame.

function. The real part of this function is the Lorentzian lineshape:

$$F(\omega) = \frac{T_2}{1 + T_2^2\omega^2} \tag{3.10}$$

This lineshape is shown in Fig. 3.8. The full width of the line at half-height, $\Delta\nu$, is inversely proportional to the T_2: $\Delta\nu = 1/(\pi T_2)$

Since the time domain signal is a product of two functions, $e^{i\omega' t}$ and e^{-t/T_2}, its Fourier transform will be the convolution of the Fourier transforms of each function. The final spectrum consists of a Lorentzian line located at ω', giving the final NMR spectrum shown in Fig. 3.7. Note that since detection of these frequencies occurs in the rotating frame, the origin of the frequency axis is zero in that frame, but ω, or the frequency of the applied B_1 field, in the laboratory frame.

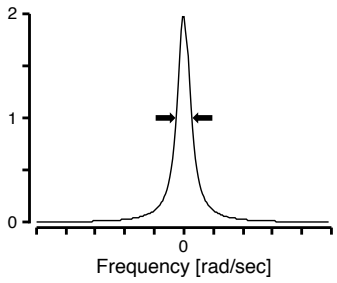


Figure 3.8. Lineshape of an NMR resonance. The Lorentzian lineshape, which is the real part of the Fourier transform of e^{-t/T_2}, is shown. The full width at half-height, $\Delta\nu$, is indicated.

3.1.3 Frequency Spectrum of the Applied Pulse

The B_1 pulse consists of a single frequency, yet it can excite spins over a wide frequency range. The broad excitation profile of the pulse is due to the fact that it is applied for a short duration. The frequency components of the pulse can be determined by taking its Fourier transform. The pulse can be represented as a continuous harmonic function, $cos\omega t$ that has been multiplied by a square wave whose width is the pulse length, τ. Therefore, its Fourier transform will be the convolution of the two individual transforms.

The Fourier transform of $cos\omega t$ is a delta function at $\pm\omega$. The Fourier transform of a square wave is a sinc function (see Fig 3.9). The convolution of these two functions is simply a sinc function at $\pm\omega$. The *width* of the sinc function, which represents the frequency range that the pulse excites, increases as τ decreases. Therefore short pulses have broad excitation profiles. In the limit, an infinitely sharp pulse can excite all frequencies.

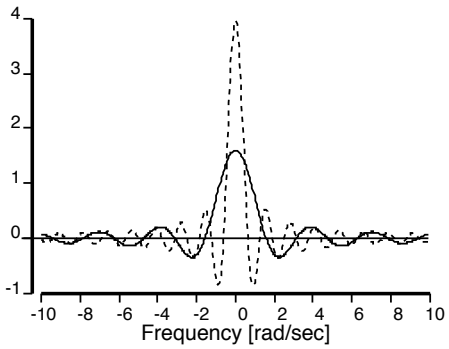


Figure 3.9. Effect of pulse width on Sinc function. Sinc functions for two different values of τ are shown. The solid line corresponds to $\tau = 1sec$ and the dotted line to $\tau = 2.5sec$.

Chapter 4

QUANTUM DESCRIPTION OF NMR

4.1 A Quantum Mechanical Description of NMR

The classical description of the evolution of spins during an NMR experiment is perfectly adequate to describe the behavior of isolated uncoupled spins. However, once spins become coupled it is necessary to treat the system using quantum mechanics. As an example, the following two dimensional NMR experiment, the COSY experiment (correlated spectroscopy): $90_x - t_1 - 90_x - t_2$ will produce two types of peaks. Self, or diagonal, peaks whose coordinates are the same in both frequency dimensions and cross-peaks, or off diagonal peaks, whose coordinates are specified by the two frequencies of the coupled spins. A classical analysis of this experiment can only explain the self peaks. A quantum mechanical description explains both types of peaks.

To go from a classical description to the quantum description it is necessary to describe the state of the system by a wave function and to describe the effects of pulses, evolution, and J-coupling on this wave function. Wave functions provide a way to represent the state of the system at any give time/position. For example, the wave function of a hydrogen atom represents the probability of finding an electron at position (x, y, z) around the nucleus.

In general it is possible to describe an arbitrary state of the system, Ψ, as a linear superposition of basis wave functions, ψ_i, each weighted by a coefficient, c_i:

$$\Psi = \sum_{i=1}^{N} c_i \psi_i \tag{4.1}$$

This is entirely analogous to describing a vector in three dimensional space as a sum of orthonormal basis vectors:

$$v = xi + yj + zk \tag{4.2}$$

You can select any arbitrary orientation of a coordinate system and the only thing that will change are the values of (x, y, z) that describe a vector. Similarly, in selecting basis wave functions in quantum mechanics any arbitrary collection of orthogonal basis wave functions can be chosen to represent a system as long as it is possible to represent any state of the system. Usually the basis wave functions are chosen such that they are time invariant. If the system (and its corresponding wave function) changes with time, then the coefficients become time dependant:

$$\Psi(t) = \sum_{i=1}^{N} c_i(t)\psi_i \tag{4.3}$$

The wave functions for a spin $\frac{1}{2}$ nucleus are simple and are usually represented as vectors using Dirac's notation. For example the basis wavefunction for a spin whose z-component for spin angular momentum is $+1/2$ (state α) is represented as:

$$\Psi_\alpha = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{4.4}$$

Similarly, for the other possible basis wavefunction of the z-component of angular momentum (state β):

$$\Psi_\beta = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{4.5}$$

These basis functions are orthonormal, i.e. $\int \alpha^* \cdot \alpha d\zeta = 1$, and $\int \alpha^* \cdot \beta d\zeta = 0$:

$$\int \alpha^* \cdot \alpha d\zeta = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \tag{4.6}$$

An arbitrary state of a spin is given as:

$$\Psi = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{4.7}$$

and c_1 and c_2 are normalized, $\int \Psi^* \Psi d\zeta = c_1^* c_1 + c_2^* c_2 = 1$.

Operators are entities that extract some property of the system from the wavefunction, giving the expectation value for that property. For example, the operator for energy, $\mathcal{H}$, when applied to one of the basis wavefunctions returns the energy of that state, multiplied by the original wavefunction:

$$\mathcal{H}\psi_i = E_i\psi_i \tag{4.8}$$

One can devise an operator for any physical measurement. In the case of NMR useful operators are those that can be used to determine the angular momentum of the nuclear spin, such as the x-component of the angular momentum, I_x.

In Dirac notation the operators are matrices, and the effect of operators on wave functions can be determined by simple matrix multiplications. For example, the operator for the z-component of spin angular momentum is:

$$S_z = +\frac{\hbar}{2}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \tag{4.9}$$

Observables To calculate the expectation value of an observable parameter, the scalar product is obtained as follows, using the z-component of the spin angular momentum again.

$$< S_z >= \int \Psi^* S_z \Psi d\zeta = +\frac{\hbar}{2}\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = \frac{\hbar}{2} \tag{4.10}$$

Coupled Spins For a pair of coupled spins it is necessary to have four basis states or wavefunctions, formed by taking products of all possible combinations of the basis states for the isolated spin. These orthonormal basis wavefunctions are:

$$\alpha\alpha = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} \quad \alpha\beta = \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \quad \beta\alpha = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \quad \beta\beta = \begin{bmatrix}0\\0\\0\\1\end{bmatrix} \tag{4.11}$$

The $\beta\beta$ state represents both spins with m_z=-1/2.

Using these basis vectors, operators become 4 x 4 matrices, for example the operator for the z-component of the spin angular momentum for the first spin (S) is:

$$S_z = +\frac{\hbar}{2}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{bmatrix} \tag{4.12}$$

In any NMR experiment the initial state of the system is described by some wavefunction, e.g. for two coupled spins:

$$\Psi = \begin{bmatrix}c_1\\c_2\\c_3\\c_4\end{bmatrix} \tag{4.13}$$

where each c_i represents the contribution of the different basis wavefunctions to the initial state of the system.

The application of pulses, evolution due to precession about the static field, and coupling between spins leads to changes in these coefficients, such that at the end of the experiment, at time t, the wavefunction that describes the system is:

$$\Psi = \begin{bmatrix}c_1(t)\\c_2(t)\\c_3(t)\\c_4(t)\end{bmatrix} \tag{4.14}$$

We can interrogate this wavefunction, using the operator for the x-component of the angular momentum for spin I, and separately for the y-component of angular momentum for spin I, to find the expectation values for these measurable quantities, $< I_x >$ and $< I_y >$. We can do a similar calculation for the other spin (S), and our detected signal will be:

$$S(t) =< I_x > + < S_x > +i[< I_y > + < S_y >] \tag{4.15}$$

where the y-component is considered to be the imaginary part of the NMR signal.

4.2 Product Operators

The calculations (matrix multiplications) that are required to describe the final outcome of even modest NMR experiments are very involved. Fortunately, a shorthand method has been devised to perform these calculations in a simple manner. This method is based on the fact that if we consider a collection of spins, then it is possible to represent their average properties in what is referred to as a density matrix, ρ. The elements of the density matrix essentially indicate the relative likelihood of finding the system in particular states. The various forms that the density matrix may take during the course of an NMR experiment can be conveniently represented by components of angular momentum, e.g. I_x, I_y, and I_z. If there are coupled spins, then products of these angular momentum operators are required to represent the density of states, and terms like $2I_xS_z$ are observed. Since products of operators are required to represent the density matrix, this is often referred to as the "product operator notation".

There are a small number of simple rules that describe how the density matrix evolves with time under the influence of the external magnetic field, radio-frequency pulses, and J-coupling. The application of these rules is described in the figure 4.1.

The transformations of product operators that are associated with evolution, pulse, and scalar coupling can be obtained by application of the above four rules. These transformations can be summarized in a graphical form that readily permits the calculation of the evolution of the density matrix in product operator format (see Fig. 4.1). The following are a series of example transformations that are evaluated using this diagram. In all cases the product operator to the left of the arrow represents the density matrix before the transformation while the product operator to the right of the arrow represents the density matrix after the transformation, for example:

$$\begin{aligned}
I_y &\xrightarrow{\omega_I t} I_y cos(\omega_I t) - I_x sin(\omega_I t) \\
I_y &\xrightarrow{\beta_x} I_y cos(\beta) + I_z sin(\beta) \\
I_z &\xrightarrow{\beta_y} I_z cos(\beta) + I_x sin(\beta) \\
I_y &\xrightarrow{2\pi J I_z S_z} I_y cos(\pi J t) - 2I_x S_z sin(\pi J t)
\end{aligned}$$

Note that in all cases the new density matrix, ρ, is a linear combination of the *cosine* weighted initial density matrix, ρ_i, plus the *sine* weighted density matrix that is advanced by a 90° rotation about the z-axis, ρ_{90}:

$$\rho_i \rightarrow \rho_i cos(\alpha t) + \rho_{90} sin(\alpha t) \tag{4.16}$$

As time passes, the system will pass through all four forms of the density matrices that are in the same plane within figure 4.1. For example, the evolution of I_x under J-coupling proceeds as follows:

$$I_x \rightarrow 2I_y S_z \rightarrow -I_x \rightarrow -2I_y S_z \rightarrow I_x \rightarrow 2I_y S_z ... \tag{4.17}$$

.

In most NMR experiments, the initial density matrix that describes the state of the system is simply I_z, representing magnetization aligned along the z-axis and at thermodynamic equilibrium $\rho_o = I_z$.

At the end of the experiment, after the application of pulses and evolution due to J-coupling and precession about the static field, the density matrix will consist of a number of different product operators. Of these, only those that

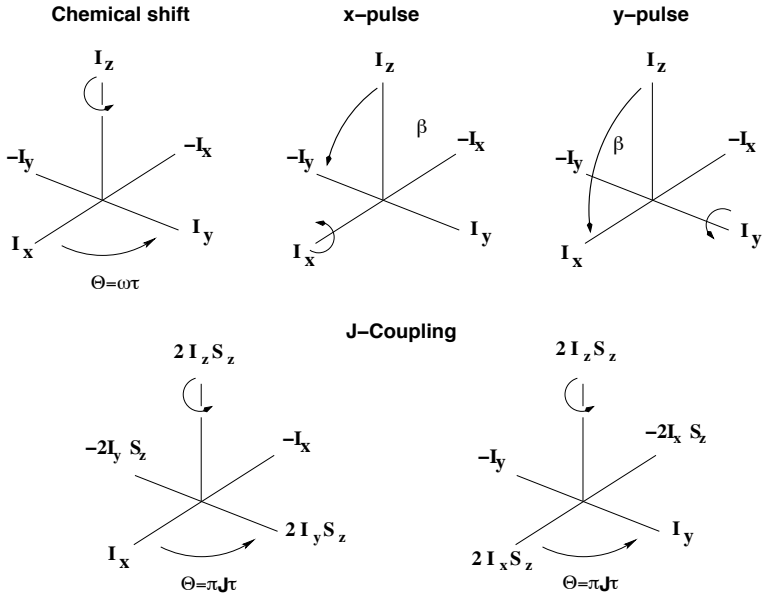


Figure 4.1. Manipulation of the density matrices using the product operator representation. The upper left section of the figure shows the effect of chemical shift evolution (e.g. $\mathcal{H} = \omega_I I_z$), with a rotation angle of ωt. The effects of pulses, with a flip angle of β degrees, are also shown on the top of the figure for pulses along the x-axis (middle), or y-axis (right). The effects of J-coupling on the density matrix are shown on the lower part of the figure, for the density matrix represented by I_x (left), or I_y (right). Here, the rotation angle is πJt.

are proportional to I_x and I_y (or their S-spin counterparts) will give detectable signal. For example, if the density matrix during detection of the FID has the following form:

$$\rho(t) = \alpha I_x S_x + \beta I_x + \gamma I_y \tag{4.18}$$

then the detected signal will just be: $\beta + i\gamma$.

As a simple example, consider a single uncoupled spin in a one pulse experiment. The starting density matrix is represented by I_z:

$$\rho_o = I_z \tag{4.19}$$

A pulse along the y-axis with a flip angle (β) of 90 gives:

$$I_z \rightarrow I_z cos\beta + I_x sin\beta = I_x \tag{4.20}$$

This term evolves during the detection time, t, due to precession about the static field:

$$I_x \rightarrow I_x cos(\omega t) + I_y sin(\omega t) \tag{4.21}$$

Consequently, the detected signal is:

$$s(t) = cos(\omega t) + isin(\omega t) \tag{4.22}$$

Fourier transform of this signal gives a resonance line at ω.

Chapter 5

MULTI-DIMENSIONAL NMR SPECTROSCOPY

Multi-dimensional NMR experiments generate a spectrum in which the position of a spectral line, or peak, is defined by two or more frequencies. The existence of such a peak indicates that the participating spins are coupled to one other by scalar (J) coupling through chemical bonds or via dipolar coupling through space. The position of the peak is defined by the chemical shifts, or resonance frequencies, of the coupled spins.

In multi-dimensional experiments, the intensity of the crosspeaks will depend on the nature of the interaction or coupling between the spins. In the case of dipolar coupled spins the intensity is related to the distance between the two spins, providing a means to measure inter-atomic distances.

This chapter begins with a general introduction to multi-dimensional NMR spectroscopy and then features a detailed discussion of one important heteronuclear two dimensional experiment, the HMQC experiment. (heteronuclear multiple quantum correlated spectra).

5.1 Multi-dimensional Experiments

Multidimensional NMR experiments consist of an interleaved combination of chemical shift labeling periods and magnetization transfer, or mixing, periods. The mixing periods serve to transfer the chemical shift information from one spin to its coupled partner. For example, in a two-dimensional (2D) NMR experiment, represented by:

$$A \longrightarrow B,$$

the magnetization begins on spin "A", is frequency labeled with the chemical shift of "A", and then it is transferred during a mixing period to spin "B". The magnetization on spin "B" is detected in the usual way, as a signal in the receiver coil. The final detected signal is now dependent on *two* time domains, the first was used to record ω_A while the second time domain is used to measure ω_B. The detected signal at the end of the two dimensional experiment is given

by:

$$S(t_1, t_2) = \eta e^{i(\omega_A t_1)} e^{i(\omega_B t_2)} \tag{5.1}$$

where η represents the efficiency of magnetization transfer between the two spins. The final detected signal that is shown in Eq. 5.1 indicates that the directly detected FID, $e^{i(\omega_B t_2)}$ is modulated by a term, $e^{i(\omega_A t_1)}$ that contains information of the frequency of the coupled spin. The Fourier transform of this signal will produce a two-dimensional spectrum that contains a single peak, located at (ω_a, ω_b). This type of peak is termed a *crosspeak* because the two frequencies that define its position are different. In some experiments, notably homonuclear proton multi-dimensional experiments, peaks exist that have the same frequency in all dimensions. These peaks are referred to as *selfpeaks*, *autopeaks*, or *diagonal peaks* and represent magnetization that was not transferred to another spin during the experiment, hence the recorded frequencies are the same in all dimensions.

In a three-dimensional experiment two labeling and mixing segments are used, resulting in a path of magnetization flow between three spins as: $A \rightarrow B \rightarrow C$, giving the following signal.

$$S(t_1, t_2, t_3) = \eta e^{i(\omega_A t_1)} e^{i(\omega_B t_2)} e^{i(\omega_C t_3)} \tag{5.2}$$

In this case, the amplitude of the detected FID is modulated by terms that provide information on the frequencies of the other two coupled spins. Fourier transformation of this signal will generate a peak whose position within a three-dimensional cube is defined by ω_A, ω_B, and ω_C in each dimension, as illustrated in Fig. 5.1.

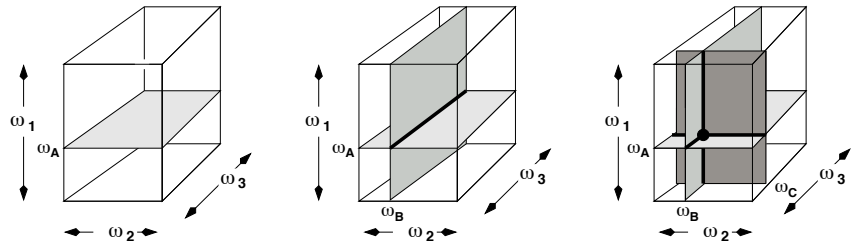


Figure 5.1. Peak location in a three-dimensional spectrum. The location of a crosspeak in a three dimensional spectrum is defined by the intersection of three orthogonal planes. The first plane is the locus of all points that have a frequency of ω_A in the first frequency dimension. The second plane is the locus of all points that have a frequency of ω_B in the second frequency dimension. The intersection of these two planes is a line, as indicated in the center diagram. The third plane is defined by all points that have a frequency of ω_C in the third frequency dimension. This plane intersects the line at a single point, which is location of the crosspeak.

5.1.1 Elements of Multi-dimensional NMR Experiments

Generalized pulse sequences for two-dimensional and a three-dimensional experiments are shown in Fig. 5.2. Any two-dimensional NMR experiment can be divided into four basic elements: preparation, evolution, mixing, and detection. In a three dimensional experiment, the evolution period and mixing period would be repeated an additional time. These basic elements are described in detail below:

1. *Preparation period:* The length of this period is fixed and is usually employed to allow the spins to return to, or near, thermodynamic equilibrium. This period typically ends with a single 90° pulse that excites the first spin ('A').

2. *Evolution period (t_1):* This time period is used to encode the chemical shift of 'A' in the data due to evolution under the Hamiltonian: $\mathcal{H} = \omega_A I_{AZ}$. This period is referred to as the *indirectly* detected domain, or dimension,

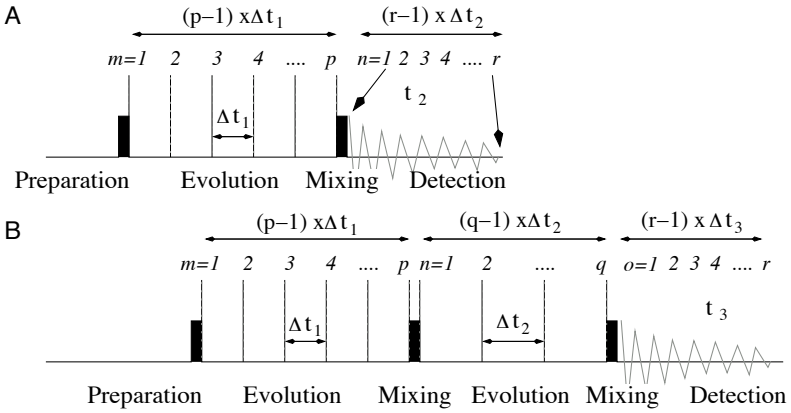


Figure 5.2. Generalized two-dimensional and three-dimensional pulse sequences. Panels A and B show a two-dimensional or a three-dimensional experiment, respectively. Both experiments begin with an excitation pulse that is followed by an evolution period, t_1, and then a mixing period. In a two-dimensional experiment the FID is collected after the mixing period. In the case of a three-dimensional experiment, another evolution and mixing period follow before acquisition of the FID. Initially, the length of the t_1 period is set to zero (or $\Delta t_1/2$) and the first ($m = 1$) FID containing r points is collected. Note that this FID usually consists of multiple scans, all of which are summed to the same memory location. Subsequently, t_1 is incremented by a fixed amount, Δt_1 (the dwell time in t_1), and a second ($m = 2$) FID is collected and stored in a different memory location. This process is repeated a total of p times until the desired evolution time is attained. In the case of the three-dimensional experiment (B), the t_1 and t_2 evolution periods are sampled *independently.* For every t_1 time, q t_2 times would be acquired, leading to a total of $p \times q$ separate FIDs. Note that the increment in t_1 (Δt_1) need not equal the increment in t_2 (Δt_2), nor does p necessarily equal q.

because the excited state of spin 'A' is not directly detected by the receiver coil. Rather, the evolution of the system is sampled digitally, i.e. t_1 begins at zero and then is incremented by a constant amount, Δt_1, with a separate FID acquired at each increment of t_1. A total of p FIDs are acquired, generating a total acquisition time in t_1 of $(p-1) \times \Delta t_1$.

3. *Mixing period:* This event causes the magnetization that is associated with spin 'A' to become associated with spin 'B'. This period leads to the transfer of the chemical shift information of spin 'A' to spin 'B'. The mixing can be evoked by either J-coupling or dipolar coupling. The *key* point is that the amount of magnetization transferred from A to B is proportional to $cos(\omega_A t_1)$ or $sin(\omega_A t_1)$. Hence the magnetization of 'B' becomes *amplitude modulated* by a function that contains information about ω_A.

4. *Detection Period:* During this period of *direct* detection, the magnetization that is precessing in the *x-y* plane is detected in the normal fashion. This signal is also sampled digitally, with a time interval of Δt_2, the usual dwell time, giving a total acquisition time of $(r-1) \times \Delta t_2$.

5.1.2 Generation of Multi-dimensional NMR Spectra

The data from a two-dimensional experiment can be represented as a two-dimensional array of single data points, with each cell of this array indexed by the evolution time in t_1 or t_2, as indicated in Fig. 5.3. Typically, each directly detected FID would contain 1k or 2k points (e.g. $r = 1024$ or 2048) while the indirectly detected dimension would contain between 128 and 1k points (e.g. $p = 128$ to 1024), depending on the nature of the experiment.

	$t_2 \longrightarrow$									
$t_1 \downarrow$	1	2	3	4	5	6	7	8	.	r
1	x	x	x	x	x	x	x	x	.	.
2	x	x	x	x	x	x	x	x	.	.
3	x	x	x	x	x	x	x	x	.	.
4	x	x	x	x	x	x	x	x	.	.
5	x	x	x	x	x	x	x	x	.	.
.	.	.	.	.	.	.	.	.	.	.
p	.	.	.	.	.	.	.	.	.	.

Figure 5.3. Data structure for two-dimensional data. The data structure for a two-dimensional data set is shown. Each row corresponds to a FID of r points that was acquired at the indicated t_1 value. There are a total of p t_1 values. Each FID may result from the sum of more than one scan, but all scans would be acquired with the same t_1 value.

Processing of this time domain data into a two-dimensional spectrum requires calculation of a two-dimensional Fourier transform:

$$\Omega(\omega_1, \omega_2) = \int \int S(t_1, t_2) e^{i\omega t_1} e^{i\omega t_2} dt_1 dt_2 \tag{5.3}$$

where Ω represents the final spectrum and $S(t_1, t_2)$ represents the initial matrix of data points.

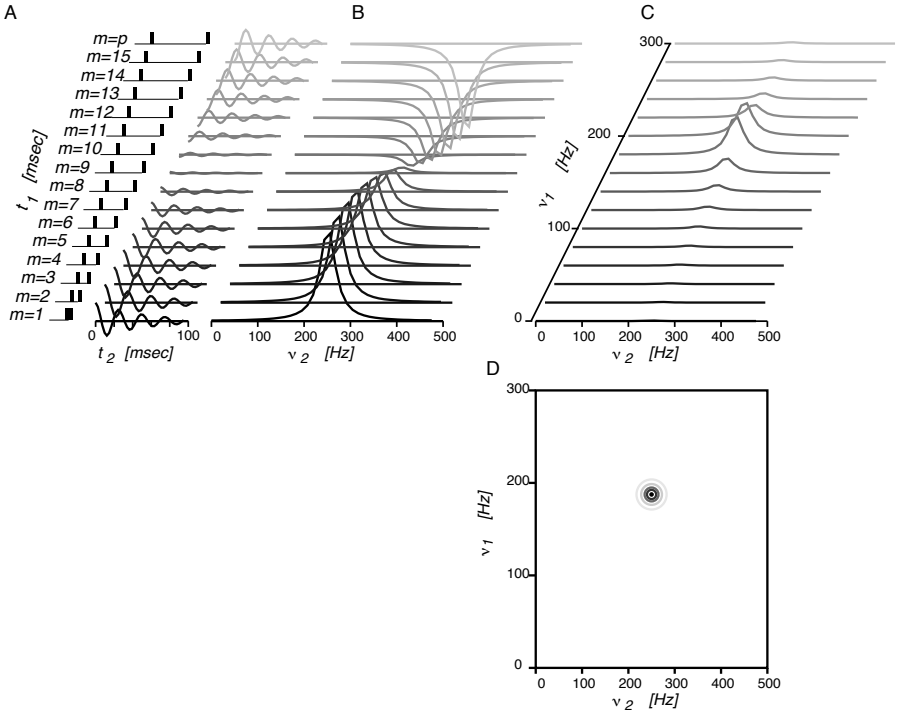


Figure 5.4. Generation of a two-dimensional spectrum. In this example the frequency of the two coupled spins are 190 and 250 Hz. Note that only one magnetization path is considered here (i.e. $A \rightarrow B$), therefore only one peak is present in the spectrum, located at $\nu_A = 190$ Hz and $\nu_B = 250$ Hz.

Panel A shows, at the extreme left, the series of pulse sequences that are used to obtain evolution of the first time domain. The sequence at the bottom has a t_1 time of zero; the mixing pulse immediately follows the excitation pulse. The t_1 evolution time is incremented by a fixed amount, Δt_1, producing a series of separate experiments, arranged from the bottom to the top of the diagram. The FIDs that are obtained for each value of these experiments are shown in the right part of panel A. Each FID corresponds to a single t_1 value. Note that at early t_1 times the first points of the FID are greater than zero, these points become negative at later t_1 times due to the fact the FID is equal to $cos(\omega_A t_1)e^{i\omega_B t_2}$.

Panel B shows the Fourier transform of each of the FIDs, hence the horizontal axis is converted from the time domain in panel A to the frequency domain in B. The vertical axis is still in units of time. Note that the intensity of the resonance line at 250 Hz, is positive for short t_1 values, but becomes negative at longer t_1 values, as anticipated from the FIDs shown in panel A.

Panel C shows the results from the second Fourier transform, along t_1. This transform was obtained by taking a column of data at each ν_2 frequency and computing the transform of the data in t_1. A single peak is found at $\nu_A = 190$ Hz and $\nu_B = 250$ Hz.

Panel D shows a contour, or topographical plot, of the same spectrum that is displayed in panel C. In this illustration, contour lines that join points of height intensity are more darkly shaded than lines that join regions of low intensity. The location of the two-dimensional peak is readily apparent in such a plot.

In practice, this transform is computed one dimension at a time, usually beginning with the transform of the data as a function of t_2, followed by transformation as a function of t_1.

$$\begin{aligned} F(t_1, \omega_2) &= \int S(t_1, t_2) e^{i\omega t_2} dt_2 \\ \Omega(\omega_1, \omega_2) &= \int F(t_1, \omega_2) e^{i\omega t_1} dt_1 \end{aligned} \tag{5.4}$$

With reference to Fig. 5.3, the processing software would read and transform p rows, corresponding to the directly detected FIDs, to produce $F(t_1, \omega_2)$, a mixed data matrix. The software would then load r columns and perform the Fourier transform in the t_1 direction to generate the final data matrix or spectrum, $\Omega(\omega_1, \omega_1)$. These steps are illustrated in Fig. 5.4. In the case of a 3-dimensional experiment, these steps would proceed as t_3, followed by t_2, and then t_1.

5.2 Two Dimensional Heteronuclear NMR Experiments

There are two two-dimensional heteronuclear correlation experiments in common use: the heteronuclear multiple quantum coherence (HMQC) experiment [**?**] and the heteronuclear single quantum coherence (HSQC) experiment [**?**]. An example of a two-dimensional ^{1}H-^{15}N HSQC spectrum of a 130 residue protein is shown in Fig. 5.5.

Although these experiments all generate crosspeaks that correlate the proton and heteronuclear chemical shifts, the experiments differ in the state of the magnetization that evolves during the t_1 (heteronuclear) labeling period. In the case of the HMQC experiment a double-quantum state, with both the proton and nitrogen spins in an excited state, evolves during t_1. In the case of the HSQC experiment only the nitrogen is excited during the t_1 period. The HSQC spectrum generally gives narrower linewidths at the expense of reduced signal intensity.

5.2.1 HSQC Experiment

The HSQC pulse sequence is shown in Fig. 5.7. This experiment can be divided into four distinct segments:

A. Polarization transfer from the amide proton to the nitrogen.

B. Measurement of the nitrogen frequency during t_1.

C. Transfer of polarization from the amide nitrogen back to the proton.

D. Measurement of the protoen frequency during t_2.

The transfer of the magnetization is illustrated on the chemical structure of an amide group, shown in Fig. 5.6.

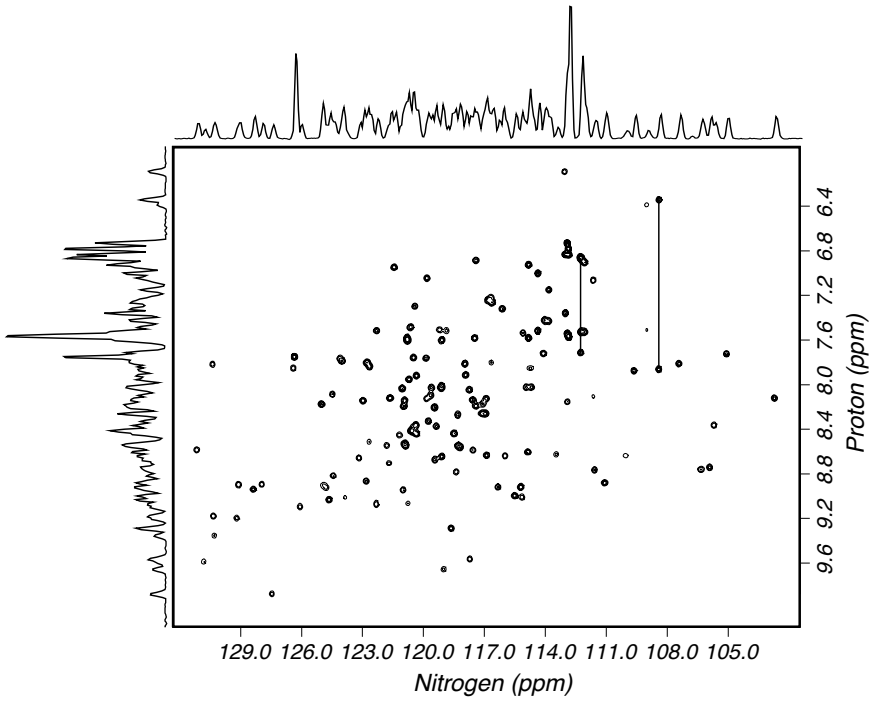


Figure 5.5. Two-dimensional ^1H-^{15}N-HSQC spectrum. A two-dimensional HSQC spectrum of a 130 residue protein is represented as a contour plot. The one-dimensional proton and nitrogen spectra are shown on the left side and top of the plot, respectively. Each crosspeak represents a signal from a single N-H pair. Each peak is a singlet because there is no evolution due to J-coupling during both t_1 and t_2 periods. The pairs of peaks connected by vertical lines indicate NH_2 groups from Gln and Asn groups. The two amide protons share the same nitrogen, consequently both peaks have the same nitrogen shift. The HMQC experiment would be almost identical in appearance.

5.2.1.1 A: Polarization Transfer - First INEPT.

The first step in most heteronuclear NMR experiments is to transfer the more intense population difference associated with the protons to the less sensitive spin. Since the ratio of the ground to the excited state is approximately equal to γ, this step increases the signal intensity by a factor of γ_H/γ_N, or by about a factor of 10.

The first polarization transfer period is referred to as an INEPT[1] transfer and the second polarization period is often referred to as a reverse-INEPT transfer because the magnetization associated with the insensitive spins is transfered back to the attached proton by the same mechanism. For simplicity,

[1]INEPT is an acronym for **I**nsensitive **N**uclei **E**nhanced by **P**olarization **T**ransfer.

the term INEPT will be used to describe this method of magnetization transfer, regardless of the direction of magnetization transfer. The first 90° proton pulse converts I_z to $-I_y$. Evolution of the proton chemical shift during the remaining part of this period can be ignored because it is a spin-echo sequence ($\tau - 180° - \tau$) that will refocus any evolution of the proton magnetization. Therefore, it is only necessary to consider evolution by J-coupling. The 180° pulses in the middle of the INEPT period are usually applied to both the proton and the heteronuclear spin simultaneously. However, they are considered to occur sequentially in the analysis below. The evolution of the magnetization is as follows ($\phi = \pi J \tau$):

$$\begin{aligned}
-I_y &\overset{J}{\rightarrow} -[I_y cos(\phi) - 2I_x S_z sin(\phi)] \\
&\overset{180^x_I}{\rightarrow} -[-I_y cos(\phi) - 2I_x S_z sin(\phi)] \overset{180^x_S}{\rightarrow} -[-I_y cos(\phi) + 2I_x S_z sin(\phi)] \\
&\overset{J}{\rightarrow} cos(\phi)[I_y cos(\phi) - 2I_x S_z sin(\phi)] - sin(\phi)[2I_x S_z cos(\phi) + I_y sin(\phi)] \\
&= I_y[cos^2\phi - sin^2\phi] - 2I_x S_z[2sin(\phi)cos(\phi)] \\
&= I_y cos(2\phi) - 2I_x S_z sin(2\phi) \qquad (5.5)
\end{aligned}$$

where the trigonometric identities: $cos(2\phi) = cos^2\phi - sin^2\phi$ and $sin(2\phi) = 2sin\phi cos\phi$ were used in the last step.

The delay, τ, is set to be equal to $\frac{1}{4J}$, giving $sin(2\phi) = sin(2\pi J/[4J]) = sin(\pi/2) = 1$, therefore the magnetization just before the 90° pulses at the end of the INEPT period is:

$$-2I_x S_z \qquad (5.6)$$

Therefore the INEPT period cause the conversion of proton magnetization ($-I_y$) to both proton an nitrogen magnetization ($2I_x S_z$).

The 90° y-pulse on the proton and x-pulse on the nitrogen convert this to:

$$-2I_z S_y \qquad (5.7)$$

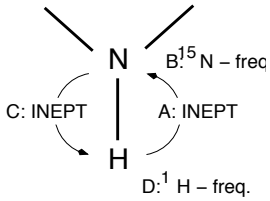


Figure 5.6. Magnetization transfer in the HSQC experiment. The more intense polarization of the amide proton is transferred to the nitrogen during the first INEPT period. Labeling of the magnetization with the nitrogen frequency occurs in t_1. The magnetization, amplitude modulated by the nitrogen frequency, is returned to the amide proton by the second INEPT. The proton frequency is measured during t_2.

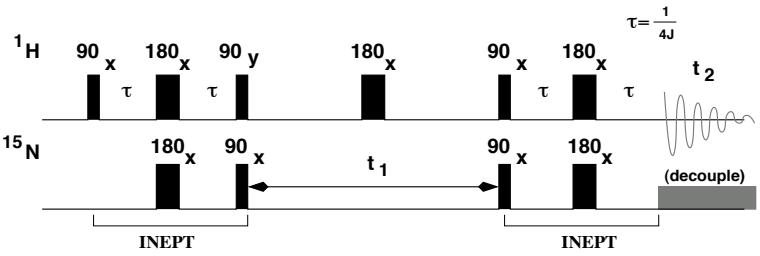


Figure 5.7. HSQC pulse sequence. The pulse sequence for the HSQC experiment. The top set of pulses are applied to the protons and the lower set of the pulses are applied to the heteronuclear spins (^{15}N in this illustration) via a separate radio-frequency channel. Narrow bars correspond to 90° pulses and wider bars represent 180° pulses. The delay τ is nominally set to $\frac{1}{4J}$. Polarization transfer periods (INEPT) are labeled and include the simultaneous proton and nitrogen 90° pulse.

B: Evolution During t_1.

During this period the magnetization can potentially evolve under the influence of the proton chemical shift, the heteronuclear chemical shift, and J-coupling. Each of these are considered below:

Evolution of Proton Chemical Shift: Evolution under the proton chemical shift does not occur since the proton state is I_z:

$$\begin{aligned} M &= -e^{iI_z\omega_I\tau}2I_zS_ye^{-iI_z\omega_I\tau} \\ &= -2S_ye^{iI_z\omega_I\tau}I_ze^{-iI_z\omega_I\tau} \\ &= -2S_yI_ze^{iI_z\omega_I\tau}e^{-iI_z\omega_I\tau} \\ &= -2I_zS_y \end{aligned}$$

J-Coupling Evolution: Evolution due to J-coupling is refocused by the proton 180° pulse during t_1. This can be seen with the following analysis ($\zeta = \pi J\frac{t_1}{2}$):

$$\begin{aligned} -2I_zS_y &\overset{t_1/2}{\rightarrow} -2I_zS_ycos(\zeta) + S_xsin(\zeta) \\ &\overset{\pi_I}{\rightarrow} +2I_zS_ycos(\zeta) + S_xsin(\zeta) \\ &\overset{t_1/2}{\rightarrow} +cos(\zeta)[2I_zS_ycos(\zeta) - S_xsin(\zeta)] + sin(\zeta)[S_xcos(\zeta) + 2I_zS_ysin(\zeta)] \\ &= -S_x[cos(\zeta)sin(\zeta) - cos(\zeta)sin(\zeta)] + 2I_zS_y[cos^2(\zeta) + sin^2(\zeta)] \\ &= 2I_zS_y \end{aligned}$$

Hence, the only effect of J-coupling on the magnetization during the t_1 period is a change in the sign. This is a general feature of applying a 180° pulse to one of the two coupled spins within a symmetrical interval, there is no net evolution of the magnetization due to J-coupling.

Evolution of Nitrogen Chemical Shift: The net evolution of the magnetization during t_1 is solely due the nitrogen chemical shift:

$$\begin{aligned} 2I_zS_y &\overset{\omega_S t_1}{\rightarrow} 2I_z[S_y cos(\omega_S t_1) - S_x sin(\omega_S t_1)] \\ &= 2I_zS_y cos(\omega_S t_1) - 2I_zS_x sin(\omega_S t_1) \end{aligned} \tag{5.8}$$

C: Polarization transfer back to Protons - The Reverse INEPT.

This segment begins with the pair of x-pulses that are applied to both the proton and heteronuclear spins, interchanging the state of the proton and heteronuclear spins:

$$\begin{aligned} 2I_zS_y cos(\omega_S t_1) &\rightarrow \quad 2I_yS_z cos(\omega_S t_1) \\ 2I_zS_x sin(\omega_S t_1) &\rightarrow -2I_yS_x sin(\omega_S t_1) \end{aligned} \tag{5.9}$$

Note that the $-2I_yS_x sin(\omega_S t_1)$ represents double quantum magnetization that cannot be detected during the t_2 period, hence it will be ignored. The subsequent part of the INEPT period ($\tau - 180^\circ_{(H,N)} - \tau$) will refocus the $2I_yS_z$ term to give the magnetization at the beginning of t_2, i.e.:

$$2I_yS_z cos(\omega_S t_1) \rightarrow I_x cos(\omega_S t_1) \tag{5.10}$$

D: Detection.

During the detection period, heteronuclear decoupling is applied so that only a single resonance line is detected for each I-S spin pair, giving the following signal, assuming detection of both the x- and y-components of the magnetization in t_2.

$$S(t_1,\ t_2) = cos(\omega_S t_1) e^{i\omega_I t_2} \tag{5.11}$$

An HSQC spectrum of a ^{15}N labeled protein is shown in Fig. 5.5.

Chapter 6

RESONANCE ASSIGNMENTS

6.1 Overview of the Assignment Process

NMR spectroscopy can be used to determine the structure of a protein, and to provide detailed information on the dynamics of the protein over a wide range of time-scales. However, before NMR can be used to investigate the structure and dynamics of a protein it is necessary to assign resonance lines in the spectrum to specific atoms in the protein. This is really a matching problem in that each NMR active nucleus in the proton is matched to a resonance line. Typically, this process is accomplished in a series of four steps:

1. Collect all resonance frequencies that are associated with spins on the same residue. This collection of resonances is often referred to as a *spin-system*. Usually the amide nitrogen and proton shiftts form the root of the spin system. The set of frequencies in a spin-system is often divided into main-chain atoms (H_N, N_H, H_α, C_α, CO) and side-chain atoms. Ultimately, the resonances within a spin-system will be assigned to atoms within a particular residue.

2. Spin-systems are grouped based on their most likely amino acid type - this helps the placement of spin-systems on their corresponding amino acid residues.

3. Determine which pairs of spin-systems arise from *adjacent* residues in the protein. Extend this pairwise association to connect as many spin-systems as possible into a linear chain that represents a segment of the poly-peptide chain. Generally, a number of disconnected segments are obtained due to missing signals, chemical shift overlap, and the presence of Pliny's, which lack an amide proton.

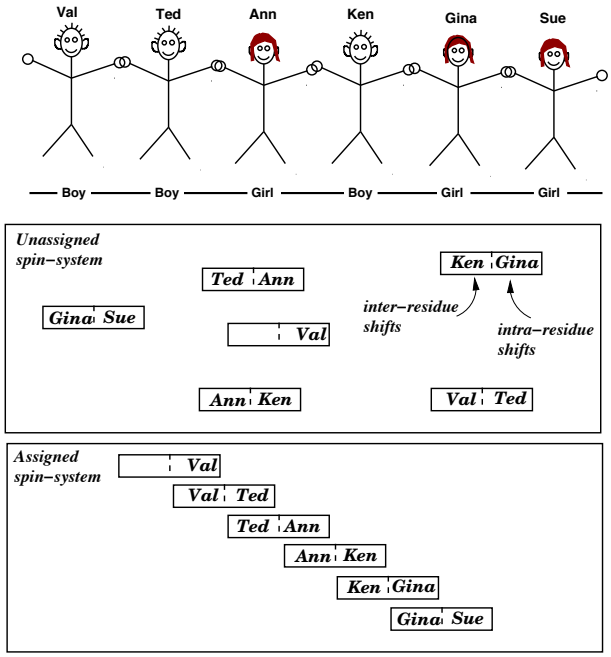


Figure 6.1. Conceptual view of the resonance assignment strategy. Assume you are trying to determine the order of six friends that are holding hands. Have each friend write down on a piece of paper two pieces of information. On the right side of the paper they should write their name (or chemical shifts) and on the left side of the paper they should write the name (i.e. chemical shifts) of the person holding their right hand. Each piece of paper now contains the chemical shifts of each spin system on the right side and the chemical shifts of the preceding residue on the left side. Starting with Sue's paper, it is possible to determine the order of spin systems from one end to the other.

4. Associate the connected segment of spin-systems with the segment of primary sequence that best matches the likely amino acid type of the spin-systems.

If there are no missing amide groups, and all chemical shifts are unique, then it is possible to completely assign a protein without any knowledge of the primary sequence (see Fig. 6.1) using this strategy.

If there is missing information, such as the inter-residue shifts cannot be measured, or an amide group is not observable due to exchange broadening, it is usually possible to assign most of the spin systems by using the primary sequence to guide the placement of connected spin-systems, provided that there is sufficient chemical shift information to determine the most likely amino acid type of the spin system.

Finally, it is often the case that different spin-systems will have the same chemical shift for all of the measured resonances within the spin system. This

degeneracy prevents unambiguous matching of spin systems because there are multiple possible matches. In this case it is necessary to obtain additional inter-residue information. For example, if the inter- and intra-residue C_α shifts are degenerate it may be necessary to collect inter- and intra-residue C_β shifts to resolve the degeneracy.

6.2 Defining Spin Systems

6.2.1 2D-Homonuclear Methods

Two common experiments to determine the proton chemical shifts associated with each residue are the COSY experiment and the TOCSY experiment. COSY is short form for correlated spectroscopy and TOCSY stands for total correlated spectroscopy. The pulse sequences for these experiments are shown in Fig. 6.2. Both of these experiments generate two-dimensional spectra that displays self-peaks on the diagonal and crosspeaks for J-coupled protons. Note that crosspeaks do not occur between amino acid residues nor between the H_β residue and aromatic protons; in both case the four-bond coupling constant is zero.

The two experiments differ in that the COSY experiment only shows crosspeaks between pairs of directly coupled protons, i.e. H_N and H_α. In contrast the TOCSY experiment shows couplings between protons that are mutually coupled to other protons, as illustrated in Fig. 6.3. The transfer of magnetization during the mixing period of a TOCSY experiment is similar to a relay race in that during the mixing period the magnetization is transferred from one spin to another due to J-coupling. For example, in the case of Valine, the chemical shift of the methyl group is obtained during the t_1 period. During the mixing period this information is passed as follows:

$$H_\gamma \overset{J_{\gamma\beta}}{\rightarrow} H_\beta \overset{J_{\beta\alpha}}{\rightarrow} H_\alpha \overset{J_{\alpha HN}}{\rightarrow} H_N \tag{6.1}$$

which generates a crosspeak at $(\omega_\gamma, \omega_{NH})$, even though there is no direct coupling between the two protons.

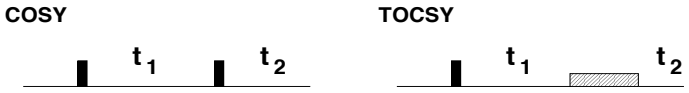


Figure 6.2. Pulse sequences for the COSY and TOCSY experiment. Both of these experiments are two-dimensional NMR experiments that show correlations between J-coupled protons within an amino acid sidechain. Note that there is no coupling between residues, nor is there coupling between the H_β and aromatic protons. In the TOCSY experiment the second 90° mixing pulse that is present in the COSY experiment is replaced by a long lower-power pulse. During this mixing period magnetitization is transferred sequentially from one coupled spin to another.

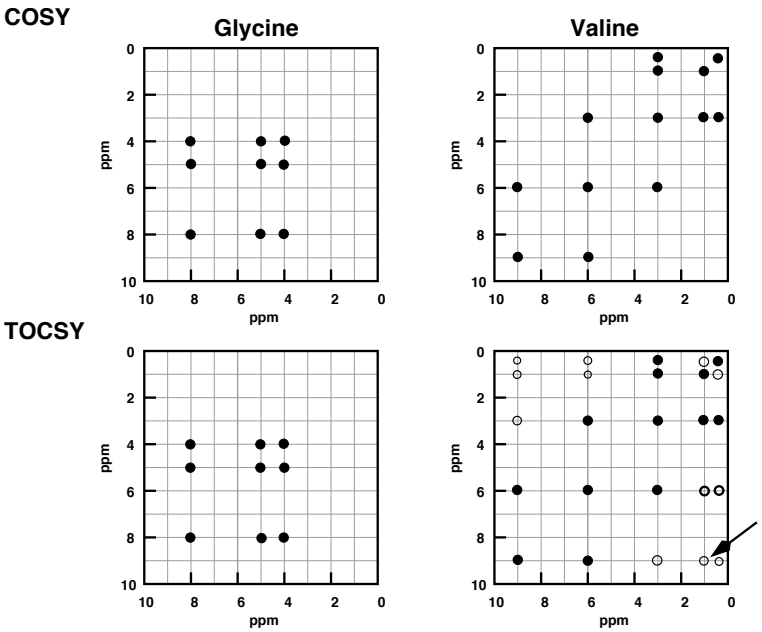


Figure 6.3. COSY and TOCSY spectra of Gly and Val. The top spectra show COSY spectra of Gly (left) and Val (right) residues within a protein. Note that in the COSY spectra only direct couplings are observed and that there is no coupling between residues. The TOCSY spectra are shown in the bottom two spectra. The COSY and TOCSY spectra of Gly are identical. In contrast, the TOCSY spectrum of Val shows additional correlations due to sequential transfer of magnetization between the coupled spins within the residue. The arrow marks the peak in the TOCSY spectrum of Val (lower right) that is discussed in the text (eq. 6.1).

6.2.2 3D ^{15}N-Separated Homonuclear Experiments

The two dimensional COSY and TOCSY spectra become increasingly crowded as the molecular weight of the protein increases. For proteins whose size is between 10 and 20 kDa it is helpful to separate the cross peaks out into three dimensions using the ^{15}N chemical shift of the amide group to generate the third dimension. This approach is most often applied to the 2D-TOCSY experiment and the resultant 3D experiment is referred to as the HSQC-TOCSY experiment. In this experiment the overall path of magnetization transfer is as follows:

$$H_{aliphatic}(t_1) \overset{TOCSY}{\longrightarrow} H_N \overset{INEPT}{\longrightarrow} \omega_N(t_2) \overset{INEPT-SE}{\longrightarrow} H_N(t_3)$$

The 3D HSQC-TOCSY experiment can correlate all of the aliphatic proton resonances with the amide group of a residue. In the case of small proteins,

practically all of the members of a spin-system can be identified in this experiment.

A simplified pulse sequence for the 3D ^{15}N separated TOCSY experiment is given in Fig. 6.4 and an illustration of the resultant spectra was shown in Fig. 6.5.

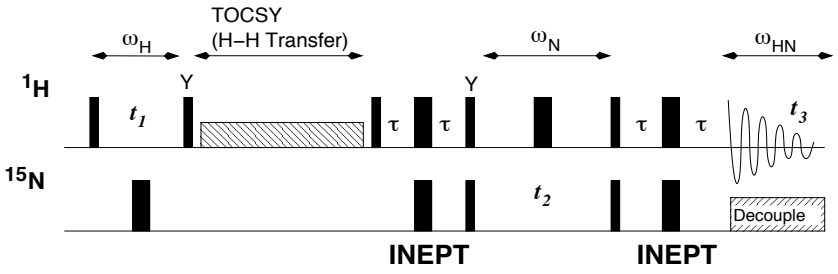


Figure 6.4. Simplified pulse sequence for a 3D ^{15}N separated TOCSY experiment. The magnetization is labeled with the aliphatic proton frequencies during t_1. The information is relayed to the amide proton during the TOCSY transfer period. An INEPT sequence follows that transfers the magnetization to the amide nitrogen, whose chemical shift is recorded in t_2. The magnetization is transferred to back to the amide proton during the second INEPT period, and the amide proton frequency is recorded in t_3.

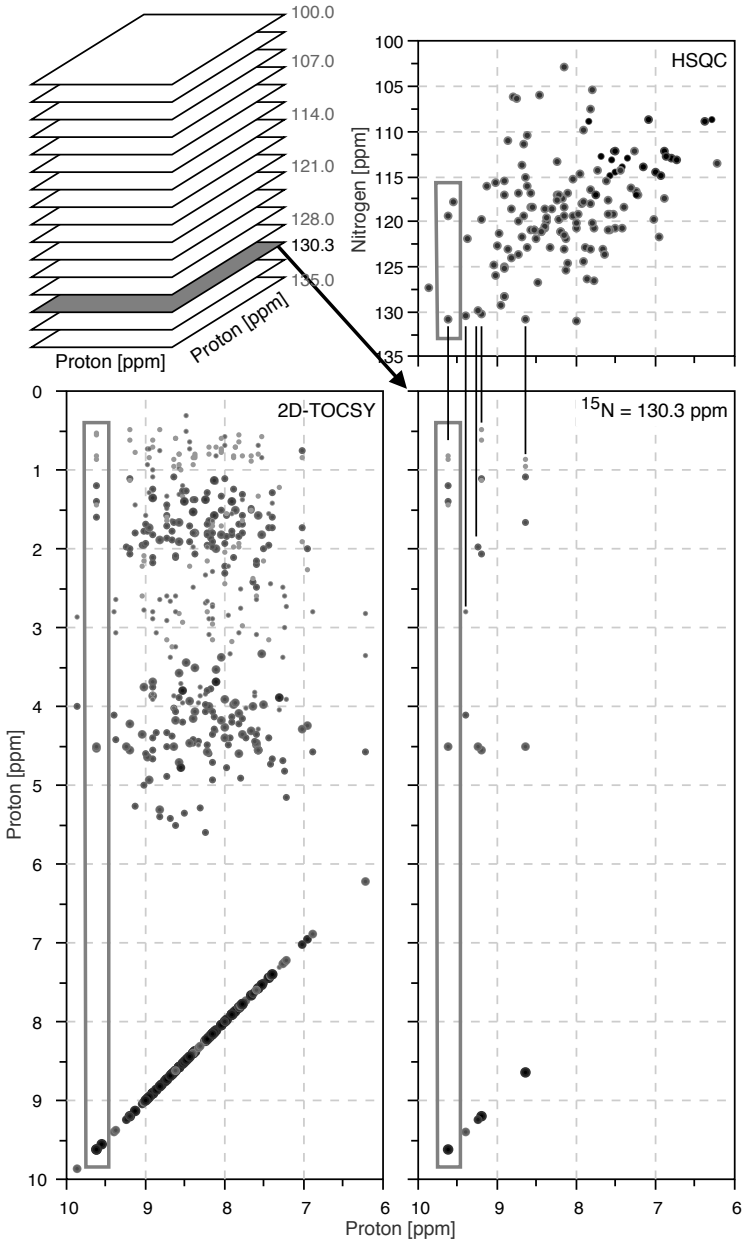


Figure 6.5. Increased resolution in a 3D-^{15}N separated TOCSY. The amide-aliphatic region of a 2D-TOCSY spectrum of a 130 residue protein is shown in the lower left. The 3D spectrum, represented as a stack of individual 2D-spectra at various ^{15}N frequencies is shown in the upper left. One of these spectra, or slices, at a nitrogen frequency of 130.3 ppm, is shown on the lower right. For reference, a 2D proton-nitrogen HSQC spectrum is shown in the upper right. The slice from the 3D experiment contains five residues with ^{15}N shifts of ≈130.3 ppm. These residues are highlighted by the vertical lines in the slice from the 3D-TOCSY spectrum. Three residues, highlighted with boxes in the 2D-TOCSY and HSQC spectra, have a proton chemical shift of ≈9.6 ppm and are degenerate in the 2D-TOCSY, but become resolved in the 3D-TOCSY spectrum. The boxed peaks in the slice from the 3D-TOCSY arise from a single residue.

6.3 Making the Connection - Determining Adjacent Spin-systems

There are two methods of determining inter-residue connectivity, through space dipolar coupling or through bond scalar coupling.

Typically, through-space connectivity is obtained using two-dimensional homonuclear NOESY experiments. The resolution, but not the information content, of the NOESY experiment is often increased by the addition of a third dimension by recording the ^{15}N shift of the amide nitrogen. Inter-residues connectivities obtained from the NOESY experiment reflect both local (e.g. sequential) as well as long-range interactions between distant residues, making the assignment process more difficult.

The use of scalar coupling across the peptide bond requires labeling with both ^{15}N and ^{13}C. Since the coupling is completely independent of the structure of the protein, and provides unambiguous inter-residue connectivities, it provides a reliable method of obtaining assignments. Another favorable attribute of triple resonance experiments is the larger J-coupling between heteronuclear spins. The larger J-couplings permits a more rapid transfer of magnetization between spins, making triple resonance experiments more sensitive in larger proteins that have shorter spin-spin relaxation times. Additionally, the presence of additional NMR active nuclei, such as ^{15}N and ^{13}C, permit the separation of the resonance signals over additional chemical shift scales, increasing the resolution of experiments. Finally, the carbon frequencies can also be used to readily identify the likely amino acid type of the spin-system.

6.3.1 Inter-residue connectivity via NOESY

The intensity of crosspeaks in the NOESY experiment is proportional to $1/d^6$, where d is the inter-proton distance. Typically, distances of 5 Å or less can be detected with this experiment. The measured inter-proton distances are usually referred to as NOEs. With respect to resonance assignments, NOEs involving the amide proton and/or the H_α proton are typically utilized because these resonances are in a less crowded region of the spectrum.

A number of inter-proton distances that are typically observed for mainchain atoms are shown in Fig. 6.6 and listed in Table 6.1. In the case of a β-sheet, the strongest NOE between adjacent residues is the NOE between the amide proton and the H_α proton of the *preceding* residue. The NOE between amide protons on adjacent residues in a β-strand conformation is comparably weak and is therefore not generally useful for confirming adjacent spin-systems. Note that the amide proton also shows two moderately intense NOEs to protons on the adjacent strand. These could be interpreted, incorrectly, as NOEs between adjacent residues, causing errors in the assignments.

In the case of the α-helix, the NOE between adjacent amide protons is strong due to the short inter-proton distance of 2.8 . The NOE between the H_α proton and the amide proton of the following residue is somewhat weaker (d = 3.5), but still useful for verifying the inter-spin-system connection. An

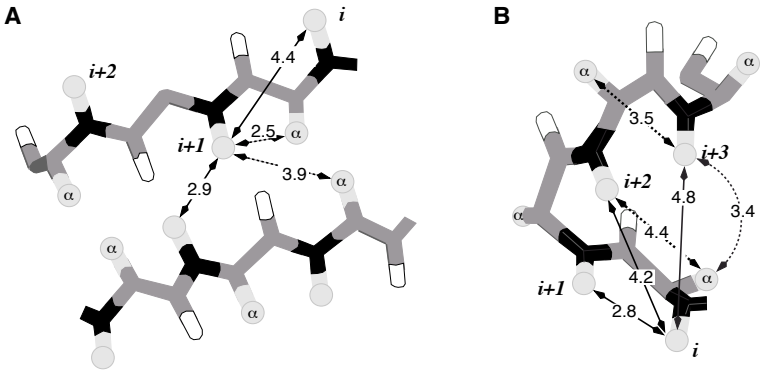


Figure 6.6. Inter-proton distances in regular secondary structures. Inter-residue distances in an β-sheet (Panel A) and an α-helix (Panel B). Inter-proton distances are shown in Angstroms. All of these distances should be detectable in a NOESY experiment, however longer distances will give weaker crosspeaks. Light gray atoms represent amide and H_α protons, dark gray atoms are carbon, white atoms with a black outline are carbonyl oxygens, black atoms are amide nitrogens.
Panel A: An amide proton will typically show four NOEs, two to the preceding residue and two to non-sequential residues across the strand. The sequential NOE between the amide and the H_α-proton of the preceding residue is intense because of the short (2.5) distance between the two protons.
Panel B: The first amide in the helix is labeled i and distances to the amide protons of the next three residues are indicated with solid arrows. The H_α proton is relatively close to the H_N proton of the next residue (3.5 , dashed line, shown for H_α^{i+2} to H_N^{i+3}) as well as to the amide proton of the $i+3$ residue.

amide proton is also close to the amide proton of the *i+2* (4.2) and the *i+3* residue (4.7), which provides redundant information regarding the sequential ordering of spin-systems.

Table 6.1. Selected inter-residue distances for sequential assignments. The first index of d_{ij} gives the atom on the *i-th* residue, the second index gives the atom on the *i+n* residue, where n varies from 1 to 4. Using the first entry of the table as an example: In an α-helix the H_α proton on the i^{th} residue is 3.5 from the amide proton (N) of the $i+1$ residue.

Secondary Structure	*Interacting Atoms*	*i+1*	*i+2*	*i+3*	*i+4*
α-helix	$d_{\alpha N}$	3.5	4.4	3.4	4.2
	d_{NN}	2.8	4.2	4.8	6.1
β-strand	$d_{\alpha N}$	2.2			
	d_{NN}	4.3			

†The numbers following the atom descriptions give the position of the residue in the turn.

6.4 Inter-residue Connectivity via Scalar Coupling

Triple-resonance experiments, involving ^{15}N, ^{13}C and ^{1}H spins, provide a much more reliable and robust method of obtaining resonance assignments. In the case of main chain assignments, the direct and relatively large scalar coupling between carbon and nitrogen across the peptide bond can be used to directly link spin systems. In the case of sidechain assignments, the large carbon-carbon coupling can be used to efficiently pass magnetization throughout the sidechain using isotropic (TOCSY) mixing. Alternatively, INEPT-like transfers between carbon atoms can be used to transfer magnetization in a well controlled fashion. Fig. 6.7 provides a summary of the heteronuclear coupling constants that can be exploited for the transfer of magnetization between spins.

Nomenclature: A large number of triple-resonance pulse sequences have been devised for mainchain assignments and a number of these are listed in Table 6.2. The pathway of magnetization transfer in a small number of these is shown in Fig. 6.8. The usual nomenclature for these experiments is to list the nuclei in the order that frequency labeling occurs in the pulse sequence. If magnetization is passed through a spin with no frequency labeling then that spin is enclosed in parenthesis. In addition, an experiment that begins with the magnetization on the amide proton, indicated as 'HN' in the experiment name, generally implies an 'out-and-back' experiment, i.e. the magnetization is returned to the amide proton for detection. For example the HN(CO)CA experiment would transfer the magnetization from the amide proton to the nitrogen, record the frequency of the N spin, pass the magnetization through the carbonyl carbon (CO), record the frequency of the C_α spin, and then return the magnetization back through the carbonyl carbon to the NH proton for detection. In contrast, if the magnetization begins elsewhere, it is usually passed in one direction to the amide proton for detection. For example the (HA)CA(CO)NH experiment would begin by transferring the α-proton magnetization to the α-carbon, followed by recording the chemical shift of the α-carbon. The magnetization is

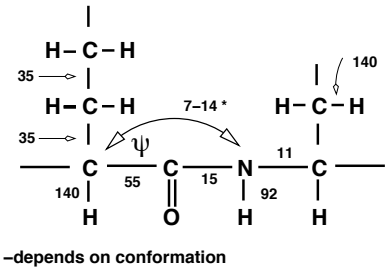


Figure 6.7. Heteronuclear scalar couplings in proteins. The coupling constants, in units of Hz, are indicated adjacent to the bond that joins the coupled spins. All of these couplings are one bond couplings, with the exception of the two bond coupling between the amide nitrogen and the C_α carbon of the preceding residue. All of the one-bond couplings are essentially independent of the secondary structure. In contrast, the two-bond coupling between the nitrogen and the C_α carbon, which depends on the ψ angle. Note that the two-bond coupling between the amide nitrogen and its own carbonyl carbon is essentially zero, thus it is only practical to directly correlate the amide nitrogen shift with the carbonyl shift of the preceding residue.

then transferred to the carbonyl carbon and then to the amide nitrogen for frequency labeling, with final detection on the amide proton on the following residue.

Each of the triple-resonance experiments listed in Table 6.2 generate a three dimensional spectrum. In cases where detection is on the amide proton, the amide nitrogen and proton frequencies generally comprise the second and third frequency dimension and the remaining dimension corresponds to the matching atom. A slice from this three dimensional spectrum at the nitrogen frequency of an amide group will show crosspeaks at the intersection of the amide proton chemical shift and the shift of the third spin (matching atom). Some experiments, such as the HNCA experiment, give both the inter- and intra-residue shifts for the matching atom, and thus will show two crosspeaks for each amide group. In order to unambiguously identify whether the peak originates from the inter- or intra-residue spin requires data from a complementary experiment that generates signals from only one of the two matching atoms. For example, the HN(CO)CA experiment complements the HNCA experiment by providing only inter-residue α-carbon shifts.

The general appearance of these spectra is illustrated in Fig. 6.9, using the HNCA and the HN(CO)CA experiments as an example. Other triple-resonance spectra would be very similar in appearance, with the significant difference

Table 6.2. Triple-resonance experiments for assignments.

Experiment	*Correlated Atoms*
Mainchain Carbon Shifts[†]	
HNCO	CO $^{i-1}$ with NH[‡]
HN(CA)CO	CO with NH
HNCA	$C_\alpha^{i,\ i-1}$ with NH
HN(CO)CA	C_α^{i-1} with NN
HNCACB	$[C_\beta C_\alpha]^{i,\ i-1}$ with NH
CBCANH	$[C_\beta C_\alpha]^{i,\ i-1}$ with NH
HN(CO)CACB	$[C_\beta C_\alpha]^{i-1}$ with NH
Mainchain Proton Shifts	
HN(CA)HA	$H_\alpha^{i,\ i-1}$ with NH
HN(COCA)HA	H_α^{i-1} with NH
HAHB(CO)NH	$[H_\beta H_\alpha]^{i-1}$ with NH
HAHBNH	$[H_\beta H_\alpha]^{i}$ & $[H_\beta H_\alpha]^{i-1}$
HCACO	C_α, CO, H_α
Other Sidechain Shifts	
HCCH TOCSY	All protons with H_α & C_α
C-C TOCSY (CO)NH	All C^{i-1} with NH
H-H TOCSY (CO)NH	All H^{i-1} with NH

[†] The mainchain includes the β-proton and carbon. [‡] "NH" implies correlation to both amide proton and nitrogen.

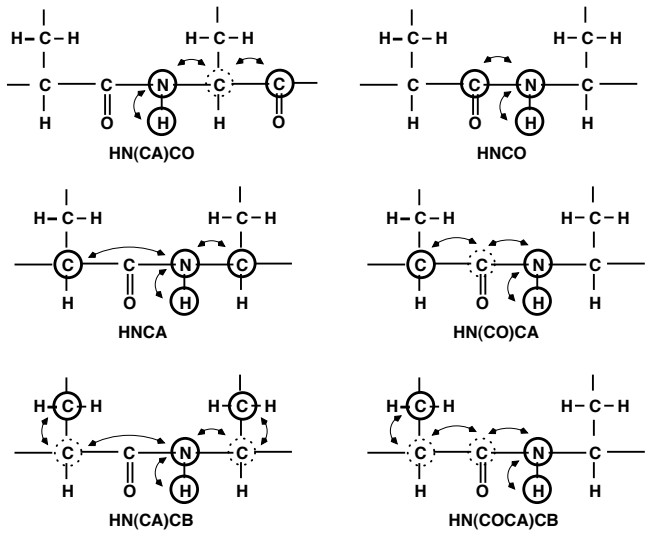


Figure 6.8. Triple-resonance experiments for main chain assignments. The magnetization transfer pathways in selected triple-resonance experiments are shown. In these diagrams the atoms that are circled define the recorded chemical shifts. The atoms that are circled with a dotted line serve to transfer magnetization between atoms, but their chemical shift is not recorded. The arrows indicate the direction of magnetization transfer. In all cases these experiments begin by transferring the amide proton polarization to the amide nitrogen. Experiments that give exclusively inter-residue shifts are on the right. These experiments generate one crosspeak per amide group. For the two experiments in the first row, the HN(CA)CO and HNCO, the matching atom is the carbonyl carbon. The HN(CA)CO can in principal give both inter- and intra-residue carbonyl shifts. However, the inter-residue peak is generally of low intensity and often not observable. The HNCO experiment gives the chemical shift of the carbonyl of the preceding residue. The HNCA and HN(CA)CB experiments give both the intra- and inter-residue α-carbon (HNCA) or β-carbon (HN(CA)CB) shifts. Thus two crosspeaks are observed per amide group. The HN(CO)CA and HN(COCA)CB experiments only correlate the inter-residue α-carbon (HN(CO)CA) or β-carbon (HN(COCA)CB) shift to the amide group, giving one crosspeak per residue.

being the carbon frequency axis. In the HNCO and HN(CA)CO experiments carbonyl shifts would be observed while in the HN(CA)CB and HN(COCA)CB experiments the β-carbon shifts would be observed.

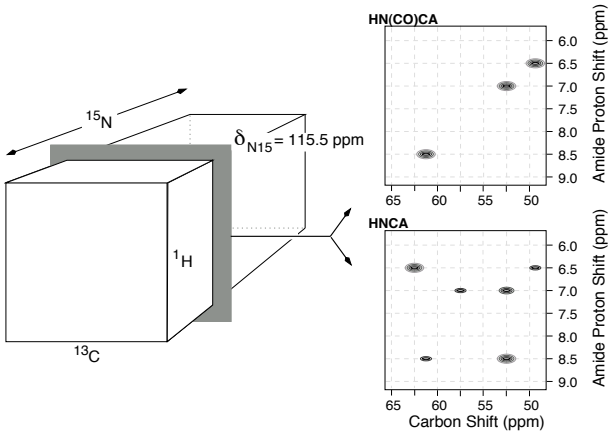
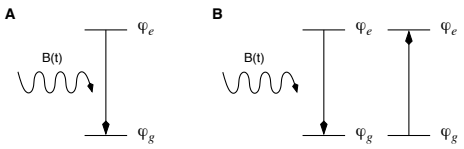


Figure 6.9. Illustration of HNCA and HN(CO)CA Spectra. The left section of the diagram shows the three-dimensional spectrum with the three frequency axis labeled. A single slice from this spectrum is taken at a nitrogen frequency of 115.5 ppm and the resultant two dimensional slices are shown on the right. The HN(CO)CA spectrum is shown on the top and the HNCA spectrum is on the bottom. This protein contains three residues whose amide nitrogen has a chemical shift of 115.5 ppm. The amide proton frequencies of these residues are 6.5, 7.0, and 8.5 ppm. The HN(CO)CA spectrum gives the chemical shifts of the α-carbon that precedes each residue. In this case the shifts are 48.75, 52.5, and 60.25 ppm for the first, second, and third amide proton shifts, respectively. The HNCA spectrum gives both the inter- and intra-residue shifts for each residue. The intra-residue crosspeak is usually more intense than the inter-residue peak because the one bond intra-residue coupling is generally usually larger than the two-bond inter-residue coupling. In this example, the opposite is true for the residue with H_N = 7.0 ppm. Using the HN(CO)CA spectrum to unambiguously identify the intra-residue peak in the HNCA spectrum gives intra-residue α-carbon shifts of 62.5, 57.5, and 52.5 ppm for the first, second, and third amide proton shifts, respectively.

Chapter 7

NUCLEAR SPIN RELAXATION AND MOLECULAR DYNAMICS

Nuclear spin relaxation is a consequence of time-dependent fluctuations in the magnetic field. These fluctuations can cause the net loss of energy from the excited state)A) or enhance the rate of spin-spin exchange (B).



The principal goal of NMR relaxation studies is to characterize molecular motion. The most common application is to determine the extent and rate of internal motion at specific sites within a protein. The type of information that is usually obtained from relaxation studies is:

1. τ_r - the overall rotational correlation time of the entire protein.
2. τ_i - the effective internal rotational correlation time at each labeled site.
3. S^2 - the order parameter for each labeled site. The order parameter indicates the extent of the internal motion: S^2=1 immobile, S^2=0 freely mobile

In relaxation measurements, the information (τ_r, τ_i, S) is obtained by comparing measured relaxation parameters to those calculated from models of the motion. Therefore, it is important to be able to predict the relaxation of the spins with reasonable accuracy. Consequently, the relaxation rates of heteronuclear spins, such as ^{15}N and ^{13}C, are generally measured because their relaxation is dominated by the *dipolar coupling* to the attached proton and by their own *chemical shift anisotropy*. The effect of both of these interactions on relaxation can be calculated with a high degree of accuracy.

7.1 Effect of Relaxation on the Evolution of Magnetization

There are two distinct relaxation mechanisms:

1. Spin-lattice relaxation. This is the rate by which the energy of the excited state is lost to the surroundings as the spins return to the original Boltzmann equilibrium. It is characterized by a time constant T_1 (see Fig. 7.2).

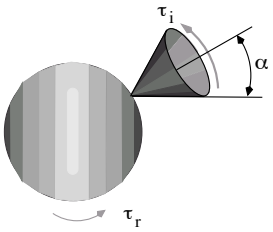


Figure 7.1. Framework for the analysis of internal motion. The rotational diffusion of the entire protein (large oval) is characterized by the overall rotation time, τ_r. The cone protruding off of the protein represents internal motion of a N-H bond. The semi-angle, α of the cone is related to the order parameter: $S = \left[\frac{1}{2}(1 + cos\alpha)cos\alpha\right]$. The rate that the N-H group diffuses within this cone is characterized by τ_i, the correlation time for internal motion.

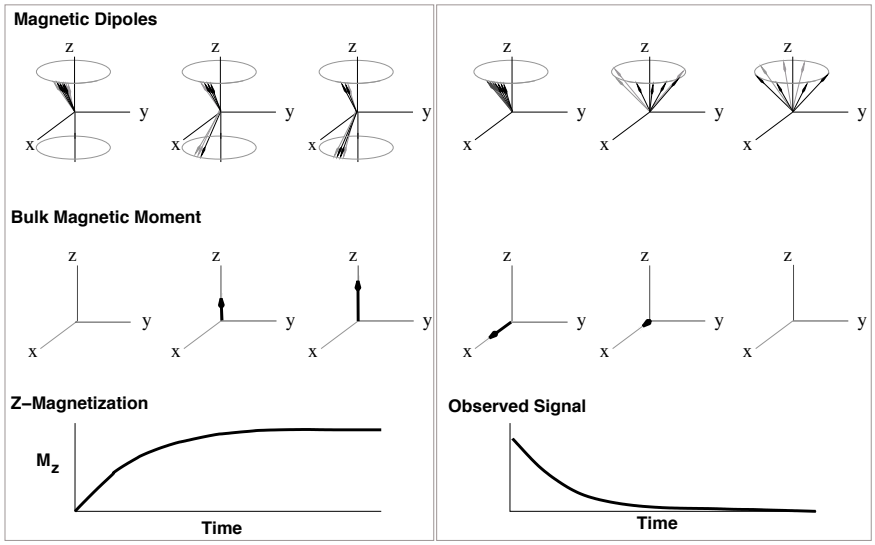


Figure 7.2. Relaxation of nuclear spins. Each panel shows, from top to bottom, the individual nuclear moments as a function of time, the bulk magnetization as a function of time, and the z-magnetization (M_z) or the bulk magnetization in the x-y plane ($\sqrt{M_y^2 + M_x^2}$), as a function of time. The left panel illustrates spin-lattice relaxation in the absence of spin-spin relaxation. The right panel shows spin-spin relaxation in the absence of spin-lattice relaxation.

2. Spin-spin relaxation. This is the rate by which a coherent population of spins becomes incoherent. Spin-spin relaxation is characterized by a time constant T_2 (see Fig. 7.2).

Measurement of the spin-lattice and spin-spin relaxation times provides information on molecule motion. In addition to these two observations, additional information on molecular motion is provided by the nuclear Overhauser effect (NOE). This is a measurement of the effect of proton irradiation on the population difference of nearby spins. The heteronuclear NOE (hnNOE) involves proton-nitrogen or proton-carbon interactions.

7.1.1 Bloch Equations

For the moment we will treat T_1 and T_2 as measurable quantities, but we will avoid any discussion about their relationship to structure and dynamics. To develop analytical equations which describe the motion of the magnetic moment in the presence of the static B_o field we will use the following approach that was originally proposed by F. Bloch in 1946 [**?**]. This approach begins with the classical description of the evolution of the magnetization:

$$\frac{d\vec{M}}{dt} = \gamma\vec{M} \times B \tag{7.1}$$

The decay of magnetization due to relaxation is added to each component of magnetization:

$$\begin{aligned}
\frac{dM_z}{dt} &= \frac{M_o - M_z}{T_1} + \gamma(M \times B)_z \\
\frac{dM_x}{dt} &= \frac{-M_x}{T_2} + \gamma(M \times B)_x \\
\frac{dM_y}{dt} &= \frac{-M_y}{T_2} + \gamma(M \times B)_y
\end{aligned} \tag{7.2}$$

The Bloch equations describe a return of the z-magnetization to the equilibrium value, M_o, with a time constant of T_1, and a decay of the transverse magnetization with a time constant of T_2. In the rotating frame these equations become:

$$\frac{\delta M_z}{\delta t} = \frac{M_o - M_z}{T_1} \tag{7.3}$$

$$\frac{\delta M_x}{\delta t} = \frac{-M_x}{T_2} + M_y(\omega_s - \omega) \tag{7.4}$$

$$\frac{\delta M_y}{\delta t} = \frac{-M_y}{T_2} - M_x(\omega_s - \omega) \tag{7.5}$$

where $\omega_s = \gamma B, \omega = -\Omega$.

These equations are most easily solved by defining a function:

$$M^+ = M_x + iM_y \tag{7.6}$$

Adding eqs. 7.4 and 7.5 gives the following:

$$\frac{\delta M^+}{\delta t} = -M^+ \left[\frac{1}{T_2} + i\omega'\right] \tag{7.7}$$

$\omega' = \omega_s - \omega$, is the precessional rate of the spin in the rotating frame. In practice, this is actually the frequency of the signal obtained from the spectrometer. The solution to Eq. 7.7 is:

$$M^+ = e^{-i\omega' t}\ e^{-t/T_2} \tag{7.8}$$

Taking into account the initial conditions immediately after a 90° x-pulse (M_x & M_z=0) gives the following final solutions:

$$\begin{aligned} M_x(t) &= \sin(\omega' t)\ e^{-t/T_2} \\ M_y(t) &= \cos(\omega' t)\ e^{-t/T_2} \\ M_z(t) &= M_o[1 - e^{-t/T_1}] \end{aligned} \tag{7.9}$$

The above shows that the magnetization along the z-axis will grow with an exponential time constant of T_1 while the transverse magnetization will decay with a time constant of T_2; exactly the behavior that is depicted in Fig. 7.2.

7.2 Measurement & Data Analysis

The usual approach for the analysis of molecular motion begins with a general model of internal motion that contains only two parameters (τ_i, S) per residue plus a value for the overal rotation time, τ_r that would apply to all residues. These parameters are obtained from the relaxation data by following the scheme that is diagrammed in Fig. 7.3 consisting of the following steps:

1. Collect T_1, T_2, and hnNOE data for each residue.
2. Select a set of initial guesses for τ_r, τ_i, and S and calculated expected T_1,T_2, and hnNOE from these parameters.
3. Compare the measured relaxation parameters to those calculated.
4. Adjust τ_r, τ_i, and S^2 until the calculated and measured relaxation rates agree.

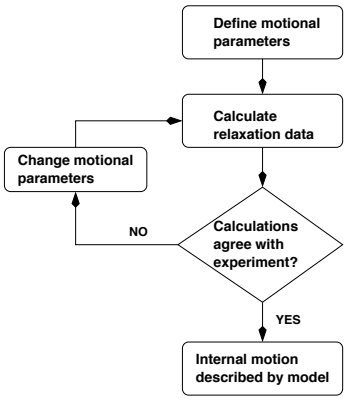


Figure 7.3. Flow chart for analysis of relaxation data. This flow chart indicates how parameters that describe internal motion are obtained from relaxation measurements.

7.3 Theory of Relaxation

Transition between the excited and ground states can occur by two independent mechanisms, spontaneous emission and stimulated emission. The rate of

spontaneous emission is proportional to the third-power of the absorption frequency. At typical proton NMR resonance frequencies, in the megahertz range, the rate of spontaneous emission is essentially zero. Therefore, the major source of relaxation for excited NMR spins is from *stimulated* emission.

Stimulated emission requires the presence of an oscillating electromagnetic field whose frequencies are matched to the absorption frequencies of the NMR transitions. The oscillating fields are created by random rotational motion of the molecule or by internal motions within the molecule. Such motions can generate a time dependent magnetic field by two mechanism, the anisotropic chemical shift (chemical shift anisotropy, CSA) that is associated with a single spin and the dipolar coupling between nearby spins. Both of these mechanism operate simultaneously for all types of spins, however dipolar coupling is the principal relaxation mechanism for the proton because of its relatively small chemical shift anisotropy.

7.3.0.1 Chemical Shift Anisotropy

The external B_o field is reduced at the nucleus by the surrounding electron density, by a shielding factor σ, giving an observed chemical shift of $\omega_s = \gamma(1-\sigma)B_o$. If the electron density is the same in all directions (isotropic) then a change in the orientation of the spin will have no effect on the shielding of the nucleus, and therefore no effect on the magnetic field at the nucleus.

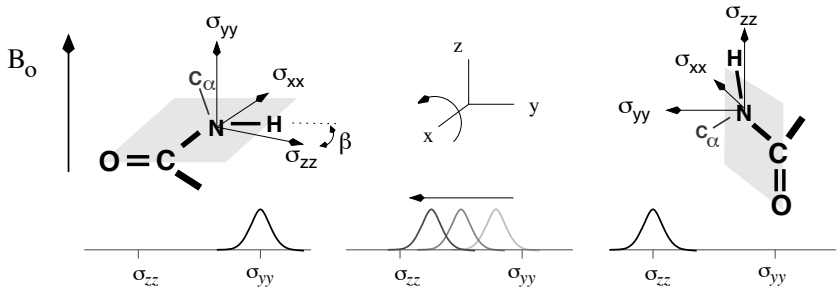


Figure 7.4. Effect of orientation on the amide chemical shift. The orientation of the peptide group affects the observed chemical shift of the amide nitrogen. The chemical shift can be obtained from the chemical shift tensor. This tensor, in the principal axis system is superimposed on the molecular structure. The z-axis of the PAS lies in the plane of the peptide bond (shaded gray) and is rotated away from the N-H bond vector by the angle β. The x-axis of the PAS is in the plane of the peptide bond and the y-axis is perpendicular to the peptide plane. If the tensor is axially symmetric, $\delta_{zz} = \delta_{\parallel}$ and $\delta_{\perp}$ lies in the plane defined by δ_{xx} and δ_{yy}. In the leftmost structure the orientation of the amide group is such that δ_{yy} is parallel to the static field (B_o). Therefore the measured chemical shift is δ_{yy}, as illustrated in the spectrum in the lower part of the figure. As the molecule rotates about the x-axis the peak position will move from δ_{yy} to δ_{zz}, as shown in the middle of the diagram. After a 90° rotation, the z-component of the chemical shift tensor becomes parallel to the static field, and the observed chemical shift is δ_{zz}.

If the shielding is anisotropic, then different orientations of the molecule will generate different magnetic fields at the nucleus. The actual field will depend on the orientation of the molecule with respect to the external magnetic field. To characterize the anisotropic nature of the shielding we will define a chemical shift tensor, $\tilde{\sigma}$, which will give the chemical shift for any given orientation of the molecule with respect to the B_o field. In one particular orientation of the molecule with respect to the magnetic field, the principal axis system (PAS), the chemical shift tensor takes on a simple form:

$$\tilde{\delta}_{PAS} = \begin{bmatrix} \delta_{xx} & 0 & 0 \\ 0 & \delta_{yy} & 0 \\ 0 & 0 & \delta_{zz} \end{bmatrix} \tag{7.10}$$

The diagonal form of the chemical shift tensor gives directly the chemical shift that would be observed if the magnetic field were along the x-axis (δ_{xx}), y-axis (δ_{yy}), or the z-axis (δ_{zz}) of the principal axis system (see Fig. 7.4).

If the molecule tumbles rapidly and isotropically in solution, then the observed *isotropic* shift is simply the average of all three components:

$$\delta_{iso} = \frac{1}{3}[\delta_{xx} + \delta_{yy} + \delta_{zz}] \tag{7.11}$$

Although the average chemical shift is time-independent, the rapid rotation of atoms with anisotropic chemical shifts will result in a change in the instantaneous chemical shift, and therefore the magnetic field, at the nucleus.

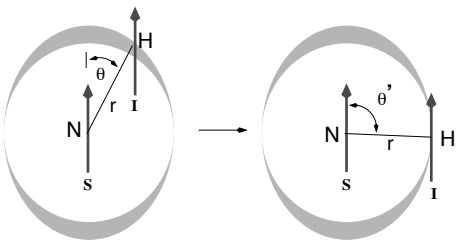


Figure 7.5. Dipole-dipole coupling. The magnetic field generated by the amide nitrogen generates an additional field, B_{dipole} at the amide proton. The strength of this field depends on the relative orientation of the two spins, as illustrated by the gray crescent shapes. When $\theta = 0$, B_{dipole} is at a maximum. The dipole field is zero when $\theta \approx 54°$, and is negative for $0 < |\theta| < 54°$. Molecular tumbling changes the relative orientation of the two spins, resulting in a change of B_{dipole}.

7.3.1 Dipolar Coupling

Dipolar coupling arises when the magnetic field of one nuclear spin affects the local magnetic field of another spin. The magnetic field generated by the I spin is given by the classic equation for a dipole field:

$$\begin{aligned} B_{dipole}(t) &= \frac{\mu}{r(t)^3}(3cos^2\theta(t) - 1) \\ &= \frac{\gamma\hbar}{r(t)^3}(3cos^2\theta(t) - 1) \end{aligned}$$

where θ and r are defined in Fig. 7.5. The energy associated with this additional magnetic field is obtained from

$E = \hbar\omega = \hbar\gamma_S B_{dipole}$:

$$\mathcal{H} = \hbar^2 \frac{\gamma_S \gamma_I}{r(t)^3}(3cos^2\theta(t) - 1) \tag{7.12}$$

The intensity of the dipole field, B_{dipole}, depends on both the orientation of the two spins and the distance between them. If either of these properties are time dependent then the magnetic field will vary with time.

Most relaxation studies take advanage of the fact the contribution of dipolar interactions to the relaxation properties of heteronuclear spins, such as ^{15}N or ^{13}C, is dominated by the attached proton due to the $1/r^3$ dependence. Consequently, dipolar fields that are generated by other protons are considered negligible. Secondly, the inter-atomic distance, r, is fixed to the bond length, consequently only fluctuations in the angular term contribute to time dependent field fluctuations.

7.3.2 Frequency Components from Molecular Rotation

Both chemical shift anisotropy and dipolar coupling generate magnetic field fluctuations at the nucleus. In order for these motions to be effective at stimulating nuclear transitions the random molecular motion must contain field fluctuations at the appropriate frequencies to stimulate transitions.

In the case of two spins that are coupled, there are six possible transitions between the four different states that represent two coupled spins (see Fig. 7.6. Four of these transitions are single quantum transitions because they involve flipping only one spin. Using an N-H pair for example, single quantum transitions will be stimulated by field fluctuation at $\omega = \omega_N$ or $\omega = \omega_H$. In addition, the double quantum (two spins flipped) and the single quantum transitions (interchange of spin states) will be stimulated by field fluctuations at $\omega = \omega_H + \omega_N$ and $\omega = \omega_H - \omega_H$, respectively. Each of these transitions

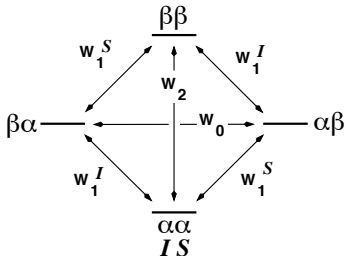


Figure 7.6. Energy level diagram for two coupled spins. The lowest energy state corresponds to both spins being in the $\alpha(m_z = +1/2)$ state. The four single quantum transitions, labeled with W_1^I or W_1^S, connect states that differ in the spin state of only one spin. There is one zero quantum transition (W_0) and one double quantum transition (W_2), corresponding to the net change in the magnetization of both spins by zero or two units, respectively.

is associated with a transition rate, W, that arises from time dependent field fluctuations.

Quantum mechanics can be used to derive the relationship between the field fluctuations and the transition rate. The end result is:

$$W \propto \int_o^t \mathcal{F}(t)\mathcal{F}(t-\tau)e^{-i\omega\tau}d\tau \tag{7.13}$$

In the above, $\mathcal{F}$ represents the time-dependent field fluctuations and the function $\mathcal{F}(t)\mathcal{F}^*(t+\tau)$ is an autocorrelation function. The autocorrelation function measures how rapidly knowledge of a prior orientation is lost by the reorienting molecule. For random fluctuations the autocorrelation function does not depend on a specific origin of time, only on the time difference between instances in time. Consequently, t can be set to zero and the auto-correlation depends only on τ. Usually the autocorrelation function is given the symbol $g(\tau)$.

To obtain the transition rates at specific frequencies it is necessary to extract the frequency components from the autocorrelation function by Fourier transformation. However, it is apparent from Eq. 7.13 that W is the Fourier transform of the autocorrelation function, therefore W already represents the frequency spectrum of the transition rates.

The Fourier transform of the normalized autocorrelation function is given a special name, the spectral density function, $J(\omega)$, because it represents the *density* of fluctuations at different frequencies. It is defined as follows:

$$J(\omega) = \int_0^\infty g(\tau)e^{-i\omega\tau}d\tau \tag{7.14}$$

An illustration of the relationship between rotational motion of the particle, magnetic field fluctuations due CSA, and the spectral density is shown in Fig. 7.7. A small molecule, which has a short correlation time, will experience rapid fluctuations in the magnetic field (Panel A, gray curve). These fluctuations will rapidly become uncorrelated, leading to a rapid decrease in the autocorrelation function (Panel B). Such rapid fluctuations contain both high and low frequency components, thus the resultant spectral density is broad and extends over a large frequency range (Panel C). In contrast, a large particle tumbles slowly, experiencing magnetic field fluctuations of lower frequencies (Panel A, black curve), hence the spectral density function will have a higher intensity at lower frequencies (Panel C).

In the case of dipolar coupling between spins, the complete expression for the transition probabilities is:

$$W(\omega) = \frac{\hbar^2\gamma_N^2\gamma_H^2}{r^6}\kappa J(\omega) \tag{7.15}$$

The κ term depends on whether the transition is a zero, single, or double quantum transition (see Fig. 7.6).

Once an autocorrelation function can be obtained, then the transitions rates can be readily calculated using the above expressions. For example, the autocorrelation and spectral density functions for random isotropic tumbling due

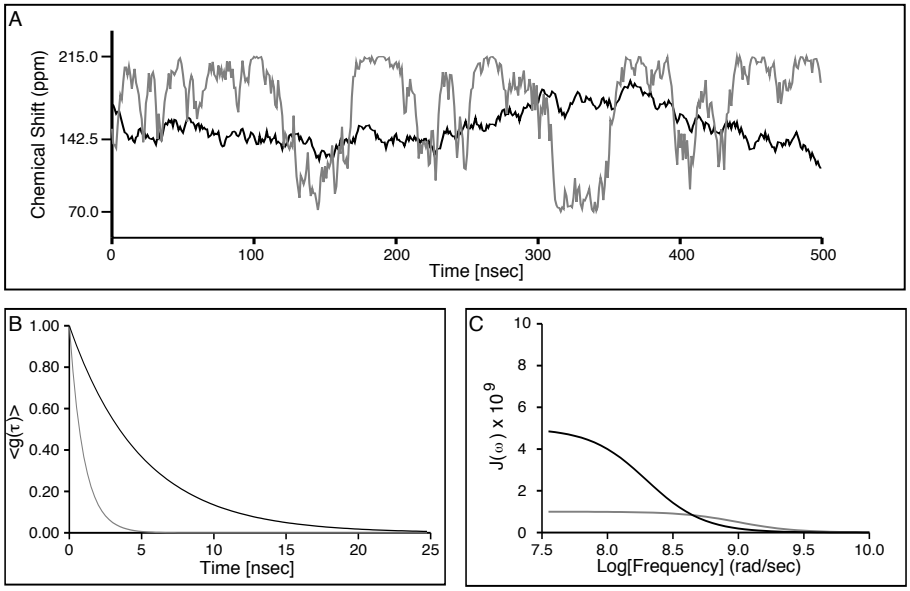


Figure 7.7. Magnetic field fluctuations and the spectral density function. The relationship between random magnetic field fluctuations, the auto-correlation function, and the spectral density function is shown. An axially symmetric chemical shift tensor with principal values of $\delta_\perp = 70$ ppm and $\delta_\parallel = 215$ ppm was used in the simulations. The fluctuations in chemical shift that occur due to random isotropic rotation of the amide group are shown in Panel A. The rotational correlation times are 1 nsec (gray line) and 5 nsec (black line). The resultant autocorrelation functions are shown in Panel B. Panel C shows the spectral density functions calculated from $< g(\tau) >$. High frequency motions (gray curve) result in smaller values of J(ω) at $\omega = 0$. However, J(ω) extends to higher frequencies.

to Brownian motion are given here:

$$g(\tau) = e^{-\tau/\tau_r} \qquad J(\omega) = \frac{\tau_r}{1 + \omega^2 \tau_r^2} \tag{7.16}$$

τ_r is the rotational correlation time. It is the time required for a molecule to rotate, on average, 1 radian and is proportional to the size of the protein and the viscosity of the solution:

$$\tau_r = 4\pi\eta a^3/3kT \tag{7.17}$$

where a is the radius of the protein, and η is the viscosity. It follows that τ_r is *approximately* 1 nsec for each 2.6 kDa of protein mass at T = 300 K.

If internal motion is present, the autocorrelation function is:

$$\begin{aligned} g(\tau) &= e^{-\tau/\tau_r}\left[S^2 + (1-S^2)e^{-\tau/\tau_i}\right] \\ &= (1-S^2)e^{-\tau/\tau_{mix}} + S^2 e^{-\tau/\tau_r} \end{aligned} \tag{7.18}$$

where S is the order parameter, τ_i is the correlation time for internal motion. This auto-correlation function gives the following spectral density:

$$J(\omega) = (1-S^2)\frac{\tau_{mix}}{1+\omega^2\tau_{mix}^2} + S^2\frac{\tau_r}{1+\omega^2\tau_r^2} \quad (7.19)$$

$$\tau_{mix} = \tau_i\tau_r/(\tau_i+\tau_r) \quad (7.20)$$

7.4 ^{1}H-^{1}H Homonuclear relaxation

7.4.1 Spin-lattice relaxation

Spin-lattice relaxation, which occurs with a time constant of T_1, is due to the stimulation of single- and double-quantum transitions as both of these provide a mechanism to re-populate the ground state. The T_1 is thus sensitive to field fluctuations at ω_s and $2\omega_s$, where ω_s is the resonance frequency of the spin. The dependence of the proton T_1 on the spectral density function is:

$$\frac{1}{T_1} = \frac{6}{20}d^2\left[J(\omega_H) + 4J(2\omega_H)\right] \quad (7.21)$$

where $d^2 = \hbar^2\gamma_H^4/r^6$ and r is the inter-proton distance.

7.4.2 Spin-spin relaxation

Spin-spin relaxation, which occurs with a time constant of T_2, is due to zero-, single-, and double-quantum transitions because all of these transitions lead to a loss of coherence in the spins. The T_2 is thus sensitive to field fluctuations that occur at $\omega = 0$ as well as at higher frequencies. The dependence of the proton T_2 on the spectral density function is:

$$\frac{1}{T_2} = \frac{3}{20}d^2\left[3J(0) + 5J(\omega_H) + 2J(2\omega_H)\right] \quad (7.22)$$

7.4.3 Molecular Weight effects on ^{1}H Relaxation

The effect of molecular weight (τ_r) and the magnetic field strength on proton relaxation times is illustrated in Fig. 7.8. Figure **??** shows the spin-spin relaxation time (T_2) for the amide nitrogen, carbonyl carbon, and C_α carbon as a function of these parameters. Note that at lower magnetic field strengths the amide nitrogen and carbonyl carbon have similar spin-spin relaxation times. In contrast, the C_α carbon is efficiently relaxed by its attached proton. However, at high magnetic field strengths, the contribution of the CSA to the carbonyl relaxation increases significantly because of the ω^2 dependence, leading to very short relaxation times for the carbonyl carbon at 900 MHz.

7.4.3.1 Effects of Molecular Weight on the Proton T_1

Since T_1 depends on $J(\omega)$ and $J(2\omega)$ the effect of the rotational correlation time on the observed T_1 can be explained by considering the intensity of the spectral density at ω and 2ω for various values of τ_r. Beginning with fast

motions, or short τ_r values, the spectral density is of low intensity at these two frequencies (solid curve in Fig. 7.8, Panel B). Therefore, spin-lattice relaxation is inefficient and the T_1 is relatively long. As the rotational motion slows, the profile of the spectral density changes becoming more intense at lower frequencies. When $\omega\tau_r$ is approximately one, the spectral density at ω and 2ω is large (dotted line in Panel B Fig. 7.8), and efficient spin-lattice relaxation occurs, giving a minimum in the T_1. As the motion becomes even slower, the intensity of $J(\omega)$ first drops at 2ω (dashed curve) and then eventually at ω. The small values of $J(\omega)$ at these two frequencies will cause inefficient spin-lattice relaxation, hence a longer T_1 for large molecules.

The magnetic field dependence on the T_1 is explained in a similar manner. For any values of τ_r corresponding to a molecular weight greater than 10 kDa, a higher magnetic field strength will cause J(ω) to be sampled at higher frequencies. Since J(ω) always decreases as ω increases, the spin-lattice relaxation will be less efficient and T_1 will become longer. Consequently, longer recycle delays need to be used at higher field strengths if the same level of steady-state magnetization is desired.

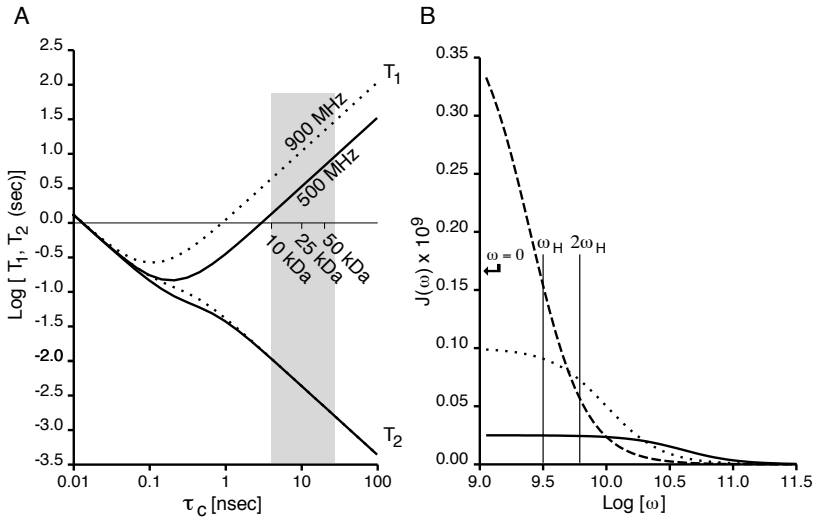


Figure 7.8. Effect of molecular weight on proton relaxation. Panel A shows proton T_1 and T_2 values as a function of the rotational correlation time, τ_r (note the *log* scale). These were computed assuming a spherical protein and an inter-proton distance of 1.7 . This value approximates the average proton density surrounding an amide proton. The individual curves for T_1 and T_2 are labeled, the solid line represents relaxation at a proton frequency of 500 MHz while the dotted line represents relaxation at 900 MHz. The gray rectangle indicates the range of molecular sizes, from 10 kDa to 60 kDa, that are routinely studied. Panel B shows spectral density functions for rotational correlation times of 0.025 (solid), 0.10 (dotted), and 0.40 nsec (dashed). These times are near the T_1 minimum. The ω_H and $2\omega_H$ values for a proton frequency of 500 MHz are indicated.

7.4.3.2 Effects of Molecular Weight on the Proton T_2

In contrast to the parabolic-like behavior of T_1, the T_2 for the proton (Fig. 7.8) decreases as the rotational correlation time increases. This relationship is a result of the contribution of J(0) to the spin-spin relaxation rate. As the rotational correlation time becomes longer, the spectral density at $\omega = 0$ also increases. In fact, for most proteins the spectral density at zero frequency dominates the spin-spin relaxation. For larger proteins, the proton T_2 can be approximated as:

$$\frac{1}{T_2} \approx \frac{3d^2}{20} 3J(0) = \frac{9d^2}{20}\tau_r \tag{7.23}$$

Since the rotational correlation time, τ_r is proportional to the molecular weight, the T_2 is inversely proportional to the molecular weight:

$$T_2 \propto 1/MW \tag{7.24}$$

as the size of the molecule increases, the T_2 becomes shorter. Recall that the linewidth is related to the T_2:

$$\Delta\nu = \frac{1}{\pi T_2} \tag{7.25}$$

Therefore larger proteins have broader lines.

7.5 Heteronuclear Relaxation

The equations that describe heteronuclear relaxation are very similar to those that describe proton-proton relaxation. The spin-lattice relaxation due to dipolar coupling is given by:

$$\frac{1}{T_1^{Dipole}} = d^2 \frac{1}{10} \left[J(\omega_I - \omega_S) + 3J(\omega_S) + 6J(\omega_I + \omega_S) \right] \tag{7.26}$$

where $d^2 = \gamma_H^2 \gamma_N^2 \hbar^2 / r^6$.

In the case of nuclei with significant CSA , it is necessary to add the contribution of the CSA to the spin-lattice relaxation [**?**]:

$$\frac{1}{T_1^{CSA}} = \frac{2}{15}\omega_S^2 \Delta\sigma^2 J(\omega_S) \tag{7.27}$$

$\Delta\sigma$ is the CSA, and ω_S is the resonance frequency of the S spin. $J(\omega_S)$ appears because field fluctuations that can cause single quantum transition of the heteronuclear spin provide a mechanism for the excited state to release energy and return to the ground state.

The overall spin-lattice relaxation rate (R_1=$1/T_1$) is the sum of the rate due to dipolar coupling and CSA. Using the amide group as an example:

$$R_1 = \frac{d^2}{10} \left[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N) \right] + \frac{2}{15}\omega_S^2 \Delta\sigma^2 J(\omega_N) \tag{7.28}$$

7.5.1 Spin-spin Relaxation

The spin-spin relaxation rate of a heteronuclear spin is affected by dipolar coupling to the attached proton [**?**], chemical shift anisotropy [**?**], as well as chemical exchange. Using the amide group as an example (S=^{15}N, I=^{1}H):

$$\begin{aligned} R_2 &= \frac{d^2}{20}\left[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)\right] \\ &\quad + \frac{1}{45}\omega_N^2 \Delta\sigma^2 \left[4J(0) + 3J(\omega_N)\right] + R_{ex} \end{aligned} \tag{7.29}$$

where $d^2 = \gamma_H^2 \gamma_N^2 \hbar^2 / r^6$.

The $J(0)$ term, which is present for both the dipolar coupling and CSA terms, represents the dephasing of the transverse magnetization by an inhomogeneous local magnetic field. The $J(\omega_N)$ and $J(\omega_N + \omega_H)$ terms represent relaxation to the ground state which also destroys the transverse magnetization. The $J(\omega_H - \omega_N)$ term represents zero-quantum mutual proton-nitrogen spin flips, which cause dephasing of the nitrogen magnetization. Finally, the presence of the $J(\omega_H)$ term indicates that transitions of the coupled proton affects the coherence of the transverse nitrogen magnetization.

7.5.2 Heteronuclear NOE

The heteronuclear NOE is a measure of the change in the steady state populations of the heteronuclear spin (^{15}N or ^{13}C) when the attached proton is saturated. Of the three types of relaxation data that are measured (T_1, T_2, and hnNOE), the hnNOE is the most sensitive to internal motion, as shown in Fig. 7.9.

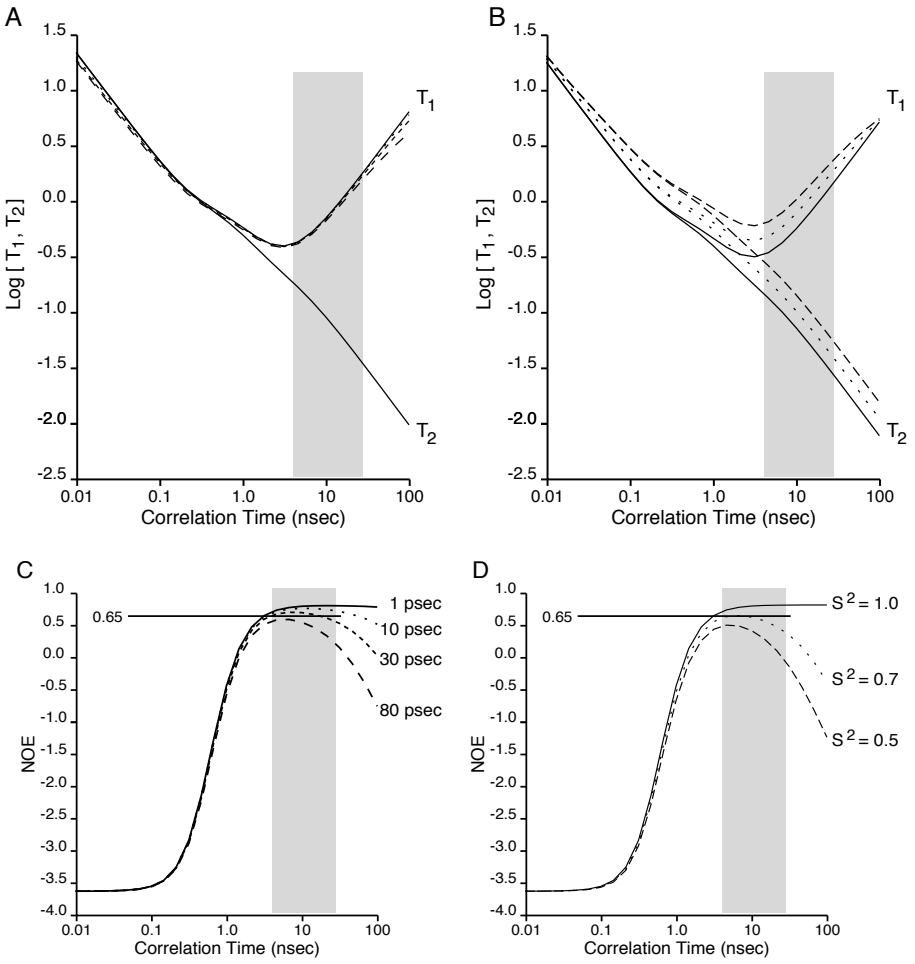


Figure 7.9. Effect of internal motion on ^{15}N relaxation. The effect of internal motion on T_1 and T_2 (panels A and B) and on the heteronuclear NOE (panels C and D). In each plot the gray box indicates the range of overall correlation time (τ_r) that are typically studied, spanning protein sizes from 10 kDa to 60 kDa. In panels A and C, the order parameter, S^2, was set to 0.8 and the correlation time for internal motion, τ_i was 1 (solid), 10 (dotted), 30 (dashed), or 80 psec (long dash). In the case of T_2, the different curves are indistinguishable. In panels B and D, the correlation time for internal motion was fixed at 30 psec and the order parameter was 1.0 (solid), 0.7 (dotted), and 0.5 (dashed). The horizontal line marked 0.65 indicates a cut-off value that is often used to determine whether a residue displays significant mobility.

Chapter 8

EXCHANGE PROCESSES

8.1 Introduction

The exchange between two or more environments can have a profound effect on the appearance of the resonance lines of the exchanging species. Under favorable conditions it may be possible to obtain information on both the forward and reverse rate-constants for the exchange reaction as well as the equilibrium population of each environment from the changes in the NMR spectrum.

When a spin exchanges between environments its spectral properties may change if the chemical shift or relaxation properties of the spin are different in each environment. Figure 8.1 shows how a conformational change in a protein can lead to a change in the chemical shift of the methyl group of a methionine residue.

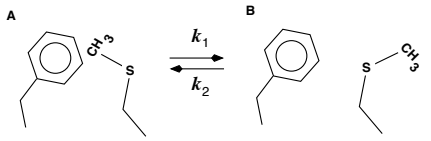


Figure 8.1. Effect of exchange on the environment of a Spin. In this illustration the methionine residue exists in two conformations. Since the environment of the methyl group is different in each environment, two different resonance frequency will be observed, ω_A and ω_B. In conformation A, the methyl group is found above an aromatic ring and will experience a ring current shift. In conformation B, the methyl group is removed from the aromatic ring, resulting in a change in its chemical shift. Note that the relaxation properties of the methy group may also differ between the two environments.

Since the exchange phenomenon involves a change in the chemical environment of a spin, leading to a chemical shift change, it is often referred to as *chemical exchange*. Usually, exchange processes are studied when there are *non*-covalent changes in the molecule. However, changes in the covalent structure of a molecule can also lead to chemical exchange.

In this chapter, three general aspects of exchange will be considered. First, we will discuss the effects of a spin sampling two distinct environments on its spectral properties. Here the goal is to measure the rate constants for exchange

and/or the equilibrium constant. In the latter part of the chapter we will extend these studies to the investigation of the kinetics and binding affinity of ligands. Finally, the exchange properties of the amide proton will be discussed with reference to measuring protein dynamics.

8.2 Chemical Exchange

The kinetics of the exchange reaction are defined by the following scheme:

$$\mathbf{A} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathbf{B}$$

The kinetic rate constant for the conversion of A to B is k_1 and the rate constant for the reverse reaction is k_2, giving an overall equilibrium constant, $\mathrm{K}_{eq} = k_1/k_2$. The fraction of the system found in each conformation is:

$$p_A = \frac{k_2}{k_1 + k_2} = \frac{1}{1 + K_{eq}} \qquad p_B = \frac{k_1}{k_1 + k_2} = \frac{K_{eq}}{1 + K_{eq}} \tag{8.1}$$

To characterize the different time scales of exchange it is useful to define an apparent exchange rate, $k_{ex} = k_1 + k_2$, and a frequency difference between the two states, $\Delta\omega = \omega_A - \omega_B$. Table 8.1 illustrates how the relationship between the apparent exchange rate, k_{ex}, and the frequency separation, $\Delta\omega$, will affect the observed spectrum. Note that the response of the system to chemical exchange depends on the ratio of the exchange rate to the frequency difference of the spins in each environment, i.e. $k_{ex}/\Delta\omega$.

8.2.1 Effect of Exchange on the NMR Spectrum

It is useful to consider in qualitative fashion two illustrative cases, fast exchange and slow exchange, to gain an understanding of the general features of chemical exchange.

Fast Exchange ($k_{ex} >> \Delta\nu$): Under conditions of fast exchange, a single resonance line is observed. The averaging of the chemical shift occurs because

Exchange Rate		*Observed Spectrum*
Very slow	$k_{ex} << \Delta\nu$	Two Resonances
Slow	$k_{ex} < \Delta\nu$	Two Broadened Resonances
Intermediate	$k_{ex} \approx \Delta\nu$	Complex Lineshape
Fast	$k_{ex} > \Delta\nu$	Single Broadened Resonance
Very fast	$k_{ex} >> \Delta\nu$	Single Resonance

Table 8.1. Summary of the effects of exchange on the properties of the NMR spectrum. The left columns describe the exchange in terms of the relationship between the frequency separation ($\Delta\nu = \frac{1}{2\pi}(\omega_A - \omega_B)$) and the apparent exchange rate constant ($k_{ex} = k_1 + k_2$). The remaining column describes the effect of exchange on the spectrum. A more comprehensive description of the methodologies that can be used to measure chemical exchange can be found in [**?**].

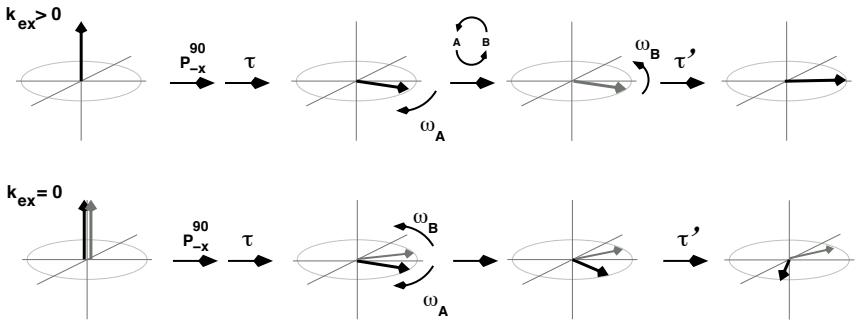


Figure 8.2. Effects of fast exchange on the environment of a spin. The upper part of the figure follows a single spin as it precesses during fast exchange. The lower part of the figure shows two spins, one in conformation A (black) and one in conformation B (gray), precessing in the absence of chemical exchange. In both cases the coordinate frame is rotating at a frequency that is midway between ω_A and ω_B and $p_A = p_B$. The left most part of the figure represents the system before the excitation pulse. Following from left to right, after the 90° pulse and a short period τ, the spins have precessed at a frequency that is representative of their environment. In the case of fast exchange, the spin that was in environment A is now found in environment B. Consequently its precessional frequency is now ω_B and the spin precesses counter-clockwise at ω_B for the next τ' period. At the end of this period the magnetization is found along the y-axis. Consequently no net precession has occurred and the *observed* resonance frequency is $\frac{1}{2}[\omega_A + \omega_B]$. In contrast, the non-exchanging spins ($k_{ex} = 0$) remain in their original environment and will continue to precess in the same direction for an additional period of τ'. Note that $\tau \neq \tau'$ because exchange is a random process, but the average time between exchange events, $\bar{\tau}$ and $\bar{\tau'}$ will be the same if $p_A = p_B$.

the spins do not exist in either environment long enough to establish an associated resonance frequency (See Fig. 8.2). Consequently, the spin precesses at a population averaged resonance frequency , ω_{obs}, and exhibits a population averaged spin-spin relaxation rate, T_2^{avg}:

$$\omega_{obs} = p_A\omega_A + p_B\omega_B \qquad \frac{1}{T_2^{avg}} = \frac{p_A}{T_2^A} + \frac{p_B}{T_2^B} \tag{8.2}$$

where p_A and p_B are the populations of each environment.

Slow Exchange ($k_{ex} << \Delta\nu$): Under conditions of slow exchange, the rate of exchange is slower than the frequency difference, in Hz, of the spin in each environment. Thus the exchange process is incapable of averaging the chemical shifts while the spins are precessing. Consequently, two resonance lines are observed, one line from the fraction of population of the spins that are found in conformation A and one line from the fraction that are found in conformation B. Since the environment of a spin determines its absorption frequency two separate resonance lines are observed, one at ω_a and one at ω_b. The integrated intensity of each line is equal to the fraction of spins in conformation A and B, respectively.

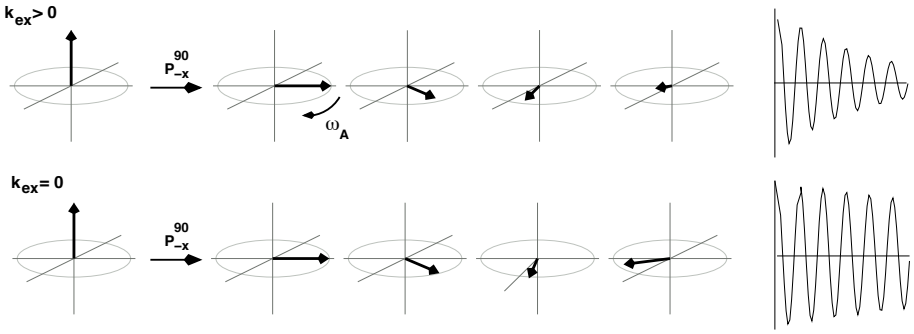


Figure 8.3. Effect of slow exchange on the relaxation rate of transverse magnetization. The top of the figure shows the bulk magnetization of spins in conformation 'A', undergoing slow exchange with conformation "B". The lower part of the figure shows the same spins in the absence of exchange. In both cases, the coordinate system is rotating at a frequency that is midway between ω_A and ω_B. The left-most section of the figure shows the bulk magnetization before the excitation pulse. The following panels show the spin precessing clockwise at a rate ω_A. The right-most section of the figure shows the observed free induction decay for the spins at ω_A. In the top section of the figure ($k_{ex} > 0$) the spins leave conformation 'A' and return from conformation 'B' a random time later. Since the precessional frequencies, ω_A and ω_B are different, the phase of the returning magnetization is no longer the same as those spins that remained in conformation A, causing randomization of the transverse magnetization. This reduces the signal, as illustrated by the shortening of the arrow that represents the x-y magnetization. This loss of magnetization increases the rate of decay of the transverse magnetization, shortening the observed T_2. In the absence of exchange, the transverse magnetization decays at the intrinsic T_2, as shown in the lower part of the figure.

Although slow exchange has no effect on the position of the resonance lines, the exchange process reduces the lifetime of the spin within a particular environment. Consequently, the resonance linewidth increases by an amount that is proportional to the rate of exchange. This line broadening can be understood by considering the effects of exchange on the magnetization as it precesses in the x-y plane after excitation (See Fig. 8.3). After excitation, the spins in environment A will precesses in the x-y plane at a frequency of ω_A. However, as time passes, some of these spins will change their environment and will begin precessing at ω_B. This change in environment results in a loss of *coherent* magnetization of the spins in conformation A, resulting in a decay of the transverse magnetization at a rate faster than the intrinsic spin-spin relaxation rate.

Since the system is at equilibrium, an equal number of spins also convert from conformation B to conformation A. This magnetization does not contribute to the magnetization originally associated with ω_A since the conversion from B to A occurs randomly. Consequently, the phase of the magnetization coming from environment B is incoherent with respect to the phase of the magnetization originally associated with environment A.

The rate at which the magnetization decays during the exchange process, $1/T_{2A}$, is given by the sum of the intrinsic spin-spin decay rate ($1/T_2^{oA} = R_2^{oA}$) and the rate at which conformation A is converted to conformation B (k_1). A similar expression can be written for the spins in environment B:

$$\frac{1}{T_2^A} = \frac{1}{T_2^{oA}} + k_1 = R_2^{oA} + k_1 \qquad \frac{1}{T_2^B} = \frac{1}{T_2^{oB}} + k_2 = R_2^{oB} + k_2 \quad (8.3)$$

The increase in T_2 leads to broadening of the resonance line by each rate constant:

$$\pi\Delta\nu_{1/2}^A = \frac{1}{T_2^{oA}} + k_1 \qquad \pi\Delta\nu_{1/2}^B = \frac{1}{T_2^{oB}} + k_2 \quad (8.4)$$

8.2.2 General Theory of Chemical Exchange

The effect of exchange on the NMR lineshape can be analyzed in a number of ways. The traditional method, attributed to H. M. McConnell [**?**], utilizes an analysis of the steady-state magnetization while the measurement frequency is changed (continuous wave spectroscopy, CW). This approach is quite straightforward and provides an simple expression for the lineshape that is suitable for direct fitting of experimental lineshapes by least-squares methods.

The final result of this analysis, under the assumption that the decay due to non-exchange processes is negligable, and that the populations of the two environments are equal, is:

$$I(\omega) \propto \frac{1}{2}k(\Delta\omega)^2 \frac{1}{(\omega_A - \omega)^2(\omega_B - \omega)^2 + 4k^2\omega^2} \quad (8.5)$$

where k is the rate constant for exchange ($k = k_1 = k_2$) and the origin of the frequency axis ($\omega = 0$) is mid-way between ω_A and ω_B.

8.2.2.1 Fast Exchange Limit

Under conditions of fast exchange, $k >> \Delta\omega$, the second term in the denominator of Eq. 8.2.3 will dominate, giving a single resonance line at $\omega = 0$, half-way between ω_A and ω_B ($\omega = p_A\omega_A + p_B\omega_B$, $p_A = p_B$ in this example). The linewidth can be obtained by investigating the behavior of the lineshape near $\omega = 0$. Here, $\omega_A - \omega$ can be approximated as ω_A, a similar approximation can be taken for ω_B.

$$\begin{aligned} I(\omega) &\propto \frac{1}{2}k(\Delta\omega)^2 \frac{1}{(\omega_A)^2(\omega_B)^2 + 4k^2\omega^2} \\ &\propto \frac{1}{2}k(\Delta\omega)^2 \frac{1}{(\Delta\omega/2)^2(\Delta\omega/2)^2 + 4k^2\omega^2} \\ &\propto \frac{1}{8k}(\Delta\omega)^2 \frac{1}{\frac{1}{k^2}(\Delta\omega)^4(\frac{1}{8})^2 + \omega^2} \end{aligned} \quad (8.6)$$

This gives a rate of decay equal to:

$$\frac{1}{T_2} = \frac{(\Delta\omega)^2}{8k} \tag{8.7}$$

In the general case, when $k_1 \neq k_2$ the observed rate of decay of transverse magnetization due to exchange is [**?**]:

$$R_2 = \frac{1}{T_2} = p_A p_B (\Delta\omega)^2 / k_{ex} \tag{8.8}$$

This equation is equivalent to Eq. 8.7, when the two rate constants are equal.

8.2.2.2 Slow Exchange Limit

Under conditions of slow exchange, $k << \Delta\omega$, we anticipate the presence of two resonance lines, one at ω_A and the other at ω_B, as predicted from the previous qualitative analysis in Section 8.2. If the behavior of the general lineshape function in the vicinity of ω_A is considered, then the lineshape function reduces to:

$$\begin{aligned} I(\omega) &\propto \frac{1}{2}k(\Delta\omega)^2 \frac{1}{(\omega_A-\omega)^2(\omega_B-\omega_A)^2 + 4k^2\omega_A^2} \\ &\propto \frac{1}{2}k(\Delta\omega)^2 \frac{1}{(\omega_A-\omega)^2(\Delta\omega)^2 + 4k^2(\Delta\omega/2)^2} \\ &\propto \frac{1}{2}k \frac{1}{(\omega_A-\omega)^2 + k^2} \end{aligned} \tag{8.9}$$

The above transformations assume that when $\omega \approx \omega_A$ both $(\omega_B - \omega)$ and ω^2 are slowly varying functions of ω, i.e. we can approximate $(\omega_B - \omega)$ as $(\omega_B - \omega_A)$, and ω^2 as ω_A^2. On the other hand, $(\omega_A - \omega)$, is a rapidly varying function when $\omega \approx \omega_A$ and must be treated exactly.

Recall that the lineshape for a free induction decay that decays with a rate of $\frac{1}{T_2}$ is:

$$\begin{aligned} I(\omega) &\propto \frac{T_2}{1 + T_2^2\omega^2} \\ &\propto \frac{1}{\omega^2 + [\frac{1}{T_2}]^2} \end{aligned} \tag{8.10}$$

Therefore, the contribution of exchange to the decay of the transverse magnetization is k, i.e.

$$R_2 = \frac{1}{T_2} = k \tag{8.11}$$

as predicted from the earlier qualitative analysis.

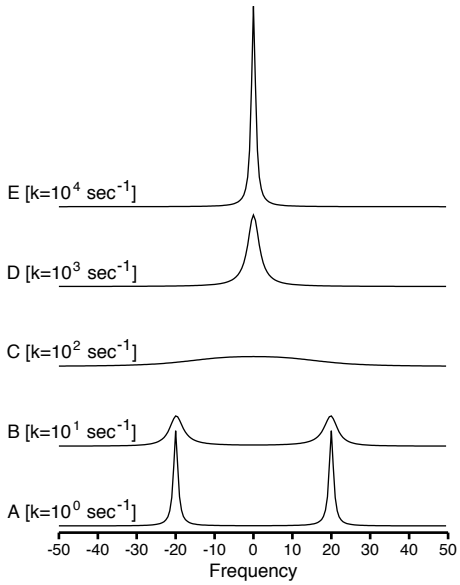


Figure 8.4. Effect of chemical exchange on lineshape. The lineshape is shown over a wide range of exchange rates. In this simulation, the population of both states were set to $\frac{1}{2}$. The frequency separation between ω_A and ω_B, $\Delta\omega$, was set to 40 Hz. The exchange rates vary from 1 $\sec^{-1}$ to 10^4 $\sec^{-1}$. Spectrum A corresponds to very slow, B to slow, C to intermediate, D to fast, and E to very fast exchange rates, as defined in Table 8.1.

8.2.3 Intermediate Time Scales

The lineshape for intermediate exchange rates is obtained by evaluation of Eq. . Numerical simulations are presented in Fig. 8.4. This simulation shows the predicted transition from two peaks in the case of slow exchange to a single peak under conditions of fast exchange.

Note that in the intermediate time scale, the resonance lines may be so extensively broadened such that no peak is observed in the actual spectrum. This situation will also hold true in a NOESY-type experiment if the dipolar coupled spins are exchanging between environments.

8.3 Ligand Binding Kinetics

The binding of a ligand, L, to a protein, P, generally causes a change in the environment of the nuclear spins on either the protein, the ligand, or both. Since a change in environment usually results in a change in chemical shift, ligand binding is equivalent to chemical exchange between two environments. In this case the two states correspond to the bound (PL) or free protein (P). The reaction for ligand binding to a single site is:

$$\mathbf{P} + \mathbf{L} \underset{\mathbf{k_{off}}}{\overset{\mathbf{k_{on}}}{\rightleftharpoons}} \mathbf{PL}$$

The equilibrium binding constant for this reaction is given by:

$$K_{eq} = \frac{k_{on}}{k_{off}} \tag{8.12}$$

where k_{on} is the second order rate constant for the binding of the ligand to the protein, and k_{off} represents the rate at which the ligand leaves the protein.

The kinetic rate equation in this situation is:

$$\begin{aligned} \frac{d[P]}{dt} &= -k_{on}[L][P] + k_{off}[PL] \\ \frac{d[PL]}{dt} &= +k_{on}[L][P] - k_{off}[PL] \end{aligned} \tag{8.13}$$

Note that the only difference between these equations and those given for general exchange in Eq. **??** is that the forward rate constant k_1, now becomes $k_{on}[L]$. Therefore, any of the previous described techniques can be applied to the study of ligand binding kinetics.

There are two key differences between general exchange and ligand binding:

1 The forward rate, k_1, can be conveniently changed by the ligand concentration, since $k_1 = k_{on}$[L]. Thus the time scale of exchange can be controlled, to some extent, by the varying the ligand concentration.

2 The chemical shifts of the unliganded and fully liganded states can be measured by acquiring spectra in the absence and presence of ligand. Consequently it is possible to obtain $\Delta\omega$, and in some cases a direct measurement of the equilibrium constant by acquiring a series of spectra at different ligand concentrations.

Figure 8.5 shows the effect of chemical exchange of the observed spectra. Three different scenarios are shown, slow (A), intermediate (B), and fast (C) exchange. Approaches to obtain the kinetic rate constants for each of these time scales are briefly discussed below.

8.3.1 Slow Exchange

In the case of slow exchange the resonance for the unliganded state simply disappears and a resonance line at the position of the liganded state appears. At any given ligand concentration the fraction of protein with bound ligand can be calculated from the integrated intensities of the two lines:

$$p_B = \frac{I_B}{I_A + I_B} \tag{8.14}$$

The dependence of p_B on the *free* ligand concentration follows the ligand binding equation:

$$p_B = \frac{K_{eq}[L]}{1 + K_{eq}} \tag{8.15}$$

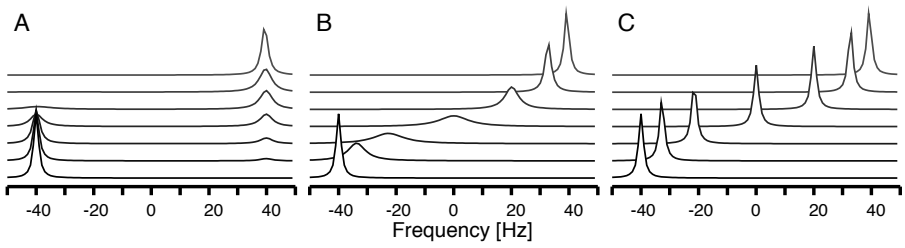


Figure 8.5. Effects of Ligand Binding on NMR Lineshapes. This figure illustrates the effect of slow (A), intermediate (B), or fast (C) exchange on the spectrum of a resonance whose frequency is changed as a result of ligand binding. All three panels correspond to an equilibrium binding constant, $K_{eq} = k_{on}/k_{off}$, of 10^4 M^{-1} but differ in the on- and off-rates. The on-rates were 10^5 $M^{-1}sec^{-1}$ (A),10^7 $M^{-1}sec^{-1}$ (B), and 10^9 $M^{-1}sec^{-1}$(C). The free and bound resonance positions are separated by 80 Hz ($\nu_A = -40$ Hz, $\nu_B = +40$ Hz). Spectra were simulated with ligand concentrations of 0, 10 μM, 30 μM, 100 μM, 300 μM, 1 mM, and 10 mM, increasing from front to back. These concentrations give the following fractions of liganded protein (p_B): 0.00, 0.09, 0.23, 0.50, 0.75, 0.91, and 0.99, respectively.

K_{eq} can be obtained by direct fitting of this function. Alternatively, it can be linearized with a Scatchard plot,

$$\frac{p_B}{[L]} = K_{eq} - p_B K_{eq} \tag{8.16}$$

If the binding is non-cooperative and to a single site, then the slope of this line is $-K_{eq}$.

The individual rate constants can be obtained by direct measurement of the linewidth at low ligand concentrations. At low ligand concentrations the lifetime of the unliganded state is shortened when the protein binds a ligand, therefore the observed spin-spin relaxation rate becomes:

$$R_2^A = R_2^{oA} + k_{on}[L] \tag{8.17}$$

The relaxation rate, R_2^{oA} in the absence of exchange is known, therefore a plot of R_2^A versus the *free* ligand concentration will have a slope of k_{on}.

8.3.2 Intermediate Exchange

In the intermediate exchange time regime the resonance line becomes very broad as ligand is added and slowly migrates from the unliganded position towards the position of the fully liganded protein. Under some conditions, the line can disappear completely during the titration, making it difficult to identify the resonance position of the fully bound species. Due to the complex nature of lineshape it is necessary to analyze these data using the complex expression for chemical exchange that was presented in Section **??**. The only modification to this equation is to replace k_1 with $k_{on}[L]$. Generally, several spectra are

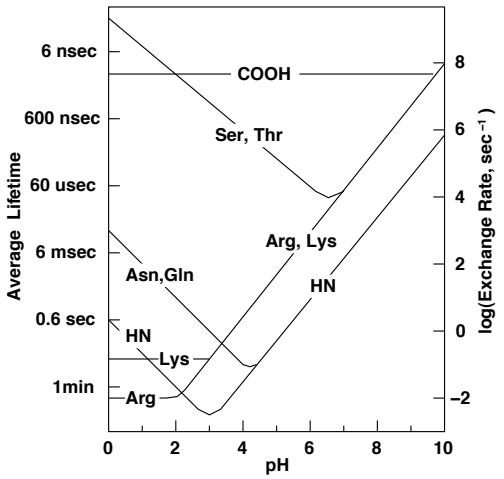


Figure 8.6. Effect of pH on hydrogen exchange rates. The hydrogen exchange rates are shown for several exchangeable groups in proteins. NH indicates the backbone amide proton. Other labels refer to sidechains. The rates assumes that the group is fully exposed to solvent. The left axis indicates the mean lifetime of the proton while the right scale gives the exchange rate.

acquired over a range of ligand concentrations and values of k_{on} and k_{off} are found that minimize the difference between the observed lineshape and that predicted from Eq. **??**.

8.3.3 Fast Exchange

In the case of fast exchange, the exchange is sufficiently fast that the chemical shift of the observed line, δ, is essentially equal to the weighted average of the initial and final states:

$$\delta = p_A\delta_A + p_B\delta_B \tag{8.18}$$

At any given ligand concentration, the amount of bound protein, p_B, is easily found:

$$p_B = \frac{\delta - \delta_A}{\delta_B - \delta_A} \tag{8.19}$$

The equilibrium binding constant can be obtained using the methods described above for slow exchange.

8.4 Measurement of Exchange: CPMG Experiments

8.4.0.1 Measuring k_{ex}: Relaxation Dispersion Experiments

Under conditions of fast exchange the apparent exchange rate constant, k_{ex}, can be obtained from relaxation dispersion experiments, or the change in the relaxation rate due to a change in the magnetic field strength. In a relaxation

dispersion experiment, the spins are subject to a transverse magnetic field, ω_T of varying strength while in the transverse (x-y) plane. The observed spin-spin relaxation rate depends on *both* the chemical exchange rate and the strength of the applied field. Because of this dependence it is possible to determine k_{ex} without relying on any other information. In practice, the spin-spin relaxation rate, R_2, is measured at a number of different transverse field strengths and the exchange rate is obtained by fitting the measured dispersion curve, $R_2(\omega_T)$, to theoretical models.

Two methods are employed to generate magnetic fields of different strengths for dispersion experiments. The first employs a series of 180° pulses, spaced τ_{cp} apart, as illustrated in Fig. 8.7. The initial implementation of this sequence by Carr and Purcell utilized the same phase for the 90° excitation and 180° refocusing pulses [**?**]. A later modification by Meiboom and Gill reduced the effects of imperfect 180° pulses by phase shifting the excitation pulse with respect to the 180° pulses [**?**]. In recognition of the individuals who developed this method, the sequence is usually referred to as a CPMG sequence. This

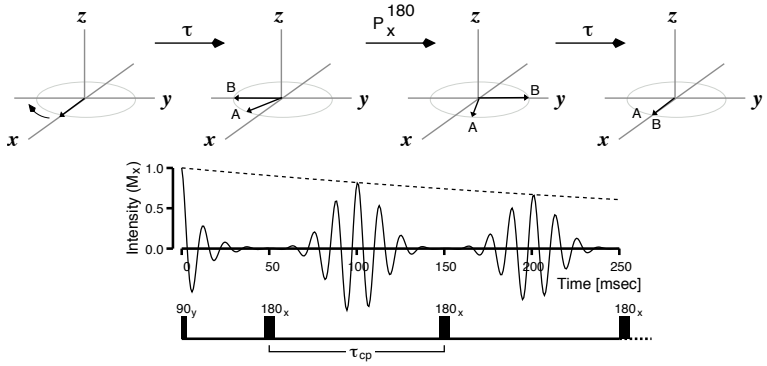


Figure 8.7. The generation of a spin-echo by a train of CPMG pulses. The top section shows the evolution of the transverse magnetization during the CPMG sequence. In this example the transverse magnetization dephases rapidly due to an inhomogeneous magnetic field. This dephasing is reversible and is refocused by a 180° pulse. The left-most section shows the magnetization immediately after the 90° pulse. During the first time period, τ, spins A and B precess at different rates due to differences in the the local magnetic field. In this example spin B precesses faster than spin A in the rotating frame. The 180° pulse rotates each spin about the x-axis. Spin A, which was lagging behind spin B, is placed ahead of spin B by the 180° pulse. In the next time period, the spins precess in the same direction as before, and at the same rate. After a period τ all of the spins will refocus to the x-axis because spins with slower precessional rates do not have to precess as far as those with higher rates. The effect of a train of such pulses, is shown in the lower part of the figure. A series of 180° pulses, spaced τ_{cp} apart, will generate a series of echos spaced τ_{cp} apart. The amplitude of which will decay according to the intrinsic spin-spin relaxation rate, R_2, as indicated by the dotted line. If chemical exchange occurs the decay rate will increase and may depend on τ_{cp}, depending on the size of k_{ex}.

pulse train refocuses the effects of magnetic field inhomogeneities on the decay of transverse magnetization and is therefore useful for measuring T_2. The 180° pulses in the CPMG sequence generate a spin-echo at a point half-way between each pulse. The amplitude of this echo decays according to the spin-spin relaxation rate, as illustrated in Fig. 8.7.

In the absence of exchange the $\tau - 180^o - \tau$ sequence will refocus any precession due to chemical shift evolution or magnetic field inhomogeneities since the precessional frequency for any given spin is constant over the duration of the experiment. However, in the case of exchange, the precessional frequency no longer is constant. Consequently, the height of the spin-echo will decrease with a rate equal to the sum of the spin-spin relaxation rate and the rate of exchange: $R_2 = R_2^o + R_{ex}$.

The second method of measuring relaxation dispersion is to apply a continuous B_1 field while the spins are transverse. Since this RF-field forces the magnetization to remain aligned with B_1, this method is referred to as *spin-locking*. This measurement of relaxation is referred to as T_1 relaxation in the rotating frame, or $T_{1\rho}$. In practice, $T_{1\rho}$ is measured at different field strengths by simply changing the strength of the B_1 field: $\omega_T = \omega_1 = \gamma B_1$. Larger spin-locking fields can be obtained using off-resonance B_1 fields.

Both CPMG methods and $T_{1\rho}$ techniques reduce the contribution of chemical exchange to the observed transverse relaxation rate. The degree of attenuation depends on the relationship between the rate of exchange and the applied field strength. A qualitative description of the effect of the transverse field strength on the contribution of chemical exchange to dephasing of the transverse magnetization is illustrated in Fig. 8.8. This figure shows the effect of increasing the transverse field strength at a fixed value of k_{ex}. In the case of CPMG methods the field strength is defined by the rate of 180° pulses, i.e. the more closely spaced pulses the higher the field. If the pulses are applied without any inter-pulse delay then a continuous B_1 is generated, which is equivalent to a spin-locking field. Thus field strengths during spin-locking are generally higher than the fields used in CPMG methods.

8.4.0.2 Quantitative Description of the CPMG Experiment

The evolution of the density matrix during the CPMG experiment has been analyzed using classical methods. The analysis considers free precession of the magnetization during the τ_{cp} between 180° pulse followed by a rotation about the x-axis of 180° by each pulse. The observed relaxation rate [**?**, **?**] is given in Eq. 8.20.

$$\begin{aligned} R_2 &= \frac{1}{2}\left[R_{2A} + R_{2B} + k_1 + k_2\right] - \frac{1}{\tau_{cp}} ln\lambda^+ \\ \lambda^+ &= ln\left[(D_+ cosh^2\xi - D_- cos^2\eta)^{1/2} + (D_+ sinh^2\xi - D_- sin^2\eta)^{1/2}\right] \\ &= (1/2) cosh^{-1}\left[D_+ cosh2\xi - D_- cos2\eta\right] \end{aligned} \tag{8.20}$$

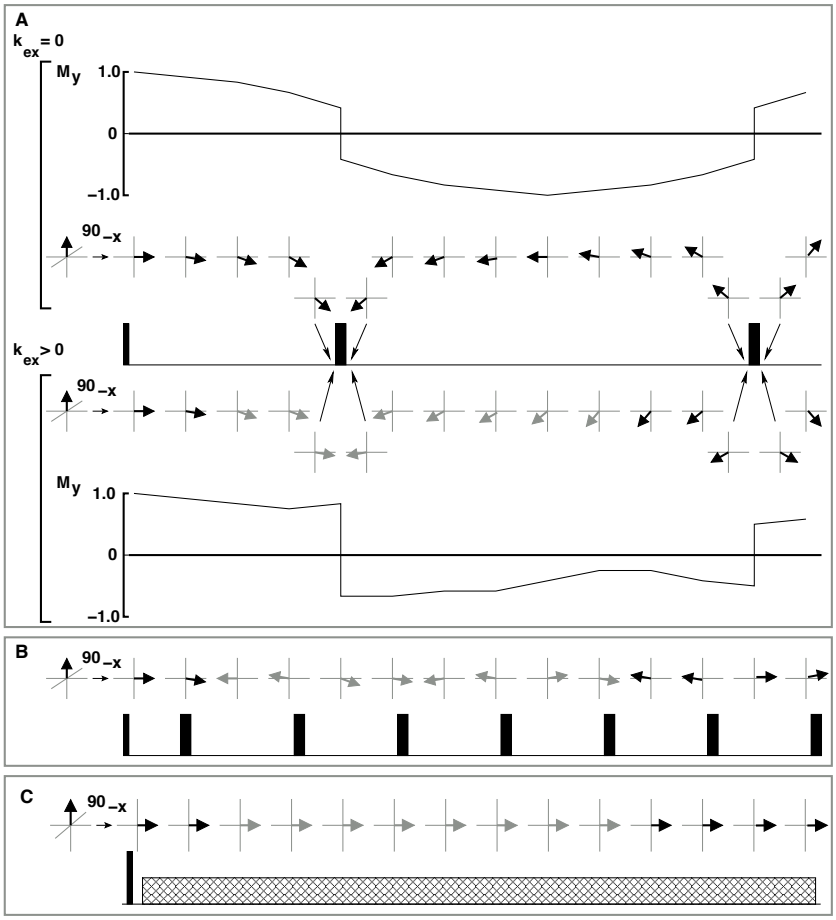


Figure 8.8. Measurement of exchange with CPMG or spin-lock methods. Panels A and B illustrate CPMG experiments while Panel C shows a spin-lock experiment. The evolution of one spin, as it exchanges between two environments, is shown by the black (environment A) or gray (environment B) arrow. In environment A the spin precesses clockwise in the x-y plane and while in environment B (gray arrow) the spin precesses counter-clockwise. The rate of spin-spin relaxation is assumed to be zero in both environments.

Panel A: The central portion of this panel shows the applied CPMG sequence. The upper part of the figure shows the evolution of a spin that does not undergo chemical exchange ($k_{ex} = 0$). The y-component of the magnetization is plotted along with the trajectory of the spin in the x-y plane. The magnetization that is precessing in the x-y plane is refocused by the first 180° pulse to give a maximum signal at τ_{cp}. In the presence of exchange ($k_{ex} > 0$), the spin begins in environment A and exchanges to B prior to the first 180° pulse. The change in precessional frequency interferes with the refocusing effect, causing an attenuation of the signal.

Panel B: The closer spacing of the 180° pulses makes it more likely that a spin will exist in a single environment during any τ_{cp} period, thus the refocusing is more effective at reducing attenuation of the signal due to exchange.

Panel C: A continuous on-resonance B_1 field along the y-axis is applied to the spins. In the rotating frame the field along the z-axis disappears, leaving only the B_1 field. The spins remain aligned, or *locked*, along B_1 in analogy to the alignment along B_o in the laboratory frame.

where,

$$D_{\pm} = \frac{1}{2}\left[\pm 1 + \frac{\psi + 2(\Delta\omega)^2}{\sqrt{\psi^2 + \zeta^2}}\right] \qquad \xi = \frac{\tau_{cp}}{\sqrt{8}}\left[+\psi + \sqrt{\psi^2 + \zeta^2}\right]^{1/2}$$

$$\eta = \frac{\tau_{cp}}{\sqrt{8}}\left[-\psi + \sqrt{\psi^2 + \zeta^2}\right]^{1/2} \qquad \zeta = 2\Delta\omega(R_{2A} - R_{2B} + k_1 - k_2)$$

$$\psi = (R_{2A} - R_{2B} + k_1 - k_2)^2 - (\Delta\omega)^2 + 4k_1k_2$$

Eq. 8.20 is valid for all exchange rates. Due to the complexity of the full expression it is helpful to gain an understanding of the response of the system by considering the simpler limiting case of fast exchange, which was derived by Luz and Meiboom [**?**] and by Allerhand and Thiele [**?**]:

$$R_2 = \frac{R_2^A + R_2^B}{2} + \frac{p_A p_B(\Delta\omega)^2}{k_{ex}}\left[1 - \frac{2}{k_{ex}\tau_{cp}}\tanh\frac{k_{ex}\tau_{cp}}{2}\right] \tag{8.21}$$

This simplification is remarkably robust in predicting the effect of exchange for experimentally accessible values of τ_{cp} (see Fig. 8.9).

The first part of the fast-exchange formula is just the averaged spin-spin relaxation rate that would be observed in the presence of very fast exchange. The second term gives the increase in the relaxation rate that is due to chemical

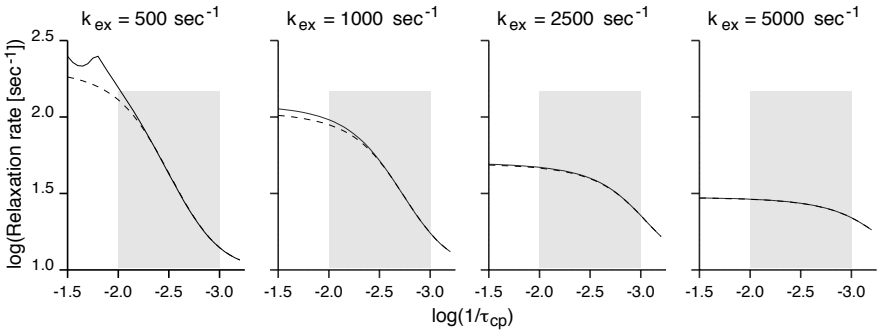


Figure 8.9. The dependence of the observed spin-spin relaxation time on the frequency of CPMG pulses. The relaxation rate, R_2^o, in the absence of exchange, is 10 sec^{-1}. This corresponds to the intersection of the x-axis with the y-axis. The frequency separation between the two states, $\Delta\omega/2\pi$ was 100 Hz and equal populations of both states were assumed ($p_A = p_B$). The solid line shows the relaxation rate calculated with the complete formula (Eq. 8.20), while the dotted line shows the result from the fast-exchange approximation, Eq. 8.21. The sum of the exchange rates ($k_{ex} = k_1 + k_2$), from left to right, are 500 sec^{-1}, 1000 sec^{-1}, 2500 sec^{-1}, and 5000 sec^{-1}. The shaded area corresponds to τ_{cp} values (10 msec, left side; 1 msec, right side) that are experimentally accessible for proteins in the $20 - 40$ kDa range. Note that the fast-exchange approximation offers a good estimation for the relaxation dispersion within the range of typical values of τ_{cp}.

exchange. When the exchange is very fast, $k_{ex} >> \tau_{cp}$, the observed increase in the relaxation rate is equal to that found for free precession in the presence of fast exchange (Eq. 8.8): $p_A p_B(\Delta\omega)^2/k_{ex}$. In contrast, when the rate of the 180° pulses becomes much faster than the exchange rate, the effect of chemical exchange on the linewidth disappears completely because:

$$\begin{aligned} lim_{\tau_{CP}\rightarrow 0}\left[1-\frac{2}{k_{ex}\tau_{cp}}tanh(\frac{k_{ex}\tau_{cp}}{2})\right]=0 \\ (lim_{x\rightarrow 0}\frac{1}{x}tanh(x)=1) \end{aligned} \tag{8.22}$$

A more complete description of the relationship between the CPMG pulse delay, the rate of chemical exchange, and the observed relaxation rate is shown in Fig. 8.9. In general, as the exchange rate increases, the dispersion, or change in R_2 versus τ_{cp} decreases. In the case of a relatively slow exchange rate, $k_{ex} = 500\ \mathrm{sec}^{-1}$, the R_2 due to exchange is large for long τ_{cp} values and is almost completely attenuated as τ_{cp} is decreased. In contrast, when the exchange rate is much faster than $\Delta\omega$ ($k_{ex} = 5000\ \mathrm{sec}^{-1}$, right panel) there is very little effect of τ_{cp} on the observed R_2. For very fast exchange, e.g. $k_{ex} = 100 \times \Delta\omega$ there is no dispersion and it is not possible to obtain information on the exchange rate with CPMG techniques.

The range of experimentally feasible τ_{cp} times is reduced by limitations on the pulse rate that can be generated by the instrument and by the intrinsic relaxation properties of the sample. Although the τ_{cp} times shown in Fig. 8.9 range from 31 msec to 0.5 msec, intervals much longer than 10 msec cannot be used with moderately sized proteins because of the relatively rapid spin-spin relaxation rate causes the transverse magnetization to decay before a single cycle of the CPMG sequence can be applied. Longer τ_{cp} values, that would extend the shaded area in Fig. 8.9 to the left, can be obtained for smaller proteins because of their long spin-spin relaxation rates (see [**?**]).

Values of τ_{cp} much shorter than 1 msec can cause sample heating problems. Furthermore, the analytical expressions shown above do not take into account finite pulse widths, thus the effect of τ_{cp} intervals that are of the same order as the 180° pulse lengths are not well represented by existing theories.

8.5 Amide Exchange

Resonance signals from the amide protons provide the cornerstone for practically all NMR experiments that are discussed in this text. Unfortunately, the amide proton readily exchanges with solvent. This exchange rate is catalyzed by both acid and base and the effect of pH on the rate of exchange is shown in Fig. 8.6. These rates correspond to groups that are fully exposed to the solvent. Rates that are several orders of magnitude slower can be observed for protons that are buried and not exposed to solvent. Since these buried amides eventually exchange with solvent, they must become transiently exposed to solvent as the protein samples multiple conformations.

If the rate at which the amide hydrogen exchanges with water is sufficiently fast, signals from the amide protons can disappear from the spectrum. For example, an amide proton resonance at 5.5 ppm is ≈500 Hz from the water resonance line on a 600 MHz (ν_H) spectrometer. If the rate of exchange is much faster than 500 sec^{-1}, the system will be in fast exchange and the observed chemical shift of the amide will be heavily weighted to that of the water due to the high concentration of solvent in the sample, i.e. the amide resonance would essentially appear at the solvent frequency.

Inspection of Fig. 8.6 shows that pH values less than approximately 6.0 are required to give an exchange rate that is slower than 500 sec^{-1} for fully exposed amides. At this pH the sidechain NH groups of arginine and lysine residues will be in fast exchange with the water resonance since their exchange rates are 100 times that of the NH proton. Consequently, their NH protons cannot be observed. The amide exchange rate of the amino terminus is similar to ϵ-amino group of lysine, therefore it is often difficult to observe resonance signals from this group as well. In a similar fashion, it would also be impossible to observe resonance signals from the hydroxyl group of serine and threonine at any pH, unless the group participates in hydrogen bonding or is buried in the protein, causing a reduction in the exchange rate.

8.5.1 Amide Exchange and Protein Dynamics

The exchange rate of a fully hydrogen bonded proton in an amide group is several orders of magnitude slower than the rate for a fully exposed, non-hydrogen bonded, amide. Yet, amide groups in helicies and sheets exchange more rapidly than would be expected for hydrogen bonded H_N protons. It is reasonably well established that the enhanced exchange rate is due to *local unfolding* of the protein. During this process, the amide proton is free to exchange with the solvent. Consequently, amide exchange rates can provide information on local unfolding in proteins.

Chapter 9

PROTEIN STRUCTURE DETERMINATION

Structure determination generates a molecular model of the protein or nucleic acid that is as consistent as possible with both the experimental data and known covalent and non-covalent features of the folded biopolymer. The most commonly used experimental constraints are:

- Inter-proton distances derived from NOESY experiments.
- Bond orientations determined from single bond residual dipolar couplings.
- Torsional angles from measurements of three bond J-couplings.
- Hydrogen bonds determined from amide exchange data.
- Peptide mainchain torsional angles from chemical shifts.

Non-experimental constraints consist of:

- Bond lengths.
- Bond angles.
- Torsional angles.
- Van der Waals Interactions.

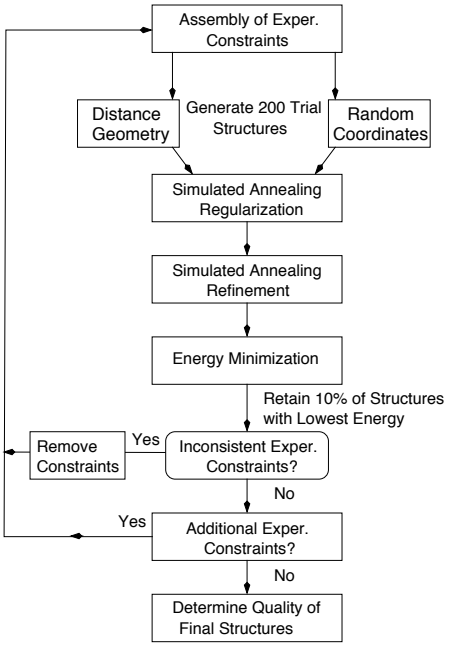


Figure 9.1. Overview of the structure determination process. Structure determination consists of a number of cycles that begin with the assembly of constraints, followed with building and refining models derived from the constraints, and then interpretation of the models to resolve errors or ambiguities in the experimental data.

The relative contribution of experimental and non-experimental constraints to the final structure are balanced by assigning an energy to both types of constraints and weighting the relative contribution of each type of constraint in an empirical manner:

$$E_{Total} = \kappa E_{Experimental} + E_{Non-experimental} \tag{9.1}$$

where κ is an empirical scaling factor.

Higher energies are associated with models whose structures gives the largest disagreement between the constraints and the structure. The refinement process seeks to create a structure that gives the lowest energy, and is therefore in best agreement with *both* experimental and non-experimental constraints.

The overall steps in structure determination are illustrated in Fig. 9.1 and outlined briefly in the following text. The first task is to assemble a collection of reliable structural constraints, these generally include inter-proton distances between protons that have unambiguous assignments. In addition, information on torsional angles from three-bond scalar (^{3}J) couplings can be utilized for the construction of the initial trial structures. Residual dipolar couplings, as well as hydrogen bonding information, are usually introduced during the latter stages of model building, when the structure is approaching its final form.

Following the assembly of constraints, 100 to 200 initial models are built for refinement. These can be generated from completely random atomic coordinates, or rough structures obtained from the NOE data via a technique called distance geometry, which converts interatomic distances to 3-D coordinates. The initial models generally show poor agreement with the experimental data and perhaps even with standard covalent and non-covalent interactions. Consequently, the models are "regularized" to produce structures that are constant with covalent geometry. Regularization is accomplished by moving the atoms to reduce the overall energy of the structure. Since some of the required changes in atomic coordinates may be large, this adjustment is usually performed using simulated annealing techniques that facilitate large changes in atomic coordinates. A number of the trial models may not converge to structures with acceptable energy and are discarded at this stage. Acceptable models are subject to additional refinement by simulated annealing to further decrease the energy of the system. The refinement is concluded with energy minimization, which performs small changes in atomic coordinates to maximize the agreement of the model with experimental data, as well as bonded and non-bonded interactions.

The refined models are ranked by energy and 5% to 10% of the lowest energy structures are selected. This ensemble of structures is inspected carefully to identify incorrect input data, such as incorrectly assigned NOEs, and such experimental constraints are removed from the data set. The ensemble of structures is also used to resolve ambiguities with the existing data, allowing the inclusion of more constraints in the next round of structure building. For example, in the case of inter-proton distances from NOE measurements, a proton

(A) with a well resolved chemical shift may show an NOE to another proton (B) whose chemical shift is degenerate with a third proton (C). In the absence of a structure it is not possible to determine if the inter-proton distance corresponds to A-B or A-C. However, the ensemble of low-energy structures may make the choice clear by comparing the distances predicted from the model structures. Intermediate structures may also be useful in identifying hydrogen bond acceptors.

After the addition of more constraints, the entire process is repeated until all experimental constraints have been exhausted. The overall quality of the ensemble of lowest energy structures is then evaluated for compliance with covalent and non-covalent energy terms as well as for agreement to the experimental data.

9.1 Energy Functions

9.1.1 Experimental Data

9.1.1.1 NOE Constraints

The intensity of a crosspeak in NOESY spectra is related to the distance between the interacting protons. However, there is considerable uncertainty associated with converting the NOE peak intensities to distances. Consequently, it is common practice to specify both a lower bound (d_{lower}) and an upper bound (d_{upper}), with the assumption that the true distance lies between the two bounds. The upper and lower bounds are generally determined from the signal-to-noise ratio in the spectrum. For example, if the uncertainty in peak intensity is ΔI, and the measured intensity is I, then $d_{lower} \propto 1/(I+\Delta I)^6$ and $d_{upper} \propto 1/(I-\Delta I)^6$. Alternatively, the lower bound is often specified as the van der Waals radii of the atom.

9.1.1.2 Residual Dipolar Coupling

The contribution of residual dipolar couplings (RDC) to the overall energy is:

$$E_{RDC} = \sum_{i=1}^{n_{RDC}} K_{RDC}(\Delta\nu_i^{Calc} - \Delta\nu_i^{Expt})^2 \tag{9.2}$$

where $\Delta\nu_i^{Calc}$ is the coupling calculated from the model, $\Delta\nu_i^{Expt}$ is the experimentally measured splitting, and the sum is over all observed dipolar couplings (n_{RDC}).

In order to compare calculated coupling to measured couplings it is necessary to know the orientation of the molecular coordinate system with respect to the magnetic field. In addition, the extent of alignment of the protein or nucleic acid (A_a, A_r) must also be known. If these are known then the expected dipolar coupling can be calculated from the molecular structure using the following equation:

$$\Delta\nu(\theta,\phi)^{Calc} = D_a\left[(3cos^2\theta - 1) + \frac{3}{2}Rsin^2\theta cos2\phi\right] \tag{9.3}$$

where θ and ϕ represent the orientation of a particular bond in the molecular coordinate system.

9.1.1.3 Torsional Angles

During the course of refinement, the expected three-bond J coupling, ${}^3J^{Calc}$ can be calculated from a torsional angle in the structure. Most 3-bond coupling constants, such as the coupling between the amide proton and the α-proton, are related to the torsional angle by the Karplus relationship [**?**]:

$$ {}^3J = A\cos^2\theta + B\cos\theta + C \tag{9.4}$$

The deviation of the calculated J-coupling constant from the measured constant contributes to the overall energy function as:

$$ E_J = \sum_{i=1}^{n_J} K_J (J_i^{Calc} - J_i^{Expt})^2 \tag{9.5}$$

where the sum is over all observed J-couplings, K_J is an empirical weighting factor, J_i^{Calc} is the coupling constant calculated from the torsional angle in the model (using Eq. 9.7), and J_i^{Expt} is the experimentally observed coupling constant.

In earlier versions of refinement software, the experimentally measured torsional angle was explicitly given, in which case the energy function is:

$$ E_{tor} = \sum_{i=1}^{n_{tor}} K_{tor} (\theta_i^{Calc} - \theta_I^{Expt})^2 \tag{9.6}$$

In this case, the torsional angle that most likely corresponds to the observed coupling constant would be specified by the user. For example, a $J_{H_N H\alpha}$ coupling of 4 Hz implies a torsional angle, ϕ of -60° (see Fig. 9.2).

Common Coupling Constants Used in Refinement: The J coupling between the H_N to the Hα proton is perhaps the most useful coupling constant because it provides information on the peptide backbone configuration (ϕ angle). The relationship between the observed coupling constant and the peptide ϕ angle is given by the following parametrized Karplus relationship:

$$ {}^3J_{H_N H_\alpha} = 6.51\cos^2\theta - 1.76\cos\theta + 1.60 \tag{9.7}$$

where $\theta = \phi - 60°$, and the constants A, B, and C have been substituted with values determined by Vuister and Bax [**?**]. A number of different values for these constants have been obtained by other investigators [**?**] which give similar values for ${}^3J_{H_N H\alpha}$. The Karplus curve for ${}^3J_{H_N H_\alpha}$ is shown in Fig. 9.2.

Note that a single value of the ^{3}J-coupling constant can correspond to as many as four distinct values of ϕ, thus a measured coupling constant may not specify a unique value for ϕ. Furthermore, the observed coupling constant may be averaged by rotation about the $N-C_\alpha$ bond at a rate that is faster than $1/J$.

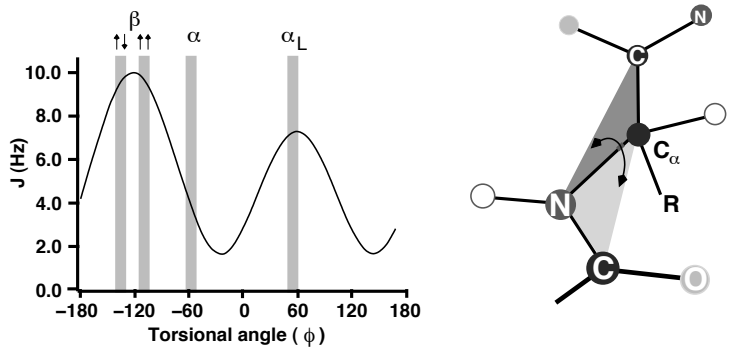


Figure 9.2. Relationship between secondary structure and $J_{N_H H\alpha}$. The observed ^{3}J-coupling versus the torsional angle, ϕ is shown. ϕ angles corresponding to common secondary structures are indicated by the vertical bars. The angle θ in Eq. 9.7 equals $\phi - 60°$. The ϕ torsional angle is the angle between the planes defined by C-N-C_α and N-C_α-C, as illustrated on the structure shown in the right-hand section of this figure.

For example, if an amino acid residue samples an α-helical configuration and a β-strand configuration with equal probability, then the observed coupling will be approximately 7 Hz, or the average of 10.0 Hz(β) and 4.0 Hz (α). Therefore, J-coupling constants that are approximately 7.0 Hz are not used in the initial model building because it is unclear whether it represents a single conformation with $\phi = -70°$ or is a result of conformational averaging. Once the conformation of the residue becomes established during refinement it may be possible to utilize these J-coupling constants as valid constraints for model building.

9.1.1.4 Hydrogen Bonding

Amide protons that participate in hydrogen bonds are usually identified by virtue of slow amide hydrogen exchange rates as well as a small temperature dependence of the amide proton chemical shift (see Fig. 9.3).

The exchange rates are readily measured by replacing the solvent with D_2O and measuring the decrease in the intensity of amide proton resonances. The exchange rate follows first-order kinetics and the rate constant is obtained by fitting the peak intensity, $I(t)$, to the following equation:

$$I(t) = I_o e^{-k_{ex} t} \tag{9.8}$$

The amide exchange rates are both acid and base catalyzed, and pH dependence of the exchange rate, for a number of exchangeable protons, is presented in Fig. 8.6. Generally, if the observed exchange rate is 10 fold slower than the rate expected for a fully exposed amide it is reasonable to assume that the amide is *either* involved in a hydrogen bond or it is simply buried in a hydrophobic region of the protein. The participation of the amide proton in an hydrogen bond can be further substantiated by the temperature dependence of the amide

proton chemical shift. Temperature coefficients that are less negative than $\approx$ 4.5 ppb/K are indicative of hydrogen bond formation [**?**], as illustrated in Fig. 9.3.

If a suitable hydrogen bond acceptor can be identified in preliminary models, for example a C=O group, then constraints involving the N-H and C=O atoms can be added to collection of constraints in the next round of structure generation.

In some cases it may be possible to unambiguously determine the presence of a hydrogen bond by detecting scalar coupling between the amide proton and the carbonyl carbon that are involved in the hydrogen bond. Since this coupling is quite weak, leading to long magnetization transfer times, experiments of this type are generally more successful with smaller proteins (e.g. less than 20 kDa) because of their longer T_2 values.

In terms of energy calculations, hydrogen bonding constraints are often represented as a pair of inter-proton distances, i.e. the distance between the amide proton and carbonyl oxygen (d_{H-O}) and the distance between the nitrogen and the carbonyl carbon (d_{NC_O}) to insure linearity of the hydrogen bond, giving the following energy function:

$$E_{h-bonds} = \sum_{i=1}^{n_{hbonds}} K_{H-bond} \left[\sum_{j=1}^{2} (d_{ij}^{Calc} - d_{ij}^{Expt})^2 \right] \tag{9.9}$$

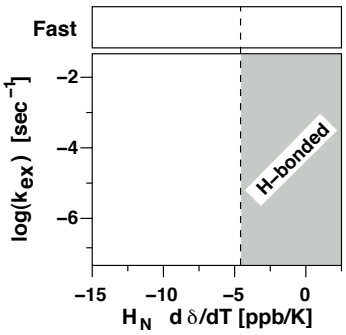


Figure 9.3. Identification of hydrogen bonds. The amide exchange rate versus the temperature dependence of the amide proton chemical shift. The range of each axis is typical for a folded globular protein at pH 5.0. An exchange rate of "fast" indicates that it was not possible to measure the rate due to rapid loss of the amide proton in D_2O. Amides that show an amide exchange rate less than 10^{-2}/sec and temperature dependence smaller than -4.5 ppb/K are likely to participate in the formation of a hydrogen bond. Amide protons that show a slow exchange rate but have a large temperature dependence of chemical shift may not be hydrogen bonded. Fig. adapted from [**?**].

where $j = 1$ represents the H-O distance and $j = 2$ represents the N-C distance and K_{H-bond} is an empirical scaling factor. K_{H-bond} is often set to the same value of K_{NOE} during refinement.

9.1.1.5 Chemical Shift Constraints

The chemical shift of backbone atoms, in particular the C_α, and carbonyl carbon, and to a lesser extent C_β, depend on the secondary structure of a residue, as indicated in Fig.9.4 [**?**]. Although many tertiary interactions also affect the chemical shift of these atoms, the change in chemical shift that is induced by the secondary structure can provide a weak constraint during refinement (see [**?**]). For example, if an alanine residue in a protein showed C_α and CO shifts that were 3 ppm below the mean chemical shift for these atoms, then the ϕ and ψ torsional angles could be constrained to favor a β-strand conformation.

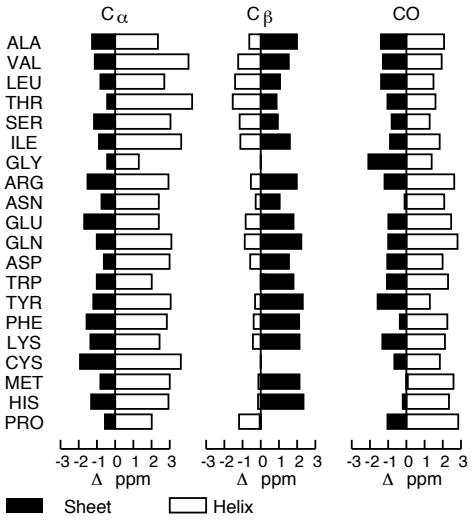


Figure 9.4. Effect of secondary structure on carbon chemical shifts. The deviation from the mean chemical shift of the C_α, C_β, and CO atoms for each residue type is illustrated by each horizontal bar. Filled bars represent residues in β-strand configuration while open bars represent residues in α-helical configuration. As an example, the chemical shift for the C_α of Ala is 52.42 ppm in a random coil, 51.15 ppm in a β-sheet, and 54.77 ppm in a helix. Therefore the change in chemical shift due to secondary structure is -1.27 ppm and 2.35 ppm from random coil for helix or sheet, respectively. Data from [**?**].

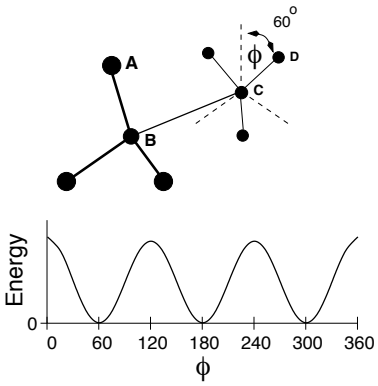


Figure 9.5. Torsional angle potential energy. A molecular fragment from a protein is shown in the upper section of the diagram. The bond connecting atoms B and C could, for example, correspond to the C_α-C_β bond in alanine. The torsional angle, ϕ is specified by the angle between the planes defined by atoms ABC and atoms BCD. In this particular configuration, $\phi = 60^\circ$. The lower portion of the figure shows how the energy varies with dihedral angle, here $n = 3$ and $\delta = 0$ in Eq. 9.12.

9.1.2 Covalent and Non-covalent Interactions

The energy associated with covalent and non-covalent interactions are defined by the following terms:

$$\begin{aligned} E_{covalent} &= E_{bonds} + E_{angles} + E_{torsional} + E_{improper} \\ E_{non-bonded} &= E_{van\ der\ Waals} + E_{Electrostatic} \end{aligned} \tag{9.10}$$

The covalent energy terms insure proper covalent bonding and molecular structure, including planarity of aromatic groups and the correct geometry of chiral centers.

Bond Lengths: Proper inter-atomic bond lengths are maintained during refinement with the following energy term:

$$E_{bonds} = \sum_{i}^{n_{bonds}} K_{bonds}(d_i^{Calc} - d_i^{Ideal})^2 \tag{9.11}$$

where n_{bonds} are the number of covalent bonds in the structure, K_{bonds} is an empirical scale factor, d_i^{Calc} is the bond length calculated from the structure during refinement and d_i^{Ideal} is the ideal covalent bond length.

Torsional Angles: Torsional angles can also contribute to the energy function. They are used to maintain the planarity of aromatic rings, favor non-eclipsed configuration of atoms, and to define the geometry of chiral centers. In the

case of X-PLOR, two energy functions can be used to specify torsional angles, the first of which is usually used to insure non-eclipsed atoms, as illustrated in Fig. 9.5:

$$\begin{aligned} E_{torsional} &= \sum K_\phi \left[1 + cos(n\phi + \delta)\right] \quad if \quad n > 0 \\ &= \sum K_\phi (\phi - \delta)^2 \quad if \quad n = 0 \end{aligned} \tag{9.12}$$

where ϕ is the torsional angle calculated from the structure at some point during refinement, K_ϕ is the weighting factor, n is the multiplicity, and δ is the phase shift.

To specify that four atoms lie in a plane, such as in an aromatic ring, n would be set to zero and δ would be 180°. Any out of plane configurations would raise the energy due to a non-zero value of $(\phi - \delta)$. To specify three equally populated rotomers, for example a CH_3 group, a value of $n = 3$ and $\delta = 0°$ would be used, giving energy minima for torsional angles, $\phi = 60°, 180°$, and 300°, as illustrated in Fig. 9.5.

The second energy expression for torsionals angles is referred to as the *improper* energy term. It has exactly the same form as in Eq. 9.12 and is generally used to maintain chirality and planarity of groups within in the structure. The availability of two distinct potential functions permits a use of different scale factors for the two types of torsional angles.

van der Waals Interactions: The non-bonded energy term contains contributions from van der Waals interactions, usually encodes a pairwise standard 6-12 Lennard-Jones potential:

$$E_{vdw} = K_{vdw} \sum_{ij}^{n_{atom}} \frac{C_{12}}{d_{ij}^{12}} - \frac{C_6}{d_{ij}^6} \tag{9.13}$$

Electrostatic Energy: The simplest form of the electrostatic energy term is given by Coulomb's Law. However, given the uncertainty of the local dielectric constant as well as the absence of solvent and counter ions in most structure refinement protocols, this term is usually set to zero.

9.2 Energy Minimization and Simulated Annealing

Structures are refined by a combination of energy minimization and simulated annealing. The overall goal is to alter the atomic coordinates of the structure to attain a final set of atomic coordinates that give the lowest energy for both experimental and non-experimental energy functions.

9.2.1 Energy Minimization

The minimum energy of the structure can be found by moving the atoms in the direction defined by the gradient of the energy:

$$\xi_i = -\frac{\partial E}{\partial x_i} \tag{9.14}$$

During the energy minimization, multiple steps of adjustment of the atomic coordinates occur. The coordinate change at each step is calculated according to the following:

$$x_i' = x_i + \xi_i \tag{9.15}$$

ξ_i is recalculated after each step in the minimization process and becomes smaller and smaller as the system moves towards the minimum in energy. The minimization proceeds for either a set number of cycles or until ξ_i drops below a predetermined level.

If the energy function is smooth and has a single global minimum, then minimization will find the true global minimum and produce a structure that is as consistent as possible with the energy function. Unfortunately, the energy surface as a function of atomic coordinates is complex and multi-valued such that a simple minimization of the energy will inevitably reach a local minimum, not the true global minimum, as illustrated in Fig. 9.6.

9.2.2 Simulated Annealing

Simulated annealing is used to overcome the problem of the structure becoming trapped in a local energy minimum. This procedure receives its name because it simulates the annealing process in alloy formation in metals. Specifically, the metal is heated to high temperatures to facilitate atomic rearrangements and then cooled or annealed to more stable structures.

In the refinement of models the atoms in the protein are given a kinetic energy, as defined by the temperature of the system:

$$\frac{1}{2}\sum_{i=1}^{N} m_i v_i^2 = \frac{3}{2} N k_b T \tag{9.16}$$

Initially, the atoms are assigned a random velocity that depends on the temperature of the system [**?**]:

$$v = \left[\frac{m}{2\pi k_b T}\right]^{3/2} e^{-3m\delta^2/2k_b T} \tag{9.17}$$

where δ is a random number from 0 to 1, T is the temperature of the system, k_b is Boltzmann's constant, and m is the mass of the atom.

Since simulating annealing generally begins at high temperatures, the atoms will have high kinetic energy, and will be able to transverse the energy barrier between minima, as illustrated in Fig. 9.6. To insure the the system will

converge on a minimum, the temperature of the system is slowly lowered at the end of the molecular mechanics calculation. Provided the temperature is lowered slowly, it is very likely that the system will anneal to a global energy minimum.

The motion, or trajectory, of the atoms during simulating annealing are determined by molecular mechanics calculations. Given a set of initial coordinates and velocities, x_o and v_o, as well as the energy of the system, the coordinates at a time Δt are calculated using Newtonian mechanics using the following expression:

$$x' = x_o + v_o \Delta t - \nabla E \frac{\Delta t^2}{2m} \tag{9.18}$$

where $-\nabla E$ is equal to the force applied to the atoms. E represents all, or a subset, of the experimental, covalent, and non-covalent energy terms discussed above. Generally, the time step, Δt is a fraction of a psec and 50-200 steps are performed at any given temperature. Multiple cycles of molecular mechanics are usually performed with any given refinement protocol. Each cycle will generally use different energy scaling factors as well as a number of other

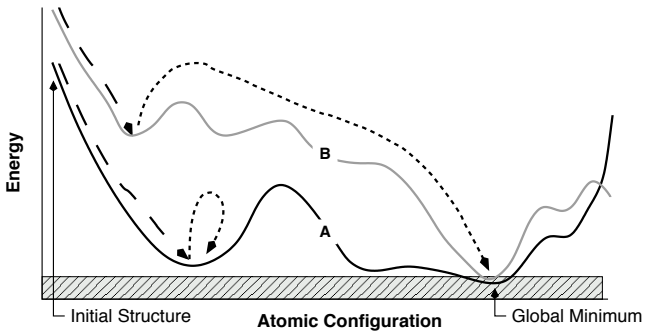


Figure 9.6. Energy changes during simulated annealing. The solid lines shown the energy of a structure as a function of its atomic coordinates. Two trial structures are shown, drawn in black (structure A) or gray (structure B). The energy of the structures immediately after generation by distance geometry or starting from random coordinates are shown on the far left of the plot. The dashed lines shows the change in energy due to regularization followed by energy minimization. Both structures reach a local minimum with reasonable covalent geometry after regularization. The dotted lines show the changes in energy that occur during additional refinement by simulated annealing. In the case of structure B (gray), the energy barriers between each local minimum can be transversed due to the high kinetic energy of the atoms during annealing, thus B eventually finds the global minimum in energy. The energy barriers surrounding the local minimum for structure A are too high, thus structure A is found at the local minimum after refinement. The gray stippled region indicates the range of energies that are considered to be acceptable after refinement; only a very small subset of all possible atomic configurations possess the indicated range of energies.

parameters, such as the van der Waals radii of atoms. Several examples of refinement protocols are presented below.

9.3 Illustrative Example of Protein Structure Determination

The following section outlines the process of structure determination for a 130 residue protein, rho130 [**?**]. This protein was chosen as an example because it consists of two sub-domains, an amino-terminal domain that is largely α-helical, and a carboxy-terminal domain that is largely β-sheet (see Fig. 9.7, Panel D). The number of contacts between the two domains is very limited, thus it is difficult to obtain a large number of inter-proton distances to constrain the structure of one domain with respect to the other. In this case the importance of including constraints from residual dipolar couplings is quite apparent, giving a precise definition of the relative orientation of the two domains with respect to each other.

Four major steps (A-D) in the process of protein structure determination are depicted here. The constraints that were used at each step are listed in Table 9.1 and the resultant structures are shown in Fig. 9.7. At the beginning of each step, 200 trial structures were built using distance geometry. Four of the lowest energy models that were obtained after the final refinement protocol are shown in Fig. 9.7.

The initial models (A) were constructed from approximately 900 inter-proton distances, which were obtained from NOESY crosspeaks involving unambiguously assigned resonance. The amide-amide distances were obtained from a ^{15}N separated three-dimensional NOESY spectrum. The amide-aliphatic distances were also obtained from this experiment. Since the aliphatic region of the proton spectrum is quite crowded the assignment of the aliphatic peaks was aided by the CN-NOESY experiment, which gives the carbon shifts of the aliphatic protons that are close to the amide protons. Distances between aliphatic protons were obtained using a ^{13}C separated NOESY experiment, similar to the ^{15}N NOESY, except that nitrogen excitation was replaced by excitation of aliphatic carbons. Inter-proton distances between aromatic protons were largely obtained from a two-dimension proton-proton NOESY acquired in D_2O, however a small number of such constraints were obtained by acquiring a three-dimensional ^{13}C separated NOESY with the carbon transmitter placed on the aromatic carbon region.

Hydrogen bond constraints could not be utilized during stage A of structure determination. Although potential hydrogen bond donors had been identified by virtue of slow amide exchange kinetics, the identification of the acceptors required initial models. A similar situation also existed for the use of ϕ torsional angle constraints. Although most of the coupling constants had been measured for the three bond H_N-H_α coupling, only approximately one-third (30/91) were above 9 Hz and could be used without concern of conformational averaging of the coupling constant.

Residual dipolar couplings for the H-N and H-C_α bond vectors were obtained from samples aligned using filamentous bacteriophage. Both couplings were obtained by measuring the oscillation of peak intensity as a function of time and then fitting the data to a damped cosine function to determine the value of the coupling constant. These constraints were not used for refinement until the final stage, when reasonably accurate structures became available.

The lowest energy model structures that were obtained from step A showed good agreement within each sub-domain of the protein. The structural similarity can be characterized by the root-mean-squared deviation RMSD. The

			A	B	C	D
I	NOE	H_N-H_N				
		Local	147	167	184	184
		Long	38	42	53	53
		H_N-H_C				
		Intra	238	238	308	308
		Local	245	247	379	379
		Long	53	55	137	137
		H_C-H_C				
		Intra	163	163	159	159
		Local	14	14	52	52
		Long	39	39	118	118
	NOE:Long/residue		1.0	1.0	2.4	2.4
	H-bond		0	36	43	43
	ϕ		30	80	91	91
	χ_1		0	0	23	23
	RDC	HN	0	0	0	63
		HC_α	0	0	0	49
II		RMSD (β)	1.3	0.90	0.70	0.39
		RMSD (α & β)	7.0	6.75	1.40	0.43

Table 9.1. Constraints used in determining structure of Rho130. The constraints used at each stage in structure determination (A through D, see Fig. 9.7) of structure determination are given in section I of this table. The NOE constraints are divided into amide-amide (H_N-H_N), amide-aliphatic (H_N-H_C), and aliphatic-aliphatic or aliphatic-aromatic or aromatic-aromatic (H_N-H_C). Each of these categories is further divided into intra-residue, local, and long-range distances. A local distance constraint involves residues that are within four residues of each other in the primary sequence. Long range distances involve residues that are more than four residues from each other. The average number of long range distance constraints/residue are also given. The number of H-bonds are also listed. The number of torsional constraints constraining the N-C_α bond (ϕ) and the C_α-C_β bond (χ_1) are also listed. Finally, the number of HN and HC_α residual dipolar couplings are listed. Section II of this table gives the root-mean-squared-difference (RMSD) between each of the four low-energy structures. The first entry is the RMSD for alignment of the β-domain while the second entry is for aligning both domains.

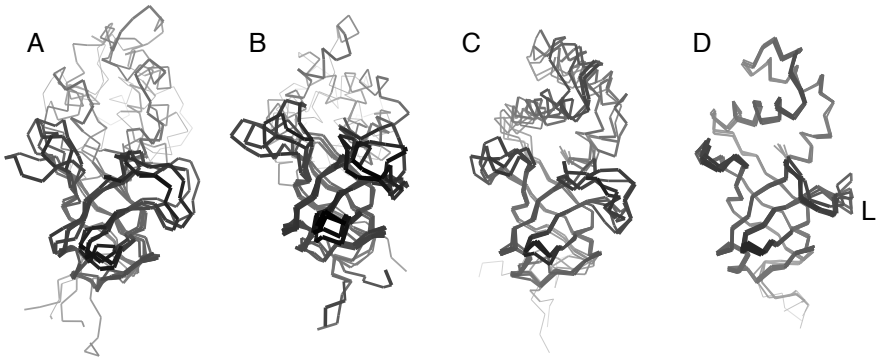


Figure 9.7. Stages in the structure determination of Rho130. Each panel shows the superposition of the four lowest energy structures after final energy minimization. The β-strand domain, which forms the lower part of the structure in this presentation, was used to define the alignment. Panel A shows the set of structural models generated from the initial set of constraints. Panel B and C show improvement in the structural models by the addition of hydrogen bonds as well as torsional angle and additional distance constraints. Panel D shows the effect of utilizing residual dipolar couplings as additional constraints. Note that the relative alignment of the two sub-domains is not well defined until Panel C and becomes more precise with the addition of the RDCs.

RMSD is a measure of the similarity of one structure to another and is defined as follows:

$$RMSD = \frac{1}{N} \sum_{i=1}^{N\ Atoms} \sqrt{[(x_{ij} - x_{ik})^2 + (y_{ij} - y_{ik})^2 + (z_{ij} - z_{ik})^2]} \qquad (9.19)$$

where x_{jk} represents the x-coordinate of the k^{th} atom in molecule j and x_{ik} is the x-coordinate of the same atom in molecule i.

The smaller the RMSD, the more similar the structures are to each other. In the case of the sub-domains, the RMSD is about 1.3 (Table 9.1). However, the relative orientation of the two domains was poorly determined, as indicated by the high RMSD for aligning the entire protein and the obvious poor alignment of structures in Panel A of Fig. 9.7.

The set of lowest energy structures from step A were inspected and a small number of inter-proton distances were added after using these structures to resolve ambiguities in the assignment of NOE crosspeaks. In addition, a total of 36 hydrogen bond acceptors were identified on the basis that the residues were in regular secondary structure. Finally, most of the ϕ torsional constraints could be used for the next stage since it was clear that the residues were in regular secondary structure. The refined structures from step B are shown in Fig. 9.7. There is a significant improvement in the sub-domain structure; the RMSD for the β region dropped from 1.3 to 0.9 . In addition, there is a modest decrease in the RMSD for overall alignment.

The structures from step B were sufficiently well defined to allow the assignment of a large number long-range distance constraints: 11 amide-amide, 82 amide-aliphatic, and 79 aliphatic-aliphatic. In addition, it was possible to identify 7 more hydrogen bonds and utilize 11 more ϕ torsional constraints. A number of torsional angle constraints, involving the C_α-C_β (χ_1) bond, could also be incorporated at this time.

The resultant structures are shown in Panel C of Fig. 9.7. The lowest energy structures were very well defined within each sub-domain, with an RMSD of 0.7 . The alignment of the overall protein was poorer showing an RMSD of 1.4 . However, the overall structure was quite acceptable. The increase in the quality of the structures in step C is due almost entirely to the increase in the number of long-range inter-proton distances, from 1.0/residue in B to 2.4/residue in stage C.

The difficulty in determining the relative orientation of the two sub-domains is not surprising given the fact that there are few distance constraints between the domains and that the information from these NOE derived distances is entirely of a local nature. The inclusion of constraints from residual dipolar coupling, as shown in D, provided information on the independent alignment of each sub-domain with respect to the direction of the applied magnetic field. A small number of such constraints are sufficient to fix the relative orientation of each sub-domain, as shown in panel D of Fig. 9.7.

The inclusion of residual dipolar couplings also increase the precision of the local geometry. The conformation of an inter-strand loop region of the protein, marked with an "L" in Panel D of Fig, 9.7, is poorly defined in C because of the lack of experimentally measured inter-proton distances in this region of the protein. However, the inclusion of several RDCs from this region causes the structures to converge to a common configuration. The RDCs also increase the overall precision of the structure, reducing the RMSD to 0.42 .

In summary, with a sufficient number of experimental constraints, it is possible to obtain structures of proteins in solution that rival structures obtained from high-resolution X-ray crystallography.