

Solid State NMR Investigation of Protein Based Biomaterials

**Pacific Hagfish Slime Thread and Recombinant Insect
Resilin**

by

Thomas A. Depew

B.A.Sc., Queen's University, 2004

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Master of Applied Science

in

The Faculty of Graduate Studies

(Physics)

The University Of British Columbia

May, 2008

© Thomas A. Depew 2008

Abstract

Nuclear Magnetic Resonance (NMR) was employed to investigate the structure and mechanics underlying the material properties of two remarkable biomaterials. Hydrated hagfish intermediate filament (IF) proteins were identified as having a two component nature, consistent with current structural models. One component is relatively rigid and immobile, the other rubbery, in which the protein backbone reorients with correlation times on the order of 60 ns. In order to investigate the role of calcium ions in the formation of hagfish slime, hagfish IFs were exposed to Ca^{2+} ions in solution. Energy dispersive X-ray spectroscopy revealed that the filaments did bind Ca ions after exposure. These results were variable and depended largely on the preparation technique.

Recombinant resilins from *Drosophila melanogaster* and *Anopheles gambiae* were shown to have a highly elastic structure. Direct polarization spectra from each protein were analysed and the majority of the ^{13}C peaks assigned successfully. Relaxation measurements report backbone correlation times on a scale of 2 to 8 ns, providing a molecular scale explanation the outstanding macroscopic resilience. Tyrosine residues in the resilin protein exhibited longer correlation times in the aromatic carbons, reflecting decreased mobility near dityrosine crosslinks.

Contents

Abstract	ii
Contents	iii
List of Tables	v
List of Figures	vi
Acknowledgements	x
1 Introduction	1
2 Nuclear Magnetic Resonance	2
2.1 Zeeman Interaction	2
2.2 Fundamentals of the NMR Experiment	4
2.3 Relaxation of Magnetization	5
2.4 NMR Spectroscopy	6
2.5 Density Matrix Representation	7
2.6 Theoretical Tools	8
2.6.1 Frame Transformations	8
2.6.2 Average Hamiltonian Theory	9
2.7 NMR Interactions	10
2.7.1 Chemical Shielding	11
2.7.2 Dipolar Coupling	14
2.7.3 Quadrupolar Interaction	17
2.8 Relaxation Theory	18
2.8.1 Correlation Function and Spectral Density	18
2.8.2 Dipolar Coupling	22
2.8.3 Chemical Shift Anisotropy	23
3 Experimental Techniques	24
3.1 Heteronuclear Decoupling	24
3.2 Magic-Angle Spinning	24
3.3 Phase Cycling	26
3.4 Cross-Polarization	26
3.5 Energy Dispersive X-ray Spectroscopy	30
4 Introduction to Proteins	30

4.1	Protein Secondary Structure	32
4.1.1	Random Coil	33
4.1.2	Alpha Helix	33
4.1.3	Beta Strand	34
5	Hagfish Slime Threads	34
5.1	Sample Preparation	37
5.2	Structure and Mechanical Properties	37
5.3	Nuclear Relaxation in Hagfish IF Proteins	41
6	Resilin: A Resilient Rubber-Like Protein	49
6.1	Chemical Shift Assignment of Recombinant Resilins	51
6.2	Secondary Structure of Recombinant Resilin	55
6.3	Relaxation Measurements	57
7	Incorporation of Metal Ions into Hagfish IF: Preliminary Work	59
7.1	Sample Preparation	60
7.2	Using REAPDOR to Determine Metal Binding Site	61
7.2.1	T_2 Relaxation in Hagfish IFs	66
7.2.2	REAPDOR Simulation	68
7.3	Energy Dispersive X-Ray Spectroscopy	69
8	Concluding Remarks	74
	References	77
A	Lineshape Simulation Code	81

List of Tables

5.1	Mechanical properties of some biomaterials and high performance synthetic materials, adapted from [8, 13].	35
5.2	Amino acid composition of hagfish IF polypeptides.	35
5.3	T_1 relaxation times for hagfish IFs. A single number indicates a one component exponential fit; two numbers represent a two component fit.	43
5.4	Simulated τ_c results for the mobile carbonyl carbons.	48
6.1	Amino Acid content of resilin samples. Residues not listed are not present in either sequence.	50
6.2	Simulated chemical shift variance.	57
6.3	rec-1 relaxation times and corresponding correlation times τ_c . Temperature dependence trend indicates that backbone and aromatic carbons are in the slow regime.	58
6.4	an-16 relaxation times and corresponding correlation times τ_c	59
7.1	T_2 relaxation times.	67
8.1	Summary of dynamic characteristics for structural proteins.	76

List of Figures

2.1	Evolution of magnetic moment about a static magnetic field.	3
2.2	Energy levels for spin-1/2 in a static magnetic field B_0	3
2.3	The Rotating Frame (x', y', z'): A frame taken to precess at ω_{rf} about the static field \mathbf{B}_0 or the lab frame z -axis.	9
2.4	A simple pulse sequence	10
2.5	Electron orbitals are influenced by the external field, \mathbf{B}_0 , inducing an opposing field, \mathbf{B}_{cs} , which causes the shielding effect.	11
2.6	A powder lineshape produced by all possible orientations in the Zeeman field. The IUPAC standard has $\delta_{11} \geq \delta_{22} \geq \delta_{33}$, the 33 component representing the lowest frequency.	14
2.7	The dipolar interaction between two spins I and S ; a result of the coupling of the magnetic fields produced by the dipole moments. θ and ϕ are the polar and azimuthal angles of the internuclear vector \mathbf{r} with respect to the static field \mathbf{B}_0	15
2.8	State transitions corresponding to the various components of the dipolar Hamiltonian.	16
2.9	The random relative motion of atoms or molecules contributes to relaxation. Here the variation of the dipolar interaction between two spins due to rotational tumbling is pictured. The local magnetic fields \mathbf{B}_{loc} , are dependent on the relative orientations of the two nuclei with magnetic moments depicted by the dark arrows.	18
2.10	Definition of τ ; a longer interval generally results in very different local fields $B(t)$. .	19
2.11	Temperature effects and τ_c, T_1 . The <i>Cold</i> and <i>Hot</i> markers denote the direction in which the T_1 value will move in each regime as the temperature is changed.	22
3.1	Magic Angle Spinning.	25
3.2	Phase cycling.	26
3.3	Bloch Decay NMR experiment employing cross-polarization pulse to induce transverse magnetization in the ^{13}C nuclei.	27
3.4	The toggling frame or interaction representation.	28
3.5	Trial data from EDX experiment showing elemental content of GTC bundle.	31
4.1	Protein amino acids are all similar, differing in the attached side chain R.	31
4.2	The peptide unit; a) the <i>trans</i> form with $\omega = 180^\circ$; b) the <i>cis</i> form with $\omega = 0^\circ$. . .	32

4.3	(a) The alpha helix structure, hydrogen bonds (dashed lines) between carbonyl oxygen and amides; (b) the single beta strand and beta sheet, showing hydrogen bonding between adjacent beta strands via the peptide unit. Image taken from [6].	33
5.1	Pacific Hagfish slime; an effective defense mechanism. Photo courtesy of D. Fudge, P. Guerette.	34
5.2	A schematic model of the polypeptide composing the hagfish slime gland thread cell (GTC). Image taken from [13].	36
5.3	X-ray diffraction results for a) unstrained b) high strain hagfish IFs. Images taken from [13].	36
5.4	The proposed nature of the transition from α -helix to β -strand under strain. Image taken from [13].	36
5.5	Stress strain curves for hagfish IFs a) α -helical, elastic region b) deformation into plastic, β -sheet region. Images taken from [13].	37
5.6	Direct Polarization results for hagfish IFs at 20°C. Arrows indicate citrate peaks. . .	38
5.7	Cross polarized experimental results for hagfish IFs at 20°C.	39
5.8	Direct Polarization results for hagfish IFs at varying temperature. Temp = 20°C, 5°C, – 15°C.	40
5.9	Temperature dependence of mobile C=O peak in DP spectra of hagfish IFs.	41
5.10	T_1 inversion recovery experiment employing cross-polarization pulse to study longitudinal relaxation of the ^{13}C nuclei. Note the $\frac{\pi}{2}$ pulses are generally not the same duration on both channels.	42
5.11	Spectrum from T_1 inversion recovery experiment with $\tau = 1\text{ms}$, Temp = 20°C. . . .	43
5.12	T_1 experimental data for (a) carbonyl, (b) α -carbons, (c) other side chain carbons at Temp = 20°C; (d) carbonyl data requires a two component fit at Temp = –20°C.	44
5.13	The Tyr ϕ -peak T_1 results for Temp = 20°C experiment.	45
5.14	A simulated lineshape overlayed on a DP spectrum collected at Temp = 20°C. A correlation time of $\tau_c = 0.06\mu\text{s}$ achieves a reasonable fit.	47
5.15	Simulated τ_c temperature dependence results displaying Arrhenius fit with $E_a = 40.3$ kJ/mol.	48
6.1	Amino acid sequence of recombinant resilin, rec-1.	50
6.2	Amino acid sequence of an-16 resilin.	50

6.3	Dityrosine complex formed via photo-crosslinking. The crosslink site is shown here at the ϵ -carbon but may also occur at the δ -carbon.	51
6.4	DP experimental spectra (a) rec-1 resilin (b) an-16 resilin. The broad peak centered at 120ppm is due to probe background.	51
6.5	rec-1 resilin DP spectra peak assignment (a) carbonyl and aromatic region (b) α - and β -carbon region.	52
6.6	an-16 resilin DP spectra peak assignment (a) carbonyl and aromatic region (b) α - and β -carbon region.	54
6.7	rec1 and an-16 DP spectra overlay (a) carbonyl and aromatic region (b) α - and β -carbon region.	55
6.8	Deviation from observed chemical shift (a) rec-1 resilin, $N = 13$ (b) an-16 resilin, $N = 8$	56
6.9	T_1 experiment employing background suppression via phase modulation.	57
6.10	T_1 spectra with $\tau = 0.01s$ and Temp = 4°C a) rec-1 resilin b) an-16 resilin.	57
7.1	3,4-dihydroxy-L-phenolalanine (L-DOPA). (a) DOPA is much like the amino acid tyrosine but has a second hydroxyl group on the aromatic ring; (b) a divalent metal (in this case Ca^{2+}) may form an ionic bond with the oxygen atoms.	60
7.2	A REAPDOR pulse sequence. The x and y are hard π pulses at $\frac{T_R}{2}$ intervals. The pulse on the S channel is an adiabatic passage recoupling pulse.	62
7.3	The adiabatic passage pulse is required to invert an adequate population of S spins.	65
7.4	An example of a REAPDOR experimental curve.	66
7.5	Hahn spin echo experiment with cross-polarization. The delays τ_1 and τ_2 are very similar, but the acquisition may start sooner than τ_1 after the π pulse to ensure the full echo is captured.	66
7.6	The spin echo. Vectors represent transverse magnetization in a frame rotating at the Larmor frequency ω_0 ; (a) Initial transverse magnetization following $\frac{\pi}{2}$ pulse; (b) after time τ the magnetization has dephased due to spins precessing at frequencies slightly different from the Larmor frequency, $\omega_i = \omega_0 + \delta\omega_i$; (c) a π pulse is applied, inverting magnetization; (d) magnetization refocuses at time τ after the π pulse.	67
7.7	Spin Echo experimental spectrum. $\nu_R = 5kHz$	68

7.8	REAPDOR Simulation for DOPA-Ca binding: $\nu_R = 5\text{kHz}$, $\nu_d = 57\text{Hz}$, $\chi = 50\text{kHz}$, $S = \frac{7}{2}$. The arrow indicates the T_2 limit for the tyrosine ϕ -carbon.	68
7.9	Trial SEM micrograph of GTC bundle with nearby CaCO_3 contaminant.	70
7.10	EDX spectra for control samples a) pure b) contaminated.	71
7.11	EDX spectra for samples treated with CaCO_3 a) positive results b) contaminated c) no apparent Ca presence.	72
7.12	EDX spectra for samples treated with CaCl_2 a) positive results b) contaminated c) no apparent Ca presence.	73
8.1	DP spectra comparison of elastin (C) with rec-1 resilin (A) and hagfish IF (B) spectra. (C) was taken from [32] and displays spectra for strained (dotted) and unstrained (solid) elastin.	75

Acknowledgments

A student undergoes many trials throughout their academic career as they progress from regurgitating simple facts, to learning to think critically and answer the questions that begin to materialize in their increasingly complex world. The student is lucky when they are allowed the opportunity to work with someone who nurtures and guides this learning process. I'm incredibly grateful to Carl Michal for being that someone. In addition to helping me over those little hurdles along the way, (and picking through some horrendously messy computer code), he has been a friend and mentor, thank you Carl.

Peers also play an important role in the students growth, growing alongside one another and helping along the way; be it with a homework problem, a helping hand while submitting your thesis, or to make you stop staring at the computer screen and drag you out to the pub for a relaxing beverage. Thanks to my lab mates, Jenny Chien-Hsin Tso, David Katz, Ronald Dong, Matthew Lam and Simon Hastings, for support and friendship.

I would like to thank my collaborators, John Gosline and the nice folks in his zoology lab. Paul Guerette, thank you for sliming me, without you and your fish I would not have had a project. To Chris Elvin and Dan Dudek I owe a great deal for providing material and support with the resilin samples.

I would also like to thank my friends and pseudo-family in Vancouver, thanks for putting up with my completely uninteresting ramblings on the topics of biophysics. Tim, Nick, Jason and Sara I wouldn't have survived out here without you.

Finally, I thank my family; William, Cathy, Sarah and David. Without you I wouldn't be here, you have provided guidance, support and unconditional love my entire life and I love you all.

1 Introduction

For many years, NMR has been used to successfully probe the nuclear environments of materials and elucidate structural information. Because the NMR spectrum allows the resolution of various chemical environments and reflects characteristics of molecular dynamics, it is ideally suited to the study of biological materials such as proteins. Chapter 2 discusses the fundamentals of NMR and Chapter 3 relates some commonly used experimental techniques in NMR as well as a brief discussion of energy dispersive x-ray spectroscopy (EDX), which was also used in this study. The two types of proteins investigated in this dissertation; Hagfish gland thread cell protein and recombinant forms of resilin protein, are the basis of promising structural biomaterials.

One of the goals of research in biomaterials is to find biologically inspired solutions to commercial, scientific and engineering problems. Many synthetic structural materials produced today have harmful side effects during manufacturing processes. Hagfish slime threads are composed of a conglomeration of dual helix protein pairs, forming an intermediate filament structure. The threads have interesting material properties, being remarkably elastic but also tough under high strain. Although the threads are not as mechanically robust as some synthetic materials, they may provide insight into structural or compositional factors that impart the desired characteristics to a material. The hagfish intermediate filaments also provide an excellent candidate for the study of intermediate filament structure and dynamics, difficult to expose in other living systems, such as within cells. In Chapter 5, a variety of NMR experiments are employed to describe the structure and mechanics of the hagfish IFs on a molecular scale. A possible key factor in the formation of the hagfish IFs is the incorporation of metal ions into the protein structure. Chapter 7 investigates the binding of Ca^{2+} to the protein using EDX and NMR experiments.

The recombinant resilin material is inspired by a rubber-like substance found in many arthropods. In nature, resilin provides a variety of energy storing functions in insects, from locomotion to sound production. The resilin material is attractive for its highly resilient properties; its ability to store and release mechanical energy over millions of cycles without appreciable loss in efficiency. This type of material could contribute substantially to devices or materials requiring a high cycle lifetime, such as biomedical implants or micromechanical actuators or sensors. Chapter 6 exposes the underlying molecular structure and dynamics that are believed to provide the mechanical properties of the material.

2 Nuclear Magnetic Resonance

Nuclear Magnetic Resonance (NMR) is a phenomenon that is exhibited by many nuclear species. An NMR active nucleus must have an intrinsic nuclear magnetic moment; this is a result of the nucleus containing an odd number of protons or neutrons. These nuclei and their chemical environments are studied by alignment of the magnetic moments in a large static magnetic field and observing the response of the nuclei to perturbing electromagnetic fields.

2.1 Zeeman Interaction

The starting point in understanding NMR is to examine the effect of a static magnetic field, \mathbf{B}_0 , on a nucleus. The nuclei interact with magnetic fields via their magnetic moment, which is defined by

$$\boldsymbol{\mu} = \gamma \mathbf{I}, \quad (2.1)$$

in which γ is the gyromagnetic ratio and \mathbf{I} is the angular momentum of the nucleus. The effect of the magnetic field is to exert a torque, \mathbf{T} , on the nuclei. This torque is defined as the time rate of change of the nuclear spin:

$$\mathbf{T} = \frac{d\mathbf{I}}{dt}, \quad (2.2)$$

and is dependent on the magnitude and direction of the static magnetic field

$$\mathbf{T} = \boldsymbol{\mu} \times \mathbf{B}_0. \quad (2.3)$$

Combining equations 2.2 and 2.3, a classical equation of motion for the nucleus is formulated:

$$\frac{d\boldsymbol{\mu}}{dt} = \gamma \boldsymbol{\mu} \times \mathbf{B}_0. \quad (2.4)$$

Solving 2.4 will show the magnetic moment precessing about the magnetic field at a rate $\omega_0 = -\gamma B_0$, called the Larmor frequency (see Figure 2.1).

The Hamiltonian for this simple system can be obtained by substituting the classical components by their quantized operator counterparts. For example, $\boldsymbol{\mu}$ becomes $\hat{\boldsymbol{\mu}} = \gamma \hbar \hat{\mathbf{I}}$ and the Hamiltonian is

$$\hat{H}_Z = -\mathbf{B}_0 \cdot \hat{\boldsymbol{\mu}}. \quad (2.5)$$

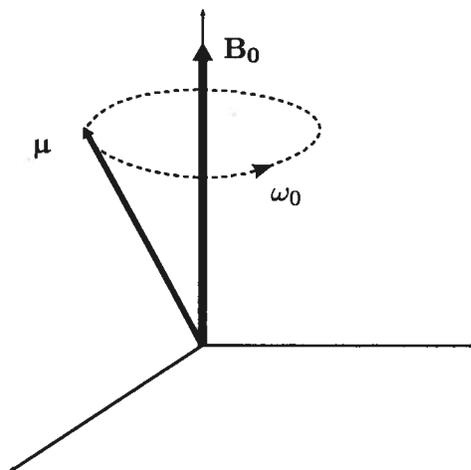


Figure 2.1: Evolution of magnetic moment about a static magnetic field.

By convention, the static magnetic field direction is assigned as the laboratory z -axis. Substituting Equation 2.1 for $\boldsymbol{\mu}$, the Zeeman Hamiltonian is written:

$$\hat{H}_Z = -\gamma\hbar B_0 \hat{I}_z. \quad (2.6)$$

The most trivial case is when dealing with a spin- $\frac{1}{2}$ nucleus, such as ^{13}C , where Equation 2.6 gives a two level eigensystem with energies $E_m = -m\gamma\hbar B_0$, with spin quantum number m denoting the eigenstate. A schematic of this two level system is shown in Figure 2.2.

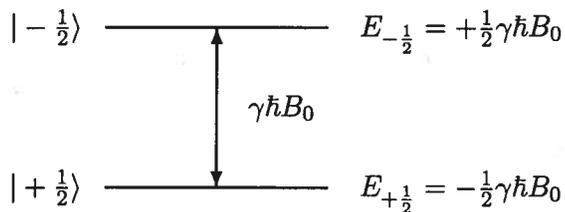


Figure 2.2: Energy levels for spin-1/2 in a static magnetic field B_0

NMR experiments deal with large numbers of quantum particles, and thus, can be described statistically in most cases. A spin ensemble, that is a collection of interacting nuclei, in the above described Zeeman field, reach a thermal equilibrium which can be characterized by the Boltzmann distribution. The population, p_m of a given energy level is

$$p_m = \frac{e^{-E_m/k_B T}}{Z}, \quad (2.7)$$

where k_B is the Boltzmann constant, T the temperature and Z is a normalization factor, the

so-called partition function:

$$Z = \sum_m e^{-E_m/k_B T}. \quad (2.8)$$

At equilibrium in non-zero temperature, there is a difference in the populations of the energy levels. This population difference elicits a net magnetization in the sample oriented parallel to the external magnetic field. The two energy levels in the spin- $\frac{1}{2}$ system correspond to magnetic moments which are anti-parallel to each other. The net sum of all the magnetic moments in the ensemble is the total magnetization $\mathbf{M}_0 = \frac{\sum_i \mu_i}{V}$ of the sample.

2.2 Fundamentals of the NMR Experiment

Most NMR experiments observe the response of the system after being subjected to a radio-frequency (rf) pulse. The pulse is administered via a current carrying coil which generates a fluctuating magnetic field, penetrating the sample inside the coil. The signal obtained is the precession of transverse magnetization created by the perturbing electromagnetic pulse. The signal is often received in the same coil; a fluctuating current created by the precessing magnetization. The applied rf pulse must have a frequency close to the Larmor frequency ($\omega_0 = \gamma B_0$) to perturb nuclei of that species.

The classical formula for a spin in a magnetic field, Equation 2.4, can be adapted to account for the average overall magnetization, \mathbf{M} in a sample:

$$\frac{d\mathbf{M}}{dt} = \gamma \mathbf{M} \times \mathbf{B}. \quad (2.9)$$

This expression is known as the Bloch Equation and gives a classical representation of a nuclear spin system. The benefit of the Bloch Equation is that it gives reasonable and intuitive results that can be calculated in a straightforward manner, without the need for involved quantum mechanical treatment. This formulation of the Bloch Equation has neglected nuclear relaxation effects which will be treated in the next section.

The applied rf-field $\mathbf{B}_1(t)$ is an oscillating field normally transverse to the static field direction (i.e. $\mathbf{B}_1 = \cos \omega_{rf} t \hat{x}$). The applied field exerts a torque on the magnetization, moving it from its equilibrium position parallel to the Zeeman field, into the xy -plane. The rate of the nutation away from the z -axis is proportional to the strength of the rf-field; $\omega_{nut} = \gamma B$. The resulting position of the magnetization is then dependent on the duration of the pulse, t_p . The tip angle can be

controlled by either the pulse length or rf-field strength:

$$\theta_p = \gamma B t_p. \quad (2.10)$$

The tip angle is an important feature to control in NMR experiments.

2.3 Relaxation of Magnetization

Relaxation is the process by which a macroscopic system returns from some non-equilibrium state to its statistical equilibrium. In the case of NMR, the macroscopic system referred to is the state of the magnetization. The processes by which relaxation can occur are many, however, the most prominent for nuclei such as ^{13}C and ^1H are the dipolar coupling and the chemical shift anisotropy. These mechanisms will be discussed in a later section.

Relaxation can be described on a phenomenological level as the time taken for the magnetization of a sample to return to equilibrium after a perturbing rf pulse. Nuclear relaxation is characterized by two parameters which describe the timescales on which the longitudinal (parallel with the Zeeman field), and transverse (perpendicular to the Zeeman field) magnetization takes to “relax”, that is, return to the thermal equilibrium. The classical equation of motion for the macroscopic net magnetization, \mathbf{M} , is the Bloch Equation (2.9). When relaxation is accounted for, the longitudinal relaxation time, T_1 and the transverse relaxation time, T_2 govern the time dependence of the magnetizations M_α , where $\alpha = x, y, z$.

$$\begin{aligned} M_x(t) &= M_x(0) \cos(\omega_0 t) e^{-\frac{t}{T_2}} \\ M_y(t) &= M_y(0) \sin(\omega_0 t) e^{-\frac{t}{T_2}} \\ M_z(t) &= M_0 + (M_z(0) - M_0) e^{-\frac{t}{T_1}}. \end{aligned} \quad (2.11)$$

Here, M_0 is the equilibrium value of the magnetization. Consequently, the Bloch Equation becomes:

$$\begin{aligned} \frac{d\mathbf{M}}{dt} &= \gamma \mathbf{M} \times \mathbf{B} - R(\mathbf{M} - \mathbf{M}_0) \\ R &= \begin{pmatrix} \frac{1}{T_2} & 0 & 0 \\ 0 & \frac{1}{T_2} & 0 \\ 0 & 0 & \frac{1}{T_1} \end{pmatrix}. \end{aligned} \quad (2.12)$$

The relaxation times T_1 and T_2 are observed characteristic timescales describing the recovery rates of the net magnetization in the sample. However, they have a link to deeper quantum mechanical processes that will be developed in Section 2.8.

2.4 NMR Spectroscopy

The precessing magnetization M after a pulse produces an oscillating current in the coil, and the resulting signal is tracked over time. This signal is known as the free induction decay (FID) of the system. Typical FID signals are on the order of MHz and are mixed down by a quadrature receiver with the rf-pulse frequency ω_{rf} . The sample will have a range of frequencies depending on the various environments of the spins under the influence of multiple interactions, as discussed in later sections. The observed frequency range $\Omega_0 = \omega_0 - \omega_{rf}$ is then centered around zero frequency and usually in the range of a few tens or hundreds of kHz. This mixing down of the acquired frequencies reduces the demand on electronics and the amount of data required. The FID is received in quadrature, the two signals are of the form

$$\begin{aligned} s_R(t) &\propto \sum_j \cos(\Omega_j t) e^{-t/T_2^j} \\ s_I(t) &\propto \sum_j \sin(\Omega_j t) e^{-t/T_2^j} \end{aligned} \quad (2.13)$$

The signals s_R , s_I should be understood as a superposition of signals with various frequencies, Ω_j . If the signals are combined into a single, complex time dependent function, $s(t) = s_R(t) + i s_I(t)$, then frequencies faster than the rf frequency can be distinguished from those lower than the rf frequency (i.e. $\Omega < 0$). This time domain signal is transferred to frequency space by Fourier transformation:

$$S(\Omega) = \int_0^\infty s(t) e^{-i\Omega t} dt. \quad (2.14)$$

The result is

$$S(\Omega) = \sum_j \frac{1/T_2^j}{(1/T_2^j)^2 + (\Omega - \Omega_j)^2} - i \frac{\Omega - \Omega_j}{(1/T_2^j)^2 + (\Omega - \Omega_j)^2} \quad (2.15)$$

These are the real (absorption) and imaginary (dispersion) part of a Lorentzian peak. Note that a fast decay will result in a broader peak, a slow decay results in a sharper peak.

2.5 Density Matrix Representation

Instead of representing the sample as an ensemble or superposition of single quantum states, it is often convenient to use *density matrix* representation, in which the spin system is described by a single density operator $\hat{\rho}$. Representing a collection of spins, each of which has a probability p_ψ of being in the state ψ , by a superposition state Ψ can be cumbersome as one must account for each spin individually. The expectation value of some operator \hat{A} in our spin ensemble must include a summation over all possible states for each spin:

$$\langle \hat{A} \rangle = \langle \Psi | \hat{A} | \Psi \rangle = \sum_{\psi} p_{\psi} \langle \psi | \hat{A} | \psi \rangle. \quad (2.16)$$

This can be simplified by representing the spin states in terms of a set of basis states, ϕ_i , convenient for our interacting spin ensemble. Now the possible states are

$$\psi = \sum_i c_{\psi i} \phi_i, \quad (2.17)$$

and the expectation value becomes

$$\langle \hat{A} \rangle = \sum_{\psi} p_{\psi} \sum_{i,j} c_{\psi i}^* c_{\psi j} \langle \phi_i | \hat{A} | \phi_j \rangle. \quad (2.18)$$

The advantage is now apparent, noting that the matrix elements $\langle \phi_i | \hat{A} | \phi_j \rangle$, are the same regardless of the state ψ we are dealing with. The ij^{th} density matrix elements are defined $\sum_{\psi} p_{\psi} c_{\psi i}^* c_{\psi j}$ and Equation 2.18 can be rewritten as

$$\langle \hat{A} \rangle = \sum_{i,j} \mathbf{A}_{ij} \rho_{ij} = \sum_i (\mathbf{A}\rho)_{ii} = Tr(\mathbf{A}\rho), \quad (2.19)$$

where \mathbf{A} is the matrix representing the operator \hat{A} in the basis ϕ_i . The expectation value of any observable can be determined by simply taking the trace of the product of the observable with the density matrix. The operator corresponding to the density matrix can be deduced from the matrix elements:

$$\hat{\rho} = \sum_{\psi} p_{\psi} |\psi\rangle \langle \psi|. \quad (2.20)$$

The Schrödinger equation can be used to formulate a relation which describes the time evolution

of the density matrix. This relation is known as the Liouville-von Neumann equation:

$$\frac{d\hat{\rho}}{dt} = \frac{i}{\hbar}[\hat{\rho}, \hat{H}]. \quad (2.21)$$

The solution of Equation 2.21 is $\hat{\rho}(t) = \hat{U}(t, 0)\hat{\rho}(0)\hat{U}^\dagger(t, 0)$, where $\hat{\rho}(0)$ is the initial density operator and $\hat{U}(t, 0)$ is a time evolution operator. The evolution operator is often termed the *propagator* and can be defined for general time-dependent Hamiltonians in the form:

$$\hat{U}(t, 0) = \hat{T}e^{-\frac{i}{\hbar} \int_0^t \hat{H}(t') dt'}, \quad (2.22)$$

where \hat{T} is the Dyson time-ordering operator which ensures that the order of propagators is properly maintained.

2.6 Theoretical Tools

When dealing with spin systems and their response to electromagnetic perturbations from the rf-field, there are a few mathematical and conceptual tricks that can be played to greatly simplify the analysis of the problem. Two that will be briefly discussed here are frame transformations and average Hamiltonian theory.

2.6.1 Frame Transformations

NMR experiments invariably entail the oscillating components of the spin operators in the static magnetic field due to Larmor precession. Since this action is universal, NMR problems are almost always simplified by examining the system in the *rotating frame*. The rotating frame is taken to rotate at the applied rf-frequency about the static lab frame z -axis (parallel to the static field \mathbf{B}_0 , see Figure 2.3).

This frame transformation is enacted analytically by subjecting the lab frame Hamiltonian to a rotation operator $\hat{R} = e^{-i\phi\hat{I}_z}$ in the following way:

$$\hat{H}' = \hat{R}^{-1}\hat{H}\hat{R} - \omega_{rf}\hat{I}_z. \quad (2.23)$$

Normally the pulses are applied on (or very close to) resonance ($\omega_{rf} \approx \omega_0$), thus the effective field \mathbf{B}_{eff} is simply that of the pulse \mathbf{B}_1 in the rotating frame. The advantage of such a transformation

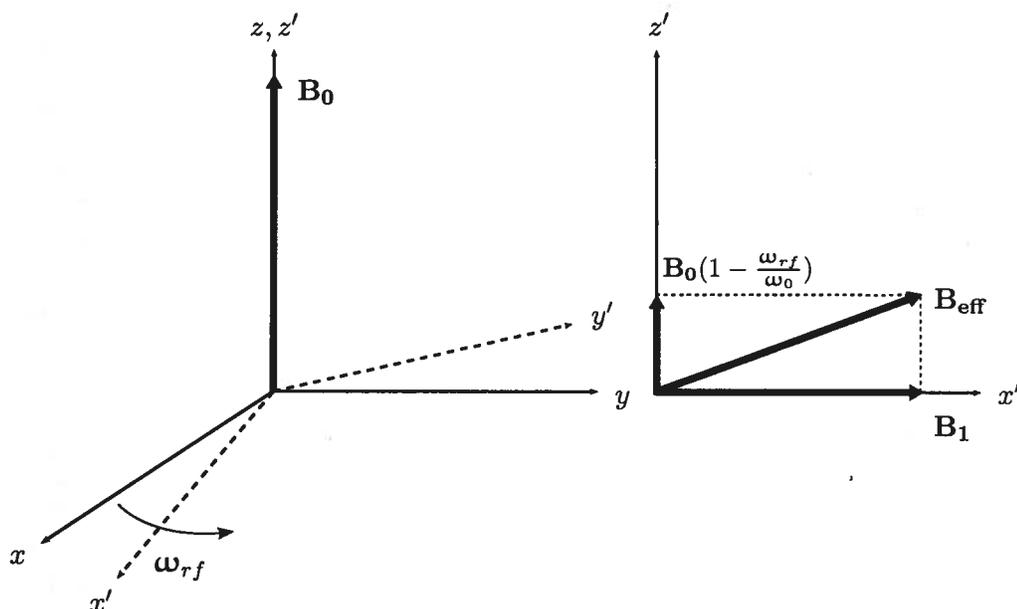


Figure 2.3: The Rotating Frame (x', y', z'): A frame taken to precess at ω_{rf} about the static field \mathbf{B}_0 or the lab frame z -axis.

is that when resonance is achieved, the only field present in the rotating frame is the applied field, \mathbf{B}_1 . Furthermore with the frame rotating at the rate of spin precession, terms associated with the static field, the Zeeman terms, are effectively nulled by the \hat{I}_z term (a result of insisting that the time-dependent Schrödinger equation is invariant under the transformation).

In the rotating frame, the magnetization is affected solely by \mathbf{B}_{eff} . It will precess about the effective magnetic field direction at a rate $\omega_p = \gamma B_{eff}$. In the lab frame, there is still precession about \mathbf{B}_0 and the resulting motion is complicated. There is nutation at a rate ω_p about an axis rotating about the lab frame z -axis at ω_0 .

2.6.2 Average Hamiltonian Theory

Because of the oscillatory nature of the interactions in NMR experiments, another convenient tool is Average Hamiltonian theory. This technique is employed by approximating periodic Hamiltonian terms as the average of the term over the required time. Average Hamiltonian theory can be especially useful for certain pulse sequences in which the sequence is made up of a chain of identical blocks. In this case, the propagator \hat{U} is described by a chronologically ordered product of propagators. The propagator for the pulse sequence shown in Figure 2.4 is

$$\hat{U}(t) = e^{-i\hat{H}_{rf}\delta_4} e^{-i\hat{H}_{fe}\tau_4} e^{-i\hat{H}_{rf}\delta_3} e^{-i\hat{H}_{fe}\tau_3} e^{-i\hat{H}_{rf}\delta_2} e^{-i\hat{H}_{fe}\tau_2} e^{-i\hat{H}_{rf}\delta_1} e^{-i\hat{H}_{fe}\tau_1} \quad (2.24)$$

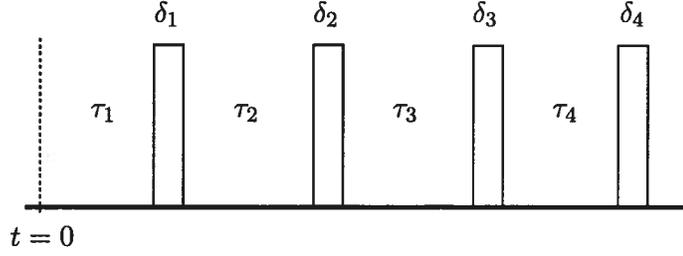


Figure 2.4: A simple pulse sequence

where \hat{H}_{rf} denotes the Hamiltonian during the rf irradiation, \hat{H}_{fe} the Hamiltonian during free evolution and δ_i, τ_i are the respective time periods over which each Hamiltonian acts. An equivalent propagator can be determined by using the Magnus expansion:

$$\begin{aligned}
 \hat{H}(t) &= \hat{H}_0 + \hat{H}_1 + \hat{H}_2 + \dots \\
 \hat{H}_0 &= \frac{1}{t} \int_0^t \hat{H}(t_1) dt_1 \\
 \hat{H}_1 &= \frac{i}{2!t\hbar} \int_0^t dt_2 \int_0^{t_2} [\hat{H}(t_2), \hat{H}(t_1)] dt_1 \\
 \hat{H}_2 &= \frac{1}{3!t\hbar^2} \int_0^t dt_3 \int_0^{t_3} dt_2 \int_0^{t_2} \left([\hat{H}(t_3), [\hat{H}(t_2), \hat{H}(t_1)]] \right. \\
 &\quad \left. + [\hat{H}(t_1), [\hat{H}(t_2), \hat{H}(t_3)]] \right) dt_1,
 \end{aligned} \tag{2.25}$$

where the \hat{H}_j represent the j^{th} order term of the Hamiltonian. The expansion takes advantage of the properties of exponential operators, i.e.

$$e^{\hat{A}} e^{\hat{B}} = \exp \left\{ \hat{A} + \hat{B} + \frac{1}{2!} [\hat{A}, \hat{B}] + \frac{1}{3!} \left([[\hat{A}, [\hat{A}, \hat{B}]] + [[\hat{A}, \hat{B}], \hat{B}]] \right) + \dots \right\} \tag{2.26}$$

Here $[\hat{A}, \hat{B}]$ denotes the commutator of operator \hat{A} with operator \hat{B} .

Equation 2.25 appears daunting, however it is frequently the case that the zeroth order term \hat{H}_0 , is sufficient to describe the resultant NMR spectrum. In fact, it is often the case that the Hamiltonians commute and all higher order terms will be zero. The zeroth order term is simply the time average of the piecewise Hamiltonians in the propagator $\hat{U}(t)$.

2.7 NMR Interactions

There are a variety of phenomena that occur in an NMR active sample. These range from interactions between adjacent spins to interactions with magnetic and electric fields, and they have a variety of effects on the resulting spectra. The following sections will discuss some of these

interactions important to the subsequent work.

2.7.1 Chemical Shielding

The chemical shielding interaction is a result of the external magnetic field. The static field induces preferential circulation of the electrons, and this motion in turn, generates additional magnetic field (Figure 2.5). This effect is primarily intramolecular; it is also seen in nearby nuclei but is normally not as dramatic due to the $1/r^3$ dependence of the magnetic potential.

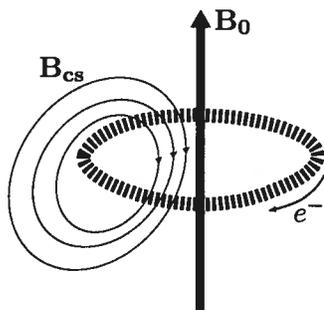


Figure 2.5: Electron orbitals are influenced by the external field, \mathbf{B}_0 , inducing an opposing field, \mathbf{B}_{cs} , which causes the shielding effect.

Generally in a large molecule, each nucleus has a slightly different chemical environment and thus experiences a slightly different local field \mathbf{B}_{cs} . This means that each spin will precess at a different frequency and a range of frequencies arises in the resulting spectrum. As alluded to above, the chemical shielding normally has an orientational dependence. This is a manifestation of the fact that the electronic environment around a nucleus is in general non-isotropic. The chemical shielding is thus characterized by a second rank tensor σ_{cs} describing the directional dependence of the orientation of the atom or molecule in the static field. In the principal axis frame (PAF) of the nucleus, σ_{cs} is diagonal:

$$\sigma_{cs}^{PAF} = \begin{pmatrix} \sigma_{xx}^{PAF} & 0 & 0 \\ 0 & \sigma_{yy}^{PAF} & 0 \\ 0 & 0 & \sigma_{zz}^{PAF} \end{pmatrix} \quad (2.27)$$

To a good approximation, the chemical shielding is linearly proportional to the static field and the Hamiltonian is

$$\hat{H}_{cs} = \gamma \hat{\mathbf{I}} \cdot \boldsymbol{\sigma} \cdot \mathbf{B}_0. \quad (2.28)$$

The shielding tensor is rarely aligned with the static field and thus the resulting frequencies pro-

duced will be dependent on the lab frame shielding components σ_{xx} , σ_{yy} , σ_{zz} . The lab frame chemical shielding tensor components can be calculated by performing a suitable Euler transformation on the PAF components. A further simplification can be made by transforming to a rotating frame as described in Section 2.6.1. The first order chemical shift Hamiltonian becomes

$$\hat{H}_{cs}^0 = \gamma \hat{I}_z \sigma_{zz} B_0, \quad (2.29)$$

where the periodic x and y components in the rotating frame average to zero. Equation 2.29 can be extended to find the first order contribution to the energy levels of the spin states:

$$\begin{aligned} E_{cs}^0 &= \gamma \hbar \sigma_{zz} B_0 \langle I, m | \hat{I}_z | I, m \rangle \\ &= \gamma \hbar \sigma_{zz} B_0 m, \end{aligned} \quad (2.30)$$

and the contribution to the frequency observed in the NMR signal is

$$\begin{aligned} \omega_{cs} &= \gamma \sigma_{zz} B_0 \\ &= -\omega_0 \sigma_{zz}. \end{aligned} \quad (2.31)$$

The total spectral frequency is the Larmor frequency plus the chemical shielding contribution, however, absolute frequencies are not measured in an NMR experiment. Rather, the frequency offsets from the applied rf frequency are measured with respect to the signal generated by a reference compound. In the case of ^1H and ^{13}C NMR, the reference compound is normally tetramethylsilane (TMS) since it has a smaller amount of shielding compared to most other compounds and the shielding is also relatively orientation independent. The displacement of the observed frequency from the reference signal is termed the *chemical shift* of the nuclear species in question. Since the frequencies obtained in an NMR experiment are field dependent, the chemical shifts δ are universally reported in parts per million (ppm). In this work, the δ -scale will be employed, following IUPAC convention:

$$\delta = \frac{\nu^{obs} - \nu^{ref}}{\nu^{ref}} \times 10^6 \quad (2.32)$$

The diagonal PAF components are conventionally labeled δ_{11} , δ_{22} , δ_{33} , where $\delta_{11} \geq \delta_{22} \geq \delta_{33}$. These components are often represented by their isotropic (δ_{iso}), anisotropic (Δ), and asymmetry

(η), values. The isotropic value is defined:

$$\delta_{iso} = \frac{1}{3} (\delta_{11} + \delta_{22} + \delta_{33}). \quad (2.33)$$

The definitions of Δ and η are dependent on the relative sizes of the components, if $|\delta_{iso} - \delta_{33}| > |\delta_{iso} - \delta_{11}|$ then

$$\begin{aligned} \Delta &= \delta_{33} - \delta_{iso} \\ \eta &= \frac{\delta_{11} - \delta_{22}}{\Delta}, \end{aligned} \quad (2.34)$$

otherwise

$$\begin{aligned} \Delta &= \delta_{iso} - \delta_{11} \\ \eta &= \frac{\delta_{33} - \delta_{22}}{\Delta}. \end{aligned} \quad (2.35)$$

If the PAF of the nucleus is oriented at respective polar and azimuthal angles, θ and ϕ , then the frequency contribution can be written in terms of the PAF components:

$$\omega_{cs} = -\omega_0 \left(\frac{1}{3} (\delta_{11} + \delta_{22} + \delta_{33}) + \delta_{11} \sin^2 \theta \cos^2 \phi + \delta_{22} \sin^2 \theta \sin^2 \phi + \delta_{33} \cos^2 \theta \right). \quad (2.36)$$

Using the definitions in Equations 2.34 and 2.33, Equation 2.36 becomes

$$\omega_{cs} = -\omega_0 \left(\delta_{iso} + \frac{1}{2} \Delta (3 \cos^2 \theta - 1 - \eta \sin^2 \theta \cos 2\phi) \right). \quad (2.37)$$

If a sample under NMR investigation had simply bare nuclei, the resulting spectrum would show an infinitely sharp single line at the resonant frequency of the nucleus. When dealing with solids, the variety of chemical environments normally encountered, especially in large organic molecules, spreads the line out into a broad peak, a collection of effective resonant frequencies, altered from the Larmor frequency by the chemical shielding. The anisotropy effectively characterizes the spread in the frequencies, see Figure 2.6. The isotropic value is the average of the principle components and the asymmetry provides a measure of the deviation from a cylindrically symmetric tensor, in which $\sigma_{22} = \sigma_{11}$.

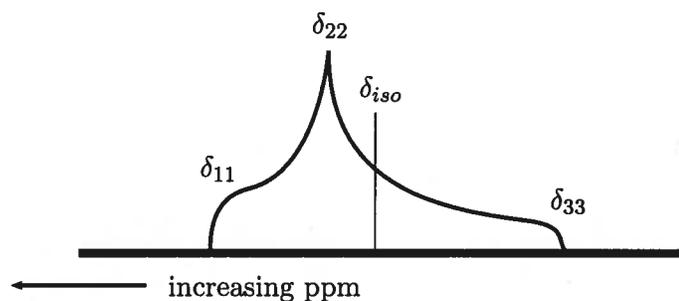


Figure 2.6: A powder lineshape produced by all possible orientations in the Zeeman field. The IUPAC standard has $\delta_{11} \geq \delta_{22} \geq \delta_{33}$, the 33 component representing the lowest frequency.

2.7.2 Dipolar Coupling

The dipolar interaction stems from the fact that a magnetic dipole produces a field of its own. This field will influence other dipoles through space, but is not normally far-reaching as it also has a $1/r^3$ dependence. Classically, the energy of interaction between two nuclei represented by magnetic dipoles μ_1 and μ_2 with an internuclear distance r is

$$U = \left(\frac{\mu_1 \cdot \mu_2}{r^3} - 3 \frac{(\mu_1 \cdot \mathbf{r})(\mu_2 \cdot \mathbf{r})}{r^5} \right) \frac{\mu_0}{4\pi}. \quad (2.38)$$

Recalling the fact that the magnetic moment is represented by the operator $\hat{\mu} = \gamma \hbar \hat{\mathbf{I}}$, the dipolar coupling Hamiltonian can be written:

$$\hat{H}_d = -\frac{\mu_0}{4\pi} \gamma_I \gamma_S \hbar \left(\frac{\hat{\mathbf{I}} \cdot \hat{\mathbf{S}}}{r^3} - 3 \frac{(\hat{\mathbf{I}} \cdot \mathbf{r})(\hat{\mathbf{S}} \cdot \mathbf{r})}{r^5} \right). \quad (2.39)$$

Here the subscripts 1 and 2 have been represented by spins I and S , respectively. Equation 2.39 is conventionally rearranged into spherical coordinates and separated into 6 terms representing various couplings between the states of the spins.

$$\hat{H}_d = -\left(\frac{\mu_0}{4\pi} \right) \frac{\gamma_I \gamma_S \hbar}{r^3} [A + B + C + D + E + F] \quad (2.40)$$

where:

$$\begin{aligned}
 A &= \hat{I}_z \hat{S}_z (3 \cos^2 \theta - 1) \\
 B &= -\frac{1}{4} (\hat{I}_+ \hat{S}_- + \hat{I}_- \hat{S}_+) (3 \cos^2 \theta - 1) \\
 C &= \frac{3}{2} (\hat{I}_z \hat{S}_+ + \hat{I}_+ \hat{S}_z) \sin \theta \cos \theta e^{-i\phi} \\
 D &= \frac{3}{2} (\hat{I}_z \hat{S}_- + \hat{I}_- \hat{S}_z) \sin \theta \cos \theta e^{i\phi} \\
 E &= \frac{3}{4} (\hat{I}_+ \hat{S}_+) \sin^2 \theta e^{-2i\phi} \\
 F &= \frac{3}{4} (\hat{I}_- \hat{S}_-) \sin^2 \theta e^{2i\phi}
 \end{aligned} \tag{2.41}$$

\hat{I}_+ , \hat{S}_+ and \hat{I}_- , \hat{S}_- are the raising and lowering operators that act on spins I and S , respectively. The polar and azimuthal angles θ and ϕ are defined in Figure 2.7. The rearrangement of the dipolar

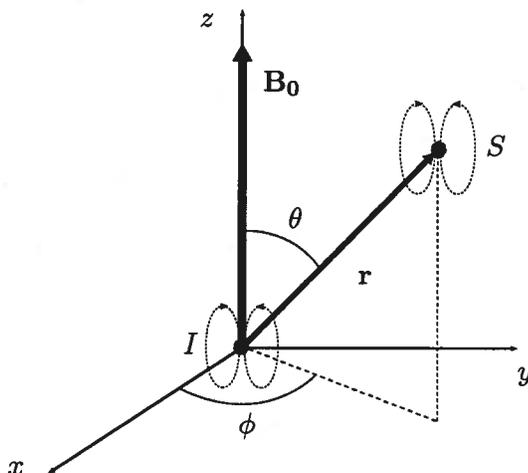


Figure 2.7: The dipolar interaction between two spins I and S ; a result of the coupling of the magnetic fields produced by the dipole moments. θ and ϕ are the polar and azimuthal angles of the internuclear vector \mathbf{r} with respect to the static field \mathbf{B}_0 .

Hamiltonian is beneficial as it separates the interactions in terms of energy level transitions. Figure 2.8 shows an energy diagram representing the states of a spin- $\frac{1}{2}$ system with two nuclei and relates the transitions to each of the components in Equation 2.41.

There are two classes of dipolar interaction; homonuclear dipolar coupling between like species of nuclei, and heteronuclear dipolar coupling between unlike species. Incidentally, homonuclear dipolar couplings would display an energy level splitting as in Figure 2.8, whereas heteronuclear energy level splittings will be different depending on the type of nuclei involved.

Both types of dipolar coupling Hamiltonians can be simplified greatly by a transformation to the rotating frame. When this is done for the homonuclear case, the terms C - F acquire periodic

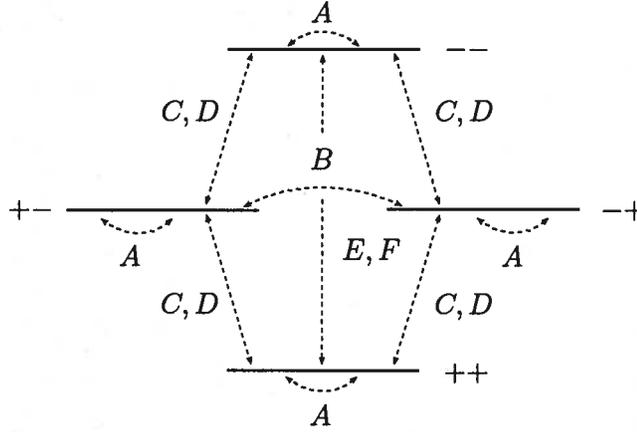


Figure 2.8: State transitions corresponding to the various components of the dipolar Hamiltonian.

time dependence with frequency ω_0 , and can be averaged away using average Hamiltonian theory (see Section 2.6). Similarly, these terms adopt periodic dependence with the heteronuclear case, however, the B term will also average away with a periodic dependence of ω_0 due to the fact that terms of the form $\hat{I}_{x/y}\hat{S}_{x/y}$ are affected by the rotation operator when I and S are not the same spin type. The first order average Hamiltonians for the dipolar interaction are thus:

$$\hat{H}_d^{homo} = -\frac{d}{2} (3 \cos^2 \theta - 1) (3 \hat{I}_z \hat{S}_z - \hat{\mathbf{I}} \cdot \hat{\mathbf{S}}) \quad (2.42)$$

$$\hat{H}_d^{hetero} = -d (3 \cos^2 \theta - 1) \hat{I}_z \hat{S}_z, \quad (2.43)$$

where we have labeled the dipolar coupling constant d :

$$d = \hbar \left(\frac{\mu_0}{4\pi} \right) \frac{1}{r^3} \gamma_I \gamma_S \quad (2.44)$$

The dipolar interaction can be expressed in tensor form $\hat{H}_d = -2\hat{\mathbf{I}} \cdot \mathbf{D} \cdot \hat{\mathbf{S}}$. In the PAF:

$$\mathbf{D}^{PAF} = \begin{pmatrix} -\frac{d}{2} & 0 & 0 \\ 0 & -\frac{d}{2} & 0 \\ 0 & 0 & d \end{pmatrix}. \quad (2.45)$$

Because the dipolar interaction depends on the internuclear distance, it can be exploited to reveal information about the structure of the sample under investigation.

2.7.3 Quadrupolar Interaction

Nuclei that have spin $I > \frac{1}{2}$, exhibit a feature known as the quadrupole interaction. This phenomenon is a result of the distribution of charge in the nucleus, giving rise to the electric quadrupole moment, which interacts with electric field gradients. An electric field gradient can arise naturally at the nuclear site by virtue of the distribution of charges in the vicinity, *i.e.* other nuclei and their electrons. The strength of the quadrupolar interaction can be quite large, often on the order of the interaction of the nucleus with an applied rf field. Thus it has a substantial effect on the spectrum and must be taken into account where necessary.

The quadrupolar Hamiltonian describes the effect of the interaction between the nuclear electric quadrupole Q and the electric field gradient V :

$$\hat{H}_Q = \frac{eQ}{2I(2I-1)\hbar} \hat{\mathbf{I}} \cdot \mathbf{V} \cdot \hat{\mathbf{I}} \quad (2.46)$$

for a spin I , where e is the charge on an electron. The expansion of Equation 2.46 is

$$\hat{H}_Q = \frac{eQ}{6I(2I-1)\hbar} \sum_{\alpha,\beta} V_{\alpha\beta} \left[\frac{3}{2} (\hat{I}_\alpha \hat{I}_\beta + \hat{I}_\beta \hat{I}_\alpha) - \delta_{\alpha\beta} \hat{I}^2 \right], \quad (2.47)$$

where α, β refer to the x, y and z coordinates of the defining reference frame, and $\delta_{\alpha\beta}$ is the Kronecker delta acquiring a value of 1 if $\alpha = \beta$ and 0 otherwise. The quadrupolar interaction can be represented in tensor form much like the chemical shift and dipolar tensors. The quadrupole coupling tensor in the PAF is

$$\chi = \frac{eQ}{\hbar} \mathbf{V} = \frac{eQ}{\hbar} \begin{pmatrix} V_{xx}^{PAF} & 0 & 0 \\ 0 & V_{yy}^{PAF} & 0 \\ 0 & 0 & V_{zz}^{PAF} \end{pmatrix}. \quad (2.48)$$

The electric field gradient tensor, \mathbf{V} is traceless in the PAF and has no isotropic component. The anisotropy, eq and asymmetry, η_Q of the interaction are conventionally defined in the same manner as the chemical shift:

$$\begin{aligned} eq &= V_{zz}^{PAF} \\ \eta_Q &= \frac{V_{xx}^{PAF} - V_{yy}^{PAF}}{V_{zz}^{PAF}}. \end{aligned} \quad (2.49)$$

The quadrupolar PAF generally does not align with the Zeeman field and the interaction Hamiltonian is quite complex. Fortunately, if the Zeeman interaction is dominant in the system, then only the secular parts of the quadrupolar Hamiltonian contribute significantly to the energies of the spin states. The first order Hamiltonian of the quadrupolar interaction has a form very similar to the chemical shift interaction:

$$\hat{H}_Q^0 = \frac{e^2qQ}{8I(2I-1)} (3 \cos^2 \theta - 1 - \eta_Q \cos(2\phi) \sin^2 \theta) \left[3\hat{I}_z^2 - I(I+1)\hat{1} \right]. \quad (2.50)$$

2.8 Relaxation Theory

The Bloch characterization of the relaxation times developed in Section 2.3 does give a general idea of the timescale of equilibration in a sample, but more information regarding molecular dynamics and structure can be obtained. The cause of relaxation for spins- $\frac{1}{2}$ is the fluctuation of magnetic fields at the site of the nucleus, primarily due to thermal motion. The fluctuating magnetic fields stem from the various interactions that the NMR nuclei can experience and a quantum mechanical treatment can illuminate dynamic characteristics of the sample.

2.8.1 Correlation Function and Spectral Density

For systems of organic molecules involving spin- $\frac{1}{2}$ nuclei, the primary mechanisms behind the fluctuating magnetic fields are the dipolar interaction and the chemical shift anisotropy. Both of these interactions are orientation dependent, so if there is any relative motion of nuclear sites, these interactions acquire some time dependence, see Figure 2.9. The time dependence of the local

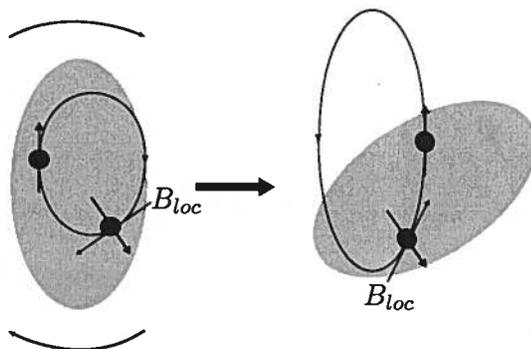


Figure 2.9: The random relative motion of atoms or molecules contributes to relaxation. Here the variation of the dipolar interaction between two spins due to rotational tumbling is pictured. The local magnetic fields \mathbf{B}_{loc} , are dependent on the relative orientations of the two nuclei with magnetic moments depicted by the dark arrows.

magnetic fields can be described by a correlation function, $G(\tau)$. $G(\tau)$ is essentially a measure of

how well the local magnetic field at a time t matches that at a time $t + \tau$. This can be described classically by the ensemble average of the product of the magnetic fields

$$G(\tau) = \overline{B(t)B(t + \tau)} \quad (2.51)$$

where the overbar denotes an average over all spins in the sample. If the interval τ is long compared to the speed of the fluctuations then the correlation function will decay rapidly. Otherwise the magnetic fields will in general be similar for most spins and the correlation function will decay more slowly. Figure 2.10 shows this relation between the timescale of fluctuation and the interval τ .

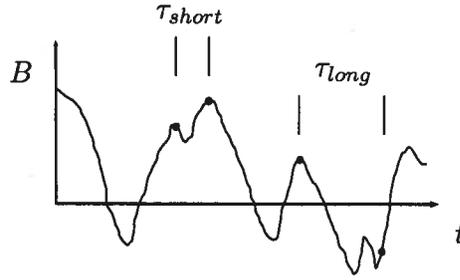


Figure 2.10: Definition of τ ; a longer interval generally results in very different local fields $B(t)$.

Quantum mechanically, we have a Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}_1$ describing the spin system subject to the Zeeman field and a randomly fluctuating Hamiltonian that accounts for the various relaxation mechanisms. In the rotating frame, the evolution of the density operator is given by the Liouville von Neumann equation:

$$\frac{d\hat{\rho}}{dt} = -i \overline{[\hat{H}_1, \hat{\rho}]}. \quad (2.52)$$

Through integration and use of the Magnus expansion, the value of the density matrix can be obtained for some arbitrary time t .

$$\hat{\rho}(t) = \hat{\rho}(0) - i \int_0^t \overline{[\hat{H}_1(t'), \hat{\rho}(0)]} dt' - \int_0^t dt' \int_0^{t'} \overline{[\hat{H}_1(t'), [\hat{H}_1(t''), \hat{\rho}(0)]]} dt'' + \dots \quad (2.53)$$

The time t is chosen such that the density operator undergoes a small change and the second order term is sufficient to describe the density operator $\hat{\rho}(t)$. The first order term vanishes as the average is zero. The form of the relaxing Hamiltonian \hat{H}_1 is:

$$\hat{H}_1(t) = \sum_m F_m \hat{T}_m e^{i\omega_0 t}, \quad (2.54)$$

where the \hat{T}_m are the spin operators and the F_m are the spatial functions which randomly change due to the thermal motion discussed above. The $e^{i\omega_0 t}$ term comes from the frame transformation on the spin operators. With this notation the correlation function, Equation 2.51 becomes

$$G(\tau) = \overline{F_m(t)F_m^\dagger(t+\tau)}. \quad (2.55)$$

The random motion may be caused by rotational, translational or vibrational motions or combinations of the three. An exponential correlation function agrees well with observed results

$$G(\tau) = Ce^{-\frac{\tau}{\tau_c}}, \quad (2.56)$$

where C is a constant depending on the interaction type which is causing the relaxation, normally itself depending on the position and orientation of nuclei involved; τ_c is referred to as the correlation time and is representative of the timescale on which the interaction Hamiltonian is changing. In theory, the correlation function need not be exponential, but it provides a simple model.

The Fourier transform of the correlation function is called the spectral density function. It describes the frequencies at which energy may be exchanged between the nucleus in question and its surroundings. These are, in fact, the transitions allowed by the Hamiltonian \hat{H}_1 . Since the correlation function is exponential, the spectral density function, labeled $J(\omega)$, is Lorentzian,

$$J(\omega) = C \int_{-\infty}^{\infty} e^{-\frac{\tau}{\tau_c}} e^{-i\omega\tau} d\tau = C \frac{2\tau_c}{1 + (\omega\tau_c)^2}. \quad (2.57)$$

The correlation function and subsequently the spectral density function materialize in Equation 2.53 when the substitution for the Hamiltonian \hat{H}_1 (Equation 2.54) is made:

$$\hat{\rho}(t) - \hat{\rho}(0) = \left[\hat{T}_m, \left[\hat{T}_m^\dagger, \hat{\rho}(0) \right] \right] \int_0^t dt' \int_0^{t'} G_{mm'}(t' - t'') e^{i(\omega_m t' - \omega_{m'} t'')} dt''. \quad (2.58)$$

The main contributions to the variation of $\hat{\rho}$ are those in which $m = m'$. A change of variables $\tau = t' - t''$ allows the realization of the correlation function in Equation 2.58:

$$\int_0^t dt' \int_0^{t'} G_{mm}(t' - t'') e^{i(\omega_m(t' - t''))} dt'' = \int_0^t (t - \tau) G_{mm}(\tau) e^{i\omega_m \tau} d\tau. \quad (2.59)$$

If the time t is assumed to be much longer than the correlation time τ_c , then $G_{mm}(\tau)$ will be

negligible for values of τ much less than the integration limit. Consequently, $t - \tau \approx t$ and the upper limit of integration can be considered infinite. Thus, the integral in 2.59 becomes the spectral density function $J(\omega)$,

$$\begin{aligned} \int_0^t (t - \tau) G_{mm}(\tau) e^{i\omega_m \tau} d\tau &\simeq t \int_0^\infty G_{mm}(\tau) e^{i\omega_m \tau} d\tau \\ &= t J_{mm}(\omega_m). \end{aligned} \quad (2.60)$$

This leads to the master equation for the evolution of $\hat{\rho}$, recalling that t is short enough that the density operator undergoes only a small change,

$$\frac{\hat{\rho}(t) - \hat{\rho}(0)}{t} \simeq \frac{d\hat{\rho}}{dt} = [\hat{T}_m, [\hat{T}_m^\dagger, \hat{\rho}]] J_{mm}(\omega_m). \quad (2.61)$$

The evolution of any physical observable $\langle \hat{Q} \rangle$ can now be determined:

$$\begin{aligned} \frac{d}{dt} \langle \hat{Q} \rangle = Tr \left\{ \hat{Q} \frac{d\hat{\rho}}{dt} \right\} &= - \sum_m Tr \left\{ \hat{Q} [\hat{T}_m, [\hat{T}_m^\dagger, \hat{\rho}]] \right\} J_{mm}(\omega_m) \\ &= - \sum_m J_{mm}(\omega_m) \left\{ \langle [[\hat{Q}, \hat{T}_m], \hat{T}_m^\dagger] \rangle - \langle [[\hat{Q}, \hat{T}_m], \hat{T}_m^\dagger] \rangle_{eq} \right\}. \end{aligned} \quad (2.62)$$

Note that to get to the second line of 2.62, the substitution $\hat{\rho} = \hat{\rho} - \hat{\rho}_{eq}$ has been made, accounting for the initial thermal equilibrium value of the density operator.

Using Equation 2.62 an expression for the longitudinal relaxation time T_1 can be derived upon insertion of the proper observable; the longitudinal magnetization described by \hat{I}_z . The relaxation time can then be related to the correlation time via the master equation.

Relaxation times have a somewhat complicated dependence on τ_c , to be discussed in the next section. A plot of T_1 versus τ_c for a typical dipolar interaction, Figure 2.11, exhibits a T_1 minimum at a correlation time just over 1 ns. To the left of the minimum, the correlation times are considered in the fast regime. This is the regime in which small molecules in non-viscous liquids generally appear. Larger molecules or viscous liquids tend to exhibit correlation timescales in the slow regime, to the right of the T_1 minimum. The dependence of relaxation time on temperature is different in the fast and slow regimes. Nuclei characterized by fast correlation times display an increase in relaxation time as the temperature is increased, whereas in the slow regime, T_1 increases as the sample is cooled.

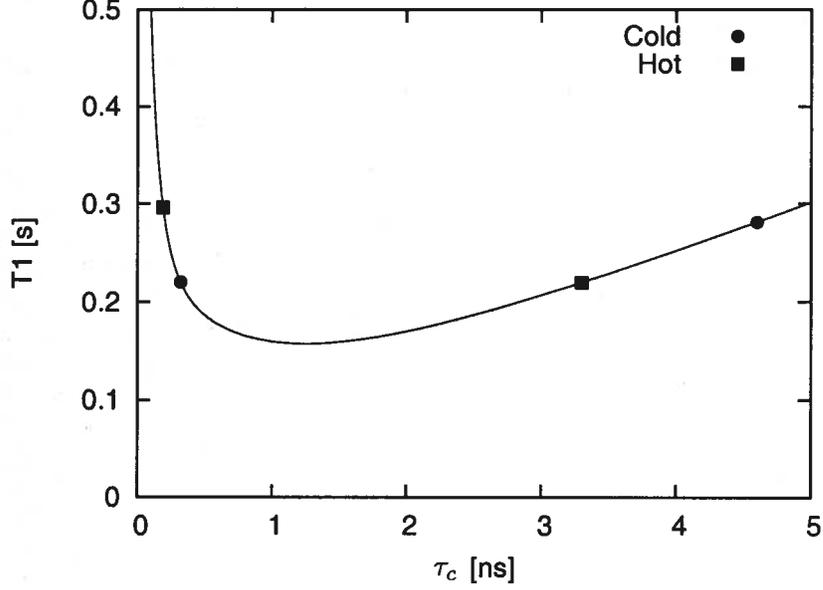


Figure 2.11: Temperature effects and τ_c , T_1 . The *Cold* and *Hot* markers denote the direction in which the T_1 value will move in each regime as the temperature is changed.

2.8.2 Dipolar Coupling

The first mechanism of relaxation that is relevant to this study is the dipolar coupling. The dipolar interaction Hamiltonian is

$$\hat{H}_d = -\frac{\mu_0 \gamma_I \gamma_S \hbar}{4\pi r_{IS}^3} \sum_{-2}^{+2} F_m(\theta, \phi) \hat{T}_m(\hat{I}, \hat{S}), \quad (2.63)$$

where, I and S denote in general two different spins. The derivation will be limited to the case where I and S are both spin- $\frac{1}{2}$. The spatial functions F_m are orientationally dependent and there is a random distribution of θ and ϕ throughout the ensemble which will ultimately result in the exponential correlation function. The spin operators \hat{T}_m are the spherical tensors:

$$\begin{aligned} F_0 &= \sqrt{\frac{3}{2}} (3 \cos^2 \theta - 1) \\ F_{\pm 1} &= \mp 3 \sin \theta \cos \theta e^{\pm i\phi} \\ F_{\pm 2} &= \frac{3}{2} \sin^2 \theta e^{\pm 2i\phi} \\ \hat{T}_0 &= \frac{1}{\sqrt{6}} (3\hat{I}_z \hat{S}_z - \hat{\mathbf{I}} \cdot \hat{\mathbf{S}}) \\ \hat{T}_{\pm 1} &= \mp \frac{1}{2} (\hat{I}_{\pm} \hat{S}_z + \hat{I}_z \hat{S}_{\pm}) \\ \hat{T}_{\pm 2} &= \frac{1}{2} \hat{I}_{\pm} \hat{S}_{\pm}. \end{aligned} \quad (2.64)$$

This is simply a different way of writing Equation 2.39, however it is worthwhile noting that in relaxation studies the non-secular terms (B-F in Equations 2.41) are important.

Using the Hamiltonian 2.63 and using the observables \hat{I}_z and \hat{S}_z , the longitudinal magnetizations, with Equation 2.62 one obtains two coupled expressions of the form

$$\begin{aligned}\frac{d}{dt}\langle\hat{I}_z\rangle &= -R_I\left(\langle\hat{I}_z\rangle - \langle\hat{I}_z\rangle_{eq}\right) - R_{IS}\left(\langle\hat{S}_z\rangle - \langle\hat{S}_z\rangle_{eq}\right) \\ \frac{d}{dt}\langle\hat{S}_z\rangle &= -R_{IS}\left(\langle\hat{I}_z\rangle - \langle\hat{I}_z\rangle_{eq}\right) - R_S\left(\langle\hat{S}_z\rangle - \langle\hat{S}_z\rangle_{eq}\right),\end{aligned}\quad (2.65)$$

with

$$\begin{aligned}R_I &= \frac{1}{10}\left(\frac{\mu_0}{4\pi}\right)^2\frac{\gamma_I^2\gamma_S^2\hbar^2}{r_{IS}^6}\tau_c\left(\frac{3}{1+(\omega_I\tau_c)^2} + \frac{1}{1+((\omega_I-\omega_S)\tau_c)^2} + \frac{6}{1+((\omega_I+\omega_S)\tau_c)^2}\right) \\ R_{IS} &= \frac{1}{10}\left(\frac{\mu_0}{4\pi}\right)^2\frac{\gamma_I^2\gamma_S^2\hbar^2}{r_{IS}^6}\tau_c\left(\frac{6}{1+((\omega_I+\omega_S)\tau_c)^2} - \frac{1}{1+((\omega_I-\omega_S)\tau_c)^2}\right).\end{aligned}\quad (2.66)$$

The terms R_α are interpreted as rates; the longitudinal relaxation time (assuming that the I spin is the species under investigation) would be $T_1 = \frac{1}{R_I}$. The terms with R_{IS} represent cross relaxation between the I and S spins. By symmetry $R_S(\omega_S, \omega_I) = R_I(\omega_I, \omega_S)$.

2.8.3 Chemical Shift Anisotropy

Relaxation due to anisotropic chemical shift is analogous to the dipolar interaction. In this situation a spin I interacts with a small field proportional to the Zeeman field $\mathbf{H} = H_0\hat{z}$. The Hamiltonian is

$$\hat{H}_{cs} = \sum_{-2}^{+2} F_m \hat{T}_m, \quad (2.67)$$

with

$$\begin{aligned}F_0 &= \frac{1}{2}\gamma_I\delta_z \\ F_{\pm 1} &= 0 \\ F_{\pm 2} &= \frac{1}{2\sqrt{6}}\gamma_I\delta_z\eta \\ T_0 &= 3\hat{H}_z\hat{I}_z - \hat{\mathbf{I}} \cdot \hat{\mathbf{H}} \\ T_{\pm 1} &= \frac{\sqrt{6}}{2}\left(\hat{H}_z\hat{I}_{\pm} + \hat{H}_{\pm}\hat{I}_z\right) \\ T_{\pm 2} &= \frac{\sqrt{6}}{2}\hat{H}_{\pm}\hat{I}_{\pm}\end{aligned}\quad (2.68)$$

The chemical shift tensor components δ_α are those of a traceless tensor and are related to the lab frame components discussed in Section 2.7.1:

$$\delta_\alpha = \sigma_\alpha - \sigma_{iso}, \quad (2.69)$$

and

$$\eta = \frac{\delta_y - \delta_x}{\delta_z}. \quad (2.70)$$

Using the same technique as in Section 2.8.2, one can determine the expressions for the relaxation due to chemical shift anisotropy. The longitudinal relaxation time is

$$\frac{1}{T_1} = \frac{3}{10} \gamma_I^2 H_0^2 \delta_z^2 \left(1 + \frac{\eta^2}{3}\right) \tau_c \frac{1}{1 + (\omega_0 \tau_c)^2}. \quad (2.71)$$

For a detailed derivation of the correlation function, spectral density functions and their relation to the master equations the reader is referred to [1] and [27].

3 Experimental Techniques

The following sections outline some of the experimental techniques used in the NMR experiments, and an explanation of energy dispersive x-ray spectroscopy.

3.1 Heteronuclear Decoupling

Decoupling is used to reduce broadening of spectral lines caused by heteronuclear broadening of dilute spins by nearby abundant spins, such as the situation of ^{13}C and ^1H in organic materials. To employ heteronuclear decoupling, high power rf irradiation is simply applied to the abundant spin channel during acquisition on the observe channel. A simple description of decoupling is that the high power irradiation causes the abundant spins to undergo transitions at a rate determined by the rf amplitude. For protons, the time averaged dipolar coupling contribution will be zero since the spin quantum number, m_H oscillates between $\pm \frac{1}{2}$.

3.2 Magic-Angle Spinning

Magic-angle spinning (MAS) is used frequently in solid-state and protein NMR. The primary effect is the removal of the chemical shift anisotropy but it also aids in removal of homo- and

heteronuclear dipolar coupling. The result is an improvement in resolution, narrowing lines that have been broadened by these interactions. Figure 3.1 displays the basic setup; the sample is contained within a rotor at an angle θ_M with respect to the lab z -axis. The rotor spins at a speed

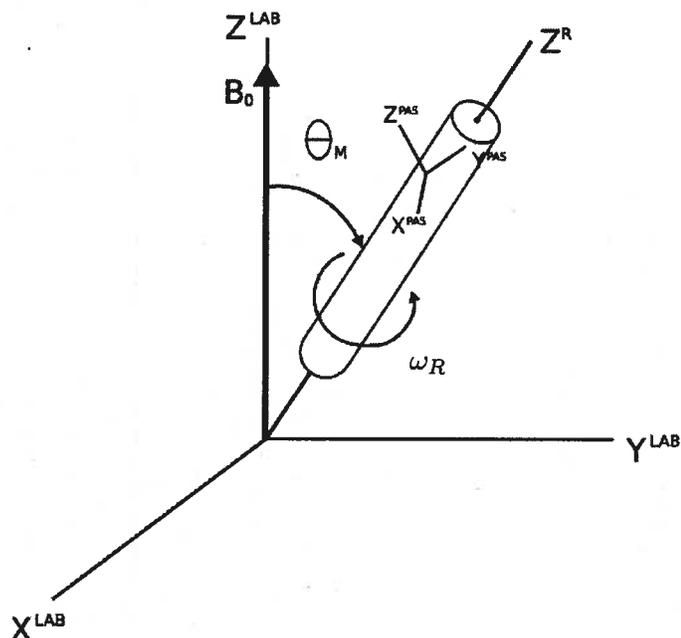


Figure 3.1: Magic Angle Spinning.

ω_R . If the angle between the z -axis in the PAF and the lab frame z -axis is θ , then the first order molecular orientation dependence for the discussed interactions will have the form $3 \cos^2 \theta - 1$. In a powder sample, all orientations θ are represented resulting in a broad lineshape. If the sample is spun, θ will become time dependent. Furthermore, if the spinning rate is fast compared to the anisotropy of the interaction, only the average of the orientation dependence, $\langle 3 \cos^2 \theta - 1 \rangle$ will contribute to the spectrum. The average value can be shown to be

$$\langle 3 \cos^2 \theta - 1 \rangle = \frac{1}{2} (3 \cos^2 \theta_M - 1) (3 \cos^2 \beta - 1), \quad (3.1)$$

where β is the angle between the z -axis in the PAF and the rotor axis. If the rotor angle, θ_M is set to 54.74° then the term $(3 \cos^2 \theta_M - 1) = 0$. Thus, the average $\langle 3 \cos^2 \theta - 1 \rangle$ is zero also. This means that the anisotropy will be averaged away to zero, leaving just the narrow isotropic peak.

If the spinning speed is not greater than the magnitude of the anisotropy of the interaction being spun out, then spinning sidebands will appear. The spinning sidebands are sharp lines that radiate out from the central isotropic line at intervals equal to the spinning rate. These lines will be seen in later spectra in Section 5.3. Spinning sidebands can be problematic if they obscure

desired signals, or if multiple sidebands from different peaks overlap.

3.3 Phase Cycling

Phase cycling is a technique used in almost all pulsed NMR experiments to eliminate undesirable signals due to effects such as field inhomogeneity or background signals. It can also alleviate problems caused by imperfect pulse lengths, especially in experiments requiring long pulse trains. In Chapter 7, such an experiment is considered, where uncompensated pulse imperfections have a dramatic effect on the performance of the sequence. Phase cycling is the strategy used to mitigate these problems.

Figure 3.3 shows a situation where the pulse length is just slightly too long. A π_x (or 180_x) pulse is one that rotates the magnetization by 180° about the x -axis. Here the pulse is imperfect, causing the magnetization to precess an extra amount dx , past the $-z$ axis. If the next pulse were

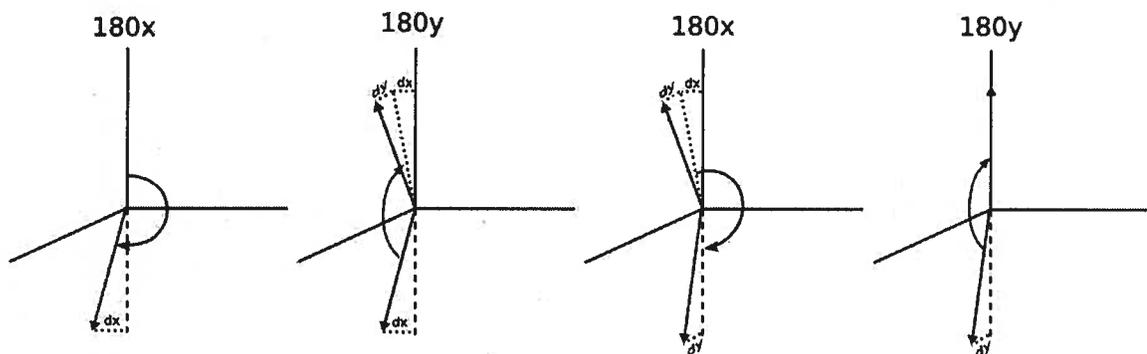


Figure 3.2: Phase cycling.

another π_x pulse, then it would compound the discrepancy and the magnetization would move further and further away from the desired orientation on the $\pm z$ -axis. To correct this, the phase of the pulses is alternated, first nutating magnetization about the x -axis and then the y -axis. As can be seen, the result is the correction of the discrepancy after a series of four pulses.

Background signals, or magnetization induced in materials outside the coil (not part of the sample) can be selectively removed from the spectrum through careful consideration of phase cycling.

3.4 Cross-Polarization

A cross-polarization (CP) experiment, Figure 3.3, employs simultaneous rf excitation on two channels to enable magnetization transfer, normally from an abundant species to dilute spins. For

example, excitation on the proton (^1H) and carbon (^{13}C) channels, induces the transverse magnetization of carbons for acquisition of an FID through spin-spin coupling. Here, the successful employment of CP is dependent on the close proximity of protons and ^{13}C nuclei, and therefore may provide some information about structure.

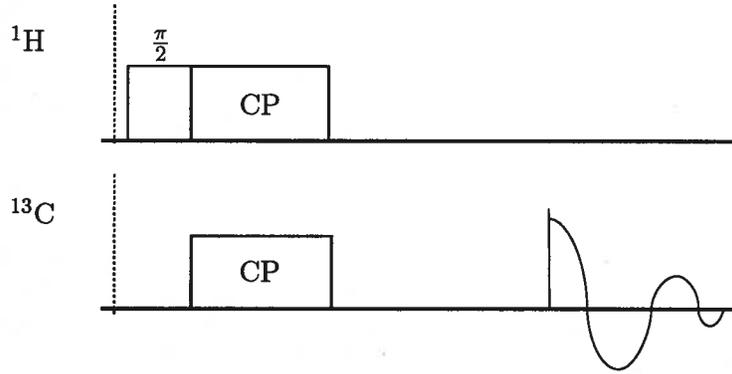


Figure 3.3: Bloch Decay NMR experiment employing cross-polarization pulse to induce transverse magnetization in the ^{13}C nuclei.

In the CP experiment an initial $\frac{\pi}{2}$ pulse creates ^1H magnetization along the $-y$ -axis of the ^1H rotating frame. The subsequent CP pulse, known as the spin-locking pulse, is applied on resonance along the $-y$ -axis. This results in a static field \mathbf{B}_H along the rotating frame $-y$ -axis, creating a situation which is analogous to the static Zeeman field, but facilitated by the applied rf-field. In essence, the CP pulse creates a new axis of quantization for the protons, with an energy gap $\Delta E_H = \hbar\gamma_H B_H$. A CP pulse is applied simultaneously at the resonance of the ^{13}C nuclei on the ^{13}C channel. This has a similar effect on the ^{13}C spin system, creating a static field in the rotating frame and a new axis of quantization for the ^{13}C spins; accordingly, $\Delta E_C = \hbar\gamma_C B_C$. When the amplitudes of these pulses are matched in a way to make the nutation frequencies equal (i.e. $\gamma_H B_H = \gamma_C B_C$), a condition arises which allows the transfer of magnetization from the protons to the ^{13}C nuclei. This can be considered a matching of the energy level separations of the rotating frame spin states of each nuclear species and is called Hartmann-Hahn matching. The magnetization transfer occurs through mutual spin flips via the heteronuclear dipolar interaction, Equation 2.43, between the protons and the nearby carbon nuclei. Since the secular Hamiltonian only contains operators \hat{I}_z and \hat{S}_z it is unaffected by the transformation to the doubly rotated frame. Though it appears that the Hamiltonian (2.43) only contains the “A”-type term (see Section 2.7.2) that would not result in magnetization transfer, the application of the pulses imparts additional interactions.

The rotating frame Hamiltonian \hat{H}_{rot} , for a ^{13}C surrounded by protons is

$$\hat{H}^{rot} = \hat{H}_{HH} + \hat{H}_{HC} + \omega_H \sum_i \hat{I}_{ix} + \omega_C \hat{S}_x, \quad (3.2)$$

where I_i and S denote the i^{th} proton and a ^{13}C nucleus, respectively. The final two terms represent the rf pulses in the rotating frame, with $\omega_H = \gamma_H B_H$ and $\omega_C = \gamma_C B_C$. The Zeeman Hamiltonian is nulled in the rotating frame transformation, thus only the homonuclear \hat{H}_{HH} and heteronuclear \hat{H}_{HC} dipolar terms appear in their rotating frame forms, Equations 2.42 and 2.43. It is assumed due to low abundance of ^{13}C that the homonuclear interaction for carbons is negligible.

At this point, it is prudent to do a further frame transformation into what is known as the *interaction representation* or the *toggling frame*. The toggling frame is an analogous extension of the rotating frame in which the applied pulse in the rotating frame is nulled by a similar rotational transformation. Much like adopting a frame rotating about the Zeeman axis at the Larmor frequency, the toggling frame is rotating about the applied field in the rotating frame, see Figure 3.4. In the present case, a doubly rotating frame (for both protons and carbon nuclei) is

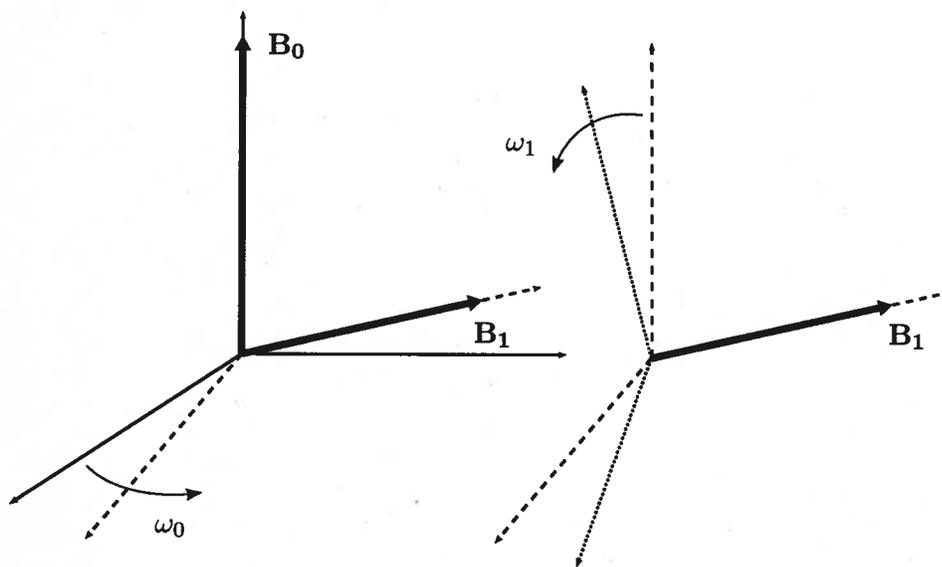


Figure 3.4: The toggling frame or interaction representation.

required, implemented by applying the rotation

$$\hat{R} = e^{-i\omega_H \sum_i \hat{I}_{ix}} e^{i\omega_C \hat{S}_x}, \quad (3.3)$$

to the rotating frame Hamiltonian, resulting in the toggling frame representation:

$$\hat{H}^{tog} = \hat{R}^{-1} \hat{H}_{rot} \hat{R} - \omega_H \sum_i \hat{I}_{ix} - \omega_C \hat{S}_x. \quad (3.4)$$

As in Section 2.6.1, the extra \hat{I}_{ix} and \hat{S}_x terms cancel perfectly with the pulse terms from Equation 3.2. The toggling frame Hamiltonian is thus rewritten:

$$\hat{H}^{tog} = \hat{H}_{HH}^{tog} + \hat{H}_{HC}^{tog}, \quad (3.5)$$

with

$$\hat{H}_{HH}^{tog} = -\frac{1}{2} \sum_{i>j} C_{ij}^{HH} (\hat{\mathbf{I}}_i \cdot \hat{\mathbf{I}}_j - 3\hat{I}_{ix}\hat{I}_{jx}) \quad (3.6)$$

$$\begin{aligned} \hat{H}_{HC}^{tog} = & -\sum_i C_i^{HC} \left[\left(\sum_i \hat{I}_{iz}\hat{S}_z + \sum_i \hat{I}_{iy}\hat{S}_y \right) \cos(\omega_H - \omega_C)t \right. \\ & \left. + \left(\sum_i \hat{I}_{iy}\hat{S}_z + \sum_i \hat{I}_{iz}\hat{S}_y \right) \sin(\omega_H - \omega_C)t \right], \end{aligned} \quad (3.7)$$

and

$$\begin{aligned} C_{ij}^{HH} &= \frac{1}{2} \frac{\mu_0}{4\pi} \frac{\gamma_H^2}{r_{ij}^3} (3 \cos^2 \theta_{ij} - 1) \\ C_i^{HC} &= \frac{1}{2} \frac{\mu_0}{4\pi} \frac{\gamma_H \gamma_C}{r_i^3} (3 \cos^2 \theta_i - 1). \end{aligned} \quad (3.8)$$

Here, r_{ij} and θ_{ij} denote the internuclear distance between protons i and j and the polar angle between the applied field and the vector \mathbf{r}_{ij} , respectively. Similarly, r_i and θ_i are the internuclear distance between the i^{th} proton and the carbon nucleus, and the polar angle between that vector and the applied field.

Only the heteronuclear dipolar interaction is significant for cross-polarization, although the homonuclear coupling between protons does have an impact on the redistribution of magnetization. The Hartmann-Hahn condition ($\omega_H \approx \omega_C$) removes the time dependence from Equation 3.7 and leaves terms of the form $\hat{I}_z\hat{S}_z$ and $\hat{I}_y\hat{S}_y$. The latter term can be represented as $\hat{I}_+\hat{S}_- + \hat{I}_-\hat{S}_+$ which is, in fact, a ‘‘B’’-type term promoting mutual spin flips between the I and S spins (see Equation 2.41 and Figure 2.8). Because the energy level separation is the same for both species under the matching condition, there is no loss of energy and there is no net change in magnetization when

there is a transfer of magnetization from the protons to the ^{13}C nuclei. Thus the heteronuclear system is allowed to come to an equilibrium; the plentiful high energy state protons redistributing magnetization to the nearby carbon nuclei. The result is a net magnetization of ^{13}C in the xy -plane which can then be observed on the carbon channel. This technique is simple to implement, and can reveal qualitative structural information regarding molecular mobility. It also provides a dramatic sensitivity enhancement over direct polarization ^{13}C spectra, for two separate reasons. The first is that the experiment may be repeated more frequently than the direct polarized version, because the ^1H T_1 is usually much shorter than the ^{13}C T_1 . The second reason is that the population difference induced by the magnetic field is much greater for ^1H nuclei than ^{13}C , because the ^1H gamma is a factor of ~ 4 greater. This enhanced population difference is then, in part, transferred to the ^{13}C by CP.

3.5 Energy Dispersive X-ray Spectroscopy

Energy dispersive x-ray spectroscopy (EDX) is a useful tool for characterizing chemical composition. The mechanics of EDX rely on the fact that different elements have different electronic structures, thereby eliciting a varied response under stimulation by an impinging beam of electrons or photons. The incoming beam particles excite and subsequently eject an electron from the inner shell of an atom in the beam path, leaving a hole. This hole is then occupied by a higher energy outer shell electron. When this electron makes its transition, energy is released in the form of a photon (in this case, an x-ray) which can be detected above the sample. The energy of the photon is of course equal to the energy gap that the electron transitioned through and is unique to a specific element. Thus the elements identity is provided by the frequency of the resulting x-ray. An example spectrum is shown in Figure 3.5.

The x -axis is the emitted photon energy and software analysis provides the elemental identification of the gaussian peaks. The y -axis represents counts of emitted photons at each energy and therefore the magnitude of the peak provides some information about the content of the element in the sample material.

4 Introduction to Proteins

Proteins are ubiquitous in nature. They fill a variety of roles from transporting material and controlling functions in cells to providing structural architecture with materials such as hair, nails

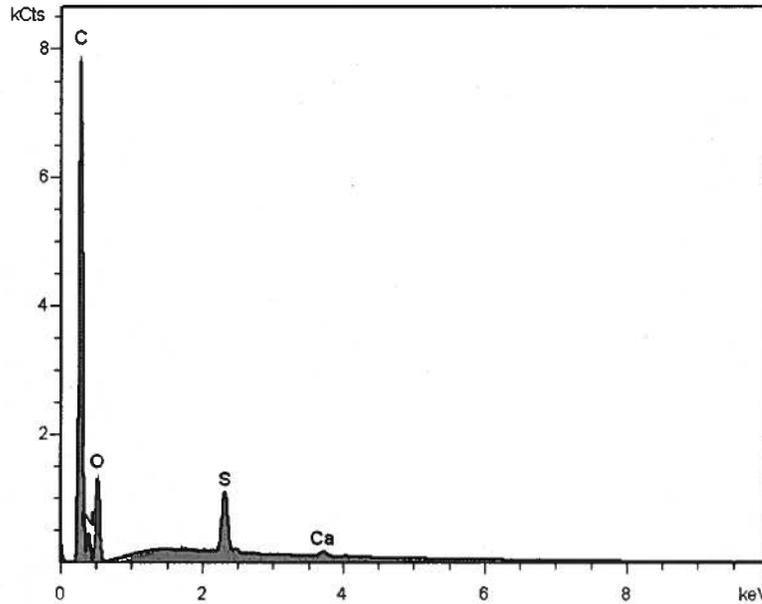


Figure 3.5: Trial data from EDX experiment showing elemental content of GTC bundle.

and tendons. Their diverse biological function is facilitated through various combinations of 20 different building blocks, the amino acids. It is the sequence of amino acids into polymeric chains that controls the properties and three-dimensional structures that the protein displays.

The amino acids share a common repeating backbone, see Figure 4.1, which provides the linear basis for the protein. The backbone carbons are known as the carbonyl- and α -carbons. The

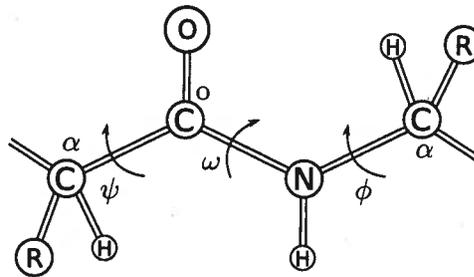


Figure 4.1: Protein amino acids are all similar, differing in the attached side chain R.

carbonyl carbon is, of course, attached to the oxygen atom via a double bond. The α -carbons root the amino acid side chain to the backbone. If they exist, the subsequent carbons are normally labeled β , γ and so on. However, not all amino acids are made up of linear carbon chains and some have alternative naming conventions.

The bond between the carbonyl carbon and the nitrogen atom is called the peptide bond and links adjacent amino acid residues together. This bond is unique in that it has partial double bond

characteristics forming a resonance with the carbonyl valency, C^o . Because of this, the peptide unit, pictured in Figure 4.2, is normally partially planar. The group of atoms forming the peptide unit is restricted by this bond and usually acts as a rigid structure. Rotation about the peptide bond is defined by an angle ω . The *trans* form is energetically favourable in most cases as there

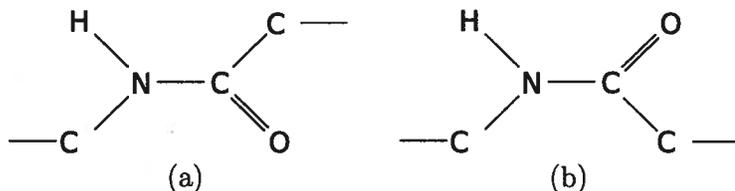


Figure 4.2: The peptide unit; a) the *trans* form with $\omega = 180^\circ$; b) the *cis* form with $\omega = 0^\circ$.

is less repulsion between non-bonded atoms. However, the amino acid residue proline, having a cyclic side chain bonded to the nitrogen atom as well as the alpha carbon, allows a stability that makes the *cis* form of comparable occurrence. The angles ϕ and ψ define torsional rotation about the $N - C^\alpha$ and $C^\alpha - C^o$ bonds, respectively.

Since NMR is sensitive to different molecular configurations it is ideal for studying proteins as the nuclei from different amino acids will exhibit different chemical shifts. Other interactions can also provide additional information, for example, the dipolar interaction can in some cases help to elucidate the three dimensional structure.

4.1 Protein Secondary Structure

The *primary* structure of proteins refers to the covalent structure defined by the amino acid sequence. There exists however, some freedom in the manner in which the covalent bonds are configured. At room temperature, bond lengths can vary by $\pm 0.05 \text{ \AA}$ and bond angles by $\pm 5^\circ$ [6]. The angles ϕ and ψ have some rotational freedom and assume different orientations dependent on the residue type. In large macromolecules such as proteins, these allowances can all contribute to a greatly flexible overall shape. Certain amino acids or combinations of amino acids can provide flexibility in the chain, or restrict motion. There are however, energetically preferential conformations that exist due to non-covalent interactions of molecules, either adjacent or distant, which govern what is known as the *secondary* structure of the protein. The various configurations of amino acids are invariably tied to the resulting protein secondary structure, often referred to as *folding*.

4.1.1 Random Coil

Random coil secondary structure refers to a dynamically changing configuration. There is rapid reorientation of the protein as the structure relies primarily on the intrinsic conformational properties of the amino acid residues. There is little or no association between distant parts of the polypeptide chain. The torsional angles ϕ and ψ freely rotate through -180° to $+180^\circ$, although the peptide bond ω is restricted to approximately -20° to $+10^\circ$.

4.1.2 Alpha Helix

The alpha helix is the most well known and easily identified of the secondary structures. The backbone forms a helix with 3.6 residues per turn and a translation of 1.5 Å per residue. The torsion angles ϕ and ψ are approximately -60° and -50° , respectively. The backbone atoms pack closely to form van der Waals interactions to favourably reduce the energy of the configuration. In addition, the carbonyl oxygen of each residue forms a hydrogen bond with the backbone-NH of the fourth residue along the chain, see Figure 4.3(a).

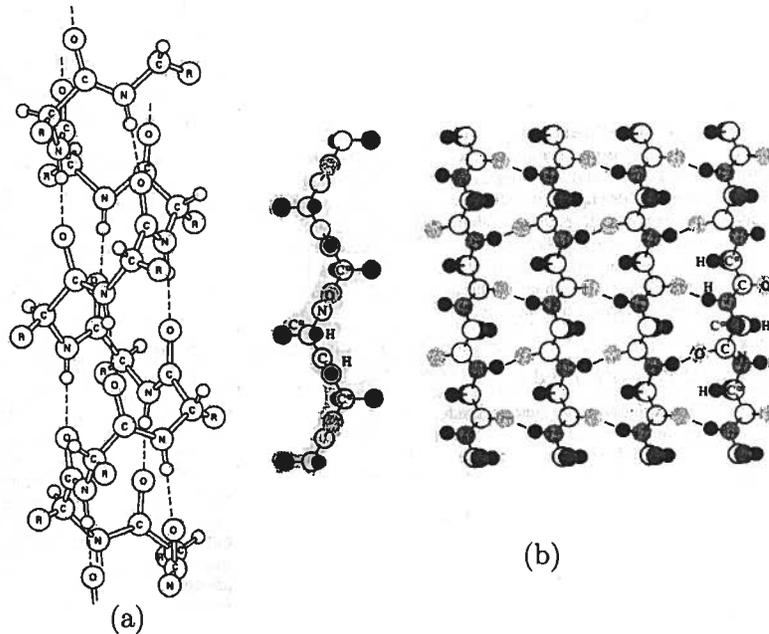


Figure 4.3: (a) The alpha helix structure, hydrogen bonds (dashed lines) between carbonyl oxygen and amides; (b) the single beta strand and beta sheet, showing hydrogen bonding between adjacent beta strands via the peptide unit. Image taken from [6].

The side chains project outward from the backbone and tilt slightly toward the amino end.

4.1.3 Beta Strand

The next most regular conformation is the beta strand, in which the protein is almost fully extended. This could be considered a type of helix with 2 residues per turn and a translation of 3.4\AA per residue. The beta strand is not energetically stable on its own, but must be incorporated into a beta *sheet*, which can be parallel or anti-parallel, see Figure 4.3(b). In the sheet, hydrogen bonds are formed between the peptide groups of adjacent beta strands. Beta sheets differ greatly from alpha helices; alpha helices can be considered basically one dimensional chains and the interactions occur between residues on the same peptide. In beta sheets, interactions occur between residues on different peptides.

5 Hagfish Slime Threads

Pacific Hagfish (*Eptatretus stouti*) are jawless, deep sea marine dwellers that have an interesting defense mechanism. When agitated the hagfish releases a combination of protein threads and vesicles composed of glyco-proteins, called mucins. These components interact with the surrounding seawater to produce a viscous slime that clogs the gills of predators, see Figure 5.1. The slime

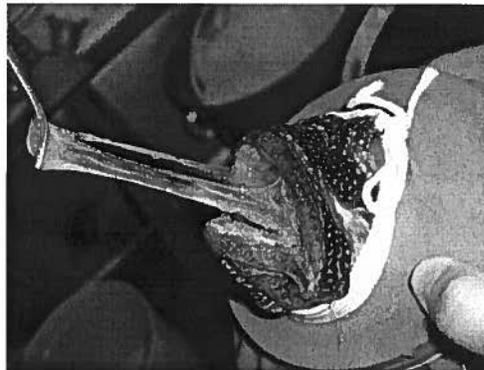


Figure 5.1: Pacific Hagfish slime; an effective defense mechanism. Photo courtesy of D. Fudge, P. Guerette.

threads are of particular interest owing to their incredible mechanical properties. Mechanical studies [13] have shown the threads to have low initial stiffness ($E_i = 6.4\text{ MPa}$) and high elasticity (up to strains of 0.34), as well as high tensile strength (180 MPa) and toughness (130 MJ/m^3). The mechanical properties of the IFs are compared to some other biomaterials and high performance fibres in Table 5.1.

The slime threads are also an excellent candidate for the study of intermediate filaments [7, 24, 25]; they offer an exposed view of the mechanics of IFs in living cells. The composition of the

Fibre Type	Modulus (MPa)	Tensile Strength (MPa)	Failure Strain (%)	Toughness (MJ/m ²)
Hagfish Thread	6.4	180	220	130
Spider Silk	10000	1800	30	130
Silkworm Silk	5000	600	12	50
Kevlar 49	130000	3600	3	50
Steel	200000	3000	2	123

Table 5.1: Mechanical properties of some biomaterials and high performance synthetic materials, adapted from [8, 13].

Amino Acid	(%) Composition		
	α	γ	total
Gly	15.9	14.6	15.2
Ser	10.3	11.9	11.1
Thr	9.0	7.5	8.2
Ala	7.5	7.8	7.6
Leu	5.9	8.3	7.1
Val	7.6	6.6	7.1
Glu	4.0	6.5	5.3
Gln	4.8	5.3	5.1
Ile	5.6	4.3	5.0
Arg	4.7	4.3	4.5
Tyr	4.0	3.7	3.9
Asp	3.7	4.0	3.9
Lys	3.6	3.2	3.4
Met	2.2	4.5	3.3
Asn	3.3	3.3	3.3
Pro	3.3	1.5	2.4
Phe	2.3	1.8	2.1
His	2.0	0.8	1.4
Trp	0.3	0.0	0.2
Cys	0.0	0.2	0.1

Table 5.2: Amino acid composition of hagfish IF polypeptides.

threads is believed to be tightly packed and oriented intermediate filaments (IFs), which are in turn made up of pairs of unlike polypeptides; α and γ . The amino acid sequences of α and γ have been deduced [24], and the amino acid content of the dual helix is listed in Table 5.2.

A model of the IFs has been proposed to explain the mechanical characteristics [14]. The model is keratin-like with a stiff central rod domain contained between flexible terminal ends and punctuated with non-helical linker regions which may provide additional flexibility. A simplified model is shown in Figure 5.2. X-ray diffraction data [13] has shown evidence of an α -helical conformation, and the central rod domain is believed to be a coiled-coil structure. The terminal ends, like the linker regions, are thought to be randomly structured. The X-ray diffraction results, displayed in Figure 5.3, also show that the protein undergoes an irreversible transition from α -helix to β -sheet under strain. The proposed effect of strain on this model is visualized in Figure 5.4. The change in structure is reflected in the stress-strain characteristics of the IFs. When the filaments

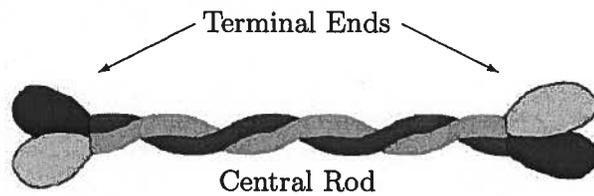


Figure 5.2: A schematic model of the polypeptide composing the hagfish slime gland thread cell (GTC). Image taken from [13].

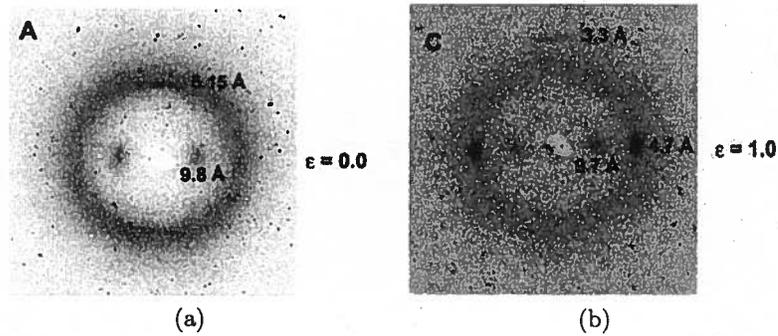


Figure 5.3: X-ray diffraction results for a) unstrained b) high strain hagfish IFs. Images taken from [13].

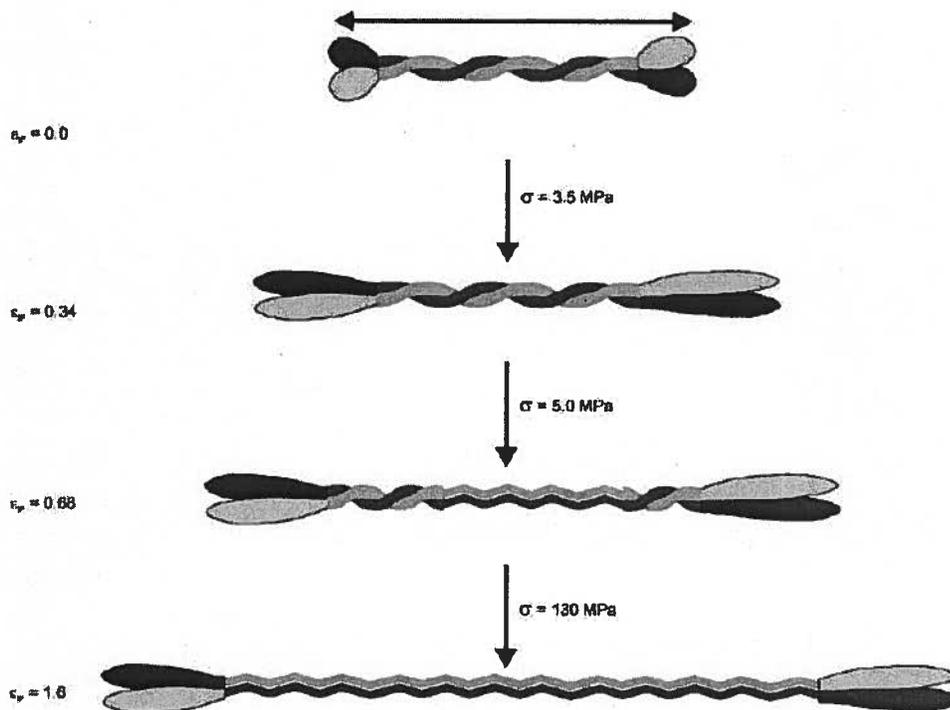


Figure 5.4: The proposed nature of the transition from α -helix to β -strand under strain. Image taken from [13].

are stressed past the point of elastic deformation, the transition to a β -sheet conformation imparts additional stiffness. This transition behaviour explains why the failure strain is so much higher

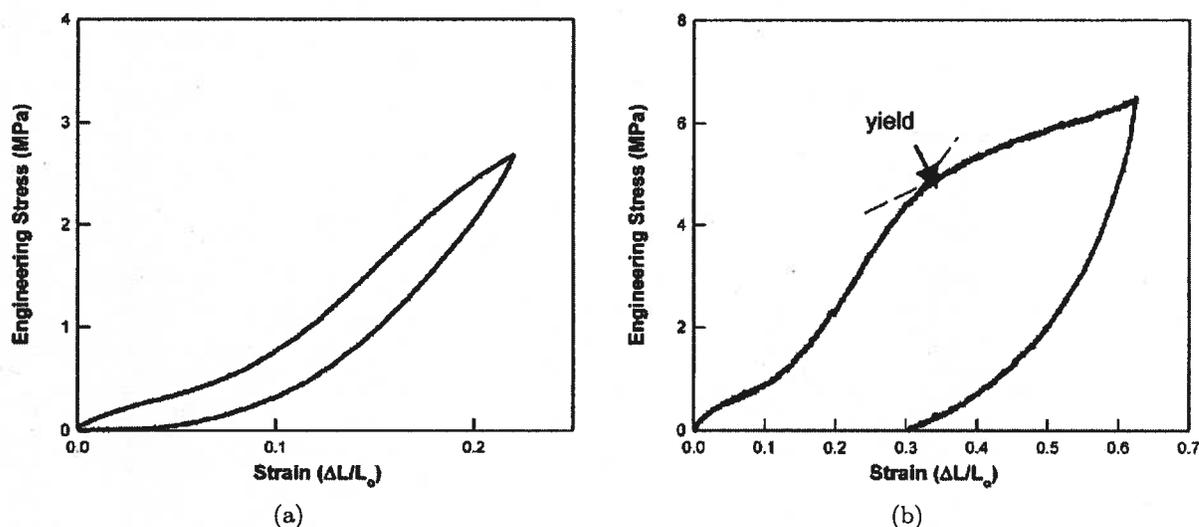


Figure 5.5: Stress strain curves for hagfish IFs a) α -helical, elastic region b) deformation into plastic, β -sheet region. Images taken from [13].

than the yield strain of the IFs.

5.1 Sample Preparation

Hagfish were captured with assistance of the Bamfield Marine Station in the Pacific Ocean near Barkley Sound. The fish were stored in a 200 L aquarium of chilled seawater. To collect slime, hagfish were anesthetized in a bucket with a buffered anesthetic (250 mg/L tetramethyl silane, MS-222; 500 mg/L sodium bicarbonate) in seawater until unresponsive. The fish were placed on a dissection tray and the skin was rinsed with de-ionized (milli-Q) water and blotted dry. Slime ejection was induced by mild electrical stimulation (8 V, 80 Hz) near the glands. The slime was collected with a plastic spatula and transferred to a buffer solution (0.9 M sodium citrate) via a 50 μ m filter to separate thread cells from mucins. A protease inhibitor cocktail (Sigma Aldrich P2714) was added to prevent protein degradation. The sample preparation techniques were slightly modified from [13] and samples were prepared by P. Guerette.

5.2 Structure and Mechanical Properties

General structural information can be obtained from quite simple NMR experiments. As alluded to in Section 3.4, magnetization transfer in a CP experiment will only occur between nuclei which are close in space. Furthermore, the CP pulse will only be effective if there is not significant relative motion between a proton and a carbon nucleus. Molecular motion disrupting the dipolar interaction will prevent efficient magnetization transfer between protons and carbon nuclei. Because of this, a

comparison of the results from direct polarization (DP) and CP experiments can suggest whether or not there may be mobility in the sample under investigation.

DP and CP experiments were performed on a Varian 400MHz (9.3T) spectrometer with an HXY MAS probe. Measurements from the hagfish IFs indicate a significant amount of motion in the backbone of the protein. A direct polarization experiment, in which the $\frac{\pi}{2}$ pulse is applied directly to the ^{13}C nuclei, shows mobility in the carbonyl peak. Figure 5.6 displays the results from a DP experiment performed at 20°C. The spectrum can be divided into four main regions;

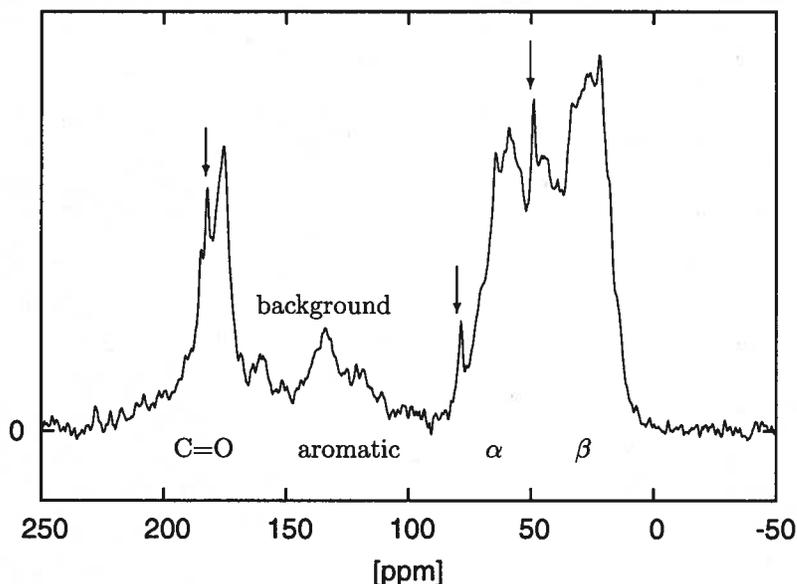


Figure 5.6: Direct Polarization results for hagfish IFs at 20°C. Arrows indicate citrate peaks.

the carbonyl region (C=O) is the furthest downfield from $\sim 150 - 250$ ppm; the α -carbons normally appear in a range between $\sim 40 - 80$ ppm; the β -carbons are generally found from $\sim 10 - 40$ ppm; the last region is the aromatic region which appears between the carbonyl and α -carbon range. Unfortunately the aromatic carbons in the hagfish IFs were not readily apparent in DP spectra, masked by a large background signal. Other carbons (γ , δ etc.) can occur in the α or β ranges and there are some exceptions, such as the Thr β -carbon which appears around 80 ppm.

The sharp peaks in the spectrum, indicated by the arrows, are signals from citrate carbons, rapidly tumbling in the buffer solution. The background signal comes from material outside of the coil, such as the end caps sealing the sample holder. The absence of a powder pattern in the carbonyl region suggest that there is significant motion. However, there is also evidence of a broad feature beneath the sharp carbonyl peak, suggesting the presence of some rigid structures in the backbone. This is confirmed in Figure 5.7, a CP spectrum under the same conditions. The

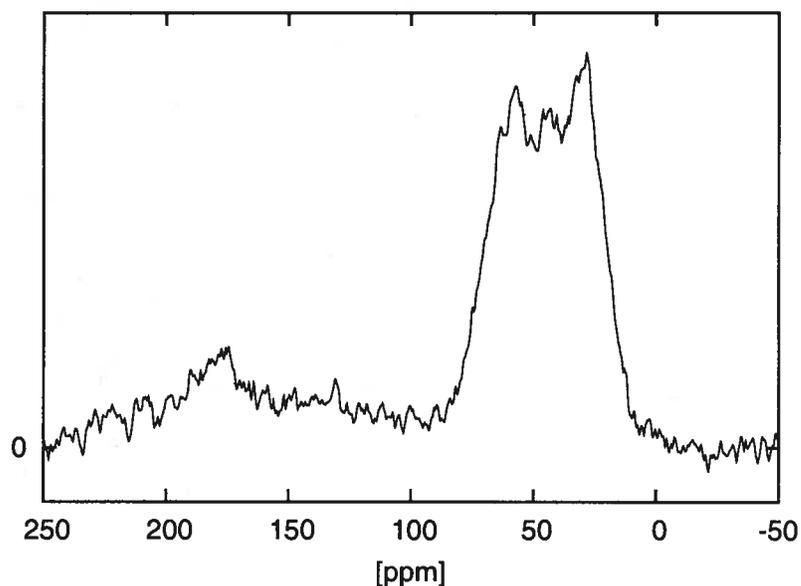


Figure 5.7: Cross polarized experimental results for hagfish IFs at 20°C.

background signal, coming from outside the coil and the citrate peaks from solution vanish in the CP spectrum. The broad portion of the carbonyl region cross polarizes well, thus remaining visible in both spectra. The highly mobile sharp peak has disappeared in the CP spectrum, rapid motion removing the possibility of magnetization transfer for these ^{13}C nuclei. These results suggest that there are different regions in the polypeptide; in some, there is high flexibility in the backbone, while others retain rigidity. These regions may contribute to the elastic and strength characteristics of the IFs.

Mobility in the peptide chain is also temperature dependent. Through analysis of the direct polarization spectra, the timescale of motion in different regions of the chain can be inferred from the size and shape of the various peaks. Direct polarization spectra were obtained at various temperatures (Figure 5.8). The carbonyl region shows a decrease in the sharp component as temperature drops. This could be attributed to restricted mobility, as the lack of rapid motion would result in dramatic broadening of the carbonyl line due to its large chemical shift anisotropy. The broad component becomes slightly more intense but as it is spread over such a large range, the increase is difficult to see. A qualitative inspection of the carbonyl region in the CP and DP results indicates that most of the carbonyl carbons are in relatively rigid backbone conformations. A small fraction, perhaps 10% based on the relative sizes of the sharp and broad peaks, can be attributed to mobile carbonyl carbons at room temperature.

In the α - and side chain carbon regions, the situation is more complicated. The chemical

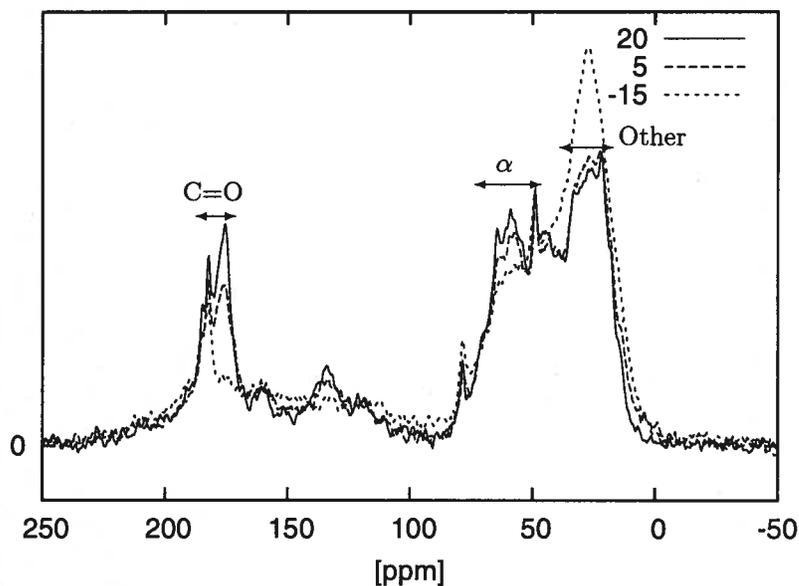


Figure 5.8: Direct Polarization results for hagfish IFs at varying temperature. Temp = 20°C, 5°C, -15°C.

shift anisotropy is not as large for these nuclei and therefore broadening due to this interaction will not be as prominent as with the C=O carbons. Furthermore, whereas the carbonyl carbon isotropic values are quite close together for the various residues in the chain, most in a range of ~ 10 ppm, the α - and side chain carbons span a larger range, closer to 40 ppm. The large number of these carbons, all with slightly different chemical environments, combine and overlap to give a large broad conglomeration of peaks. In general, the α -carbons appear to behave similar to the carbonyl carbons, the peak height dropping with the temperature. This may also be a result of sharper peaks broadening, but because of the large range of overlapping peaks and relatively small chemical shift anisotropy, it just appears that the size of the peak is decreasing. The side chain carbons, mainly β -carbons, appear to exhibit the reverse effect; the peak becoming larger as temperature decreases. Since it is unlikely to have the motion increase and result in sharper peaks as the temperature drops, another explanation is required. If the side chain carbons are in a fast motion regime, see Figure 2.11, then the T_1 times will become longer as τ_c gets shorter. It is possible that at higher temperature, the side chain carbons are not being allowed to fully relax due to a very long T_1 . This is a result of a delay between subsequent experiment scans that is not adequately long, and gives rise to a situation in which the perturbing $\frac{\pi}{2}$ pulse does not project all of the magnetization into the transverse plane. The effect is suppression of the signal from these carbons at higher temperatures. As the temperature decreases, so does T_1 allowing full relaxation

to equilibrium magnetization. Incidentally, this effect may also be the cause of the α -carbon peak behaviour, although in the opposite sense. The α -carbon T_1 s will increase with the decrease in temperature, becoming too long to allow full relaxation at lower temperatures.

The loss of mobility can be characterized well in the carbonyl region, since the broad and sharp portion of the signal can be easily separated. The integral of the sharp peak was plotted against the experimental temperature to give an impression of the temperature dependence. The results, shown in Figure 5.9, indicate that the mobility of the carbonyl carbons is largely temperature dependent and is reduced as the temperature decreases.

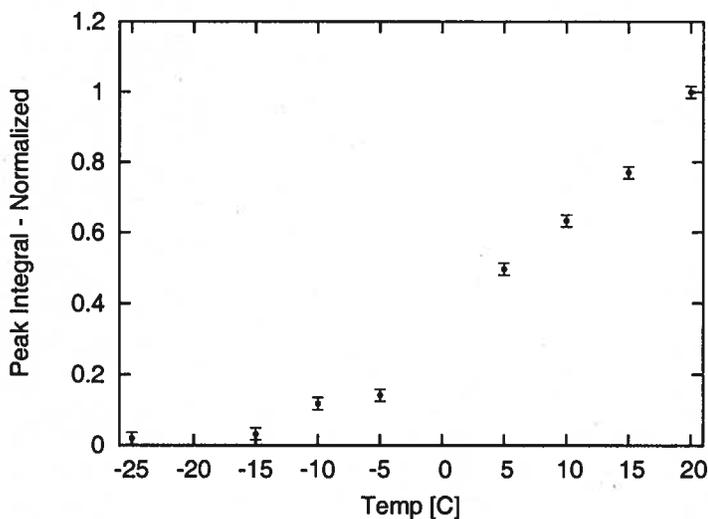


Figure 5.9: Temperature dependence of mobile C=O peak in DP spectra of hagfish IFs.

The DP and CP results give some indication that the protein backbone in hagfish IFs is divided into regions of higher mobility and regions of greater stiffness. This notion of varying degrees of flexibility supports hypotheses of previous work [13, 14] regarding the structure of the hagfish IF.

5.3 Nuclear Relaxation in Hagfish IF Proteins

In an effort to further characterize the structural composition of the IFs, relaxation experiments were implemented. The pulse sequence used was developed by Torchia [35], and incorporates ^{13}C signal enhancement via cross-polarization. A schematic of the pulse sequence is shown in Figure 5.10. The first portion of the sequence is the same as the simple cross-polarization experiment employed in Section 5.2, resulting in the ^{13}C magnetization on the $-y$ -axis. The first ^{13}C $\frac{\pi}{2}$ pulse rotates the magnetization into the $-z$ -axis, where it is allowed to relax back along the longitudinal axis due to the influence of the static field \mathbf{B}_0 for a time, τ . Next, the final $\frac{\pi}{2}$ pulse returns the

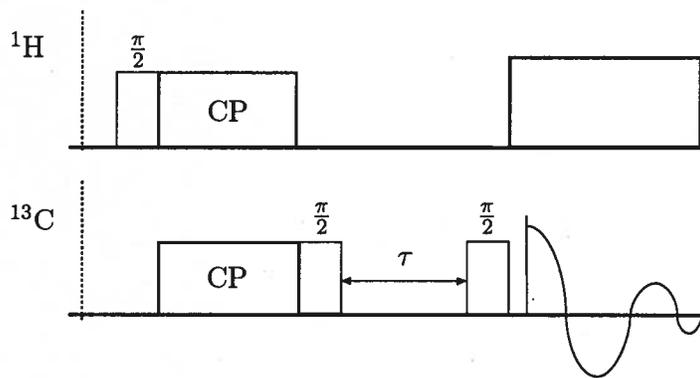


Figure 5.10: T_1 inversion recovery experiment employing cross-polarization pulse to study longitudinal relaxation of the ^{13}C nuclei. Note the $\frac{\pi}{2}$ pulses are generally not the same duration on both channels.

remaining longitudinal magnetization into the xy -plane where it provides the signal for acquisition. The pulse sequence is run many times with various delays, τ . The resulting spectra can be analysed and the intensity of the peaks is related to the delay time, decaying exponentially with a dependence on T_1 :

$$I(\tau) = I_0 e^{-\frac{\tau}{T_1}}, \quad (5.1)$$

I_0 represents the intensity of the initial measure (i.e. that with the shortest delay, τ).

Experiments were performed on a Varian 400MHz (9.3T) spectrometer in a MAS probe spinning at 6600Hz. Results on a hydrated sample of hagfish IFs are displayed for various temperatures in Table 5.3. It should be noted that since these results were obtained using a CP T_1 experiment, only signal from rigid carbons was received. Carbons in rapid motion environments will not contribute. The side chain carbons may experience fast, confined motions that still allow CP, but it is assumed that the other results refer only to the rigid components of the IFs. It should be noted that the broad carbonyl peak, from carbons in rigid structures, has become a sharp peak with spinning sidebands as a result of MAS (see Section 3.2).

The regions analysed were the carbonyl peak, two aromatic peaks believed to be tyrosine ϕ - and δ -carbons, the α -carbons and other side chain carbons (β etc.). These regions are identified on the 20°C spectrum with the shortest delay, τ (thereby providing the highest signal intensity), in Figure 5.11. Due to the poor signal to noise ratio (SNR) of the aromatic carbons, the T_1 values are only to be considered as order of magnitude estimates. The carbonyl spinning sidebands ($\text{C}=\text{O}_{SS}$) were included in the calculation of the carbonyl T_1 s. The carbonyl carbons for the most part fit a single exponential (Equation 5.1) well, see Figure 5.12a. However, this is not the case for the α -carbons

T (°C)	Carbonyl		Y ϕ		T ₁ (s)		Alpha		Other	
	%		%		%		%		%	
20	100	9.3 ± 0.6	100	4.5 ± 3.9	100	1.8 ± 0.4	74	5.8 ± 0.9	69	0.4 ± 0.1
							26	0.5 ± 0.2	31	2.9 ± 0.9
5	100	11.4 ± 0.5	100	5.3 ± 2.9	100	3.6 ± 0.8	77	7.2 ± 0.7	70	0.4 ± 0.1
							23	0.5 ± 0.2	30	3.6 ± 0.8
-5	100	11.1 ± 0.5	100	5.1 ± 1.8	100	4.5 ± 0.8	73	8.2 ± 0.9	64	0.4 ± 0.1
							27	0.9 ± 0.2	36	3.0 ± 0.5
-20	91	14.2 ± 0.6	100	5.0 ± 2.5	71	8.6 ± 2.1	77	8.7 ± 0.5	66	0.4 ± 0.1
	9	0.5 ± 0.2			29	0.4 ± 0.2	23	0.5 ± 0.1	34	3.8 ± 0.7

Table 5.3: T_1 relaxation times for hagfish IFs. A single number indicates a one component exponential fit; two numbers represent a two component fit.

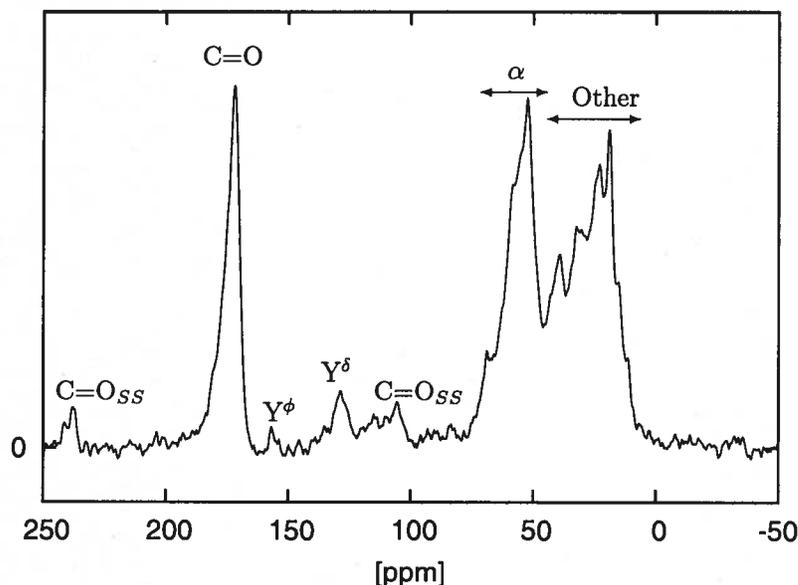


Figure 5.11: Spectrum from T_1 inversion recovery experiment with $\tau = 1\text{ms}$, Temp = 20°C .

and other side chain carbons. For the spins in these positions, the spin-lattice relaxation appears to be caused by mechanisms which operate at different rates. A model with a single exponential is not sufficient and results in a fit with non-random residuals, a more general model is needed:

$$I(\tau) = I_0^A e^{-\frac{\tau}{T_1^A}} + I_0^B e^{-\frac{\tau}{T_1^B}}. \quad (5.2)$$

Equation 5.2 accounts for a situation in which not all spins are relaxing at the same rate. If there were simply two different components, the contribution to the total intensity of the signal from the A component, I_0^A , would represent the fraction of the total population of spins relaxing at a rate $\frac{1}{T_1^A}$. The remainder of the spins would be represented by signal I_0^B and relaxing at $\frac{1}{T_1^B}$. Unfortunately, it is not possible to discern how many different rates may be present. A two-component model

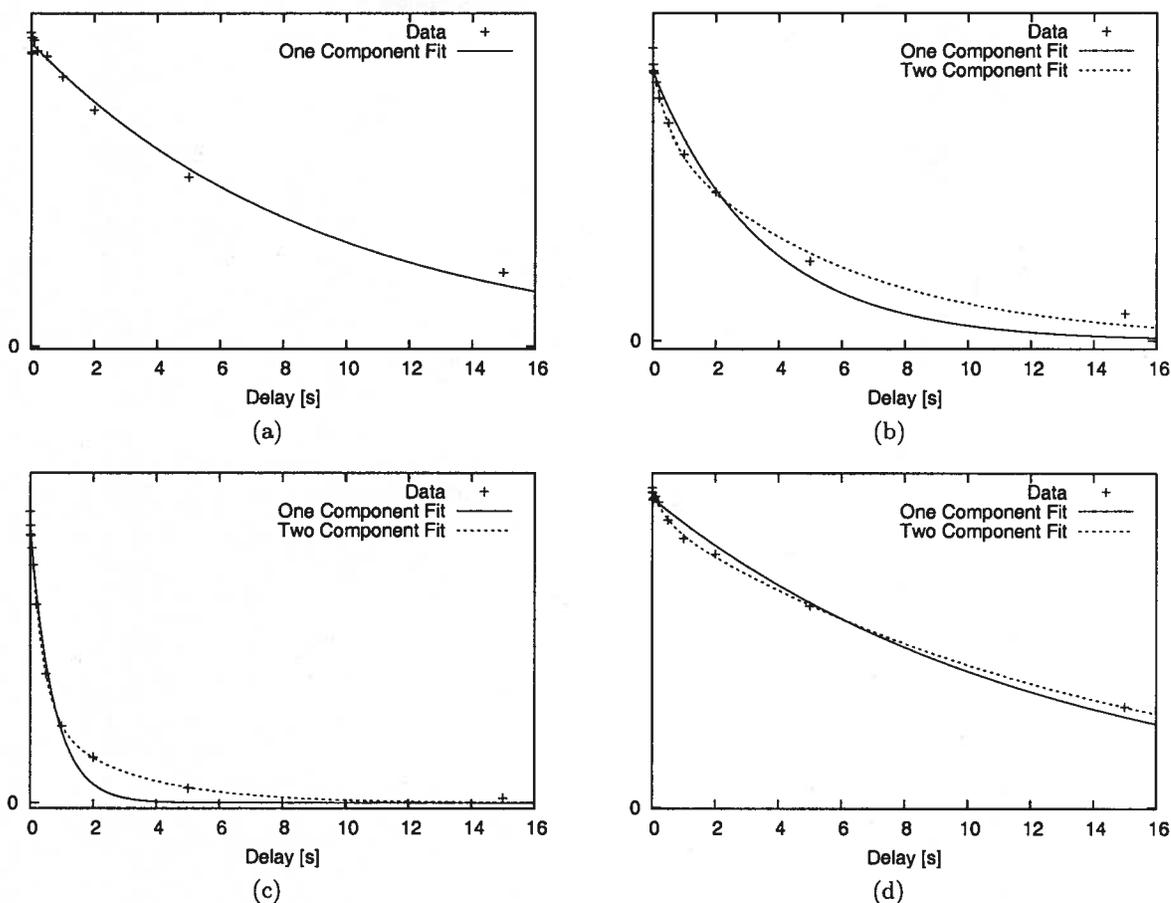


Figure 5.12: T_1 experimental data for (a) carbonyl, (b) α -carbons, (c) other side chain carbons at Temp = 20°C; (d) carbonyl data requires a two component fit at Temp = -20°C.

may fit the results of a wide distribution of relaxation rates [34]. However, due to the nature of exponential functions, the two component fit can provide a reasonable estimate of the breadth of the distribution of relaxation rates. Furthermore, the prefactors in Equation 5.2 will represent the relative amounts of spins relaxing at the respective rates. If the prefactor for the shorter T_1 is larger, it means that more of the spins are relaxing at faster rates, but no further information should be inferred regarding the nature of the distribution. The percentage of the signal relaxing at each rate is displayed before the relaxation time in Table 5.3. At low temperatures (-20°C), a single component model begins to fail for the carbonyl signal (Figure 5.12d). This may be a symptom of colder temperatures further freezing out of motion in the backbone. It is possible that at this low temperature, the cold temperatures have restricted the mobile portions of the backbone to a point where they appear in the CP spectrum.

The data suggests that the rigid parts of the backbone are not homogeneous. There is a division of the rates of signal relaxation indicating that there are at least two different regions.

For simplicity, the components can be considered limit guidelines for the possible distribution of relaxation times; the shorter relaxation times will be referred to as the *fast* limit component and the longer times the *slow* limit component.

The tyrosine aromatic carbon T_1 s exhibited the largest uncertainties as the information available in the signal was lost in noise. This is due to the minimal occurrence of tyrosine in the hagfish IFs (3.85%) and compounded by the low natural abundance of ^{13}C nuclei. In many cases, there was no apparent indication that a two component model would fit the data any better than the single component model (see Figure 5.13). The relative uncertainties in relaxation times ranged from roughly 30-90%.

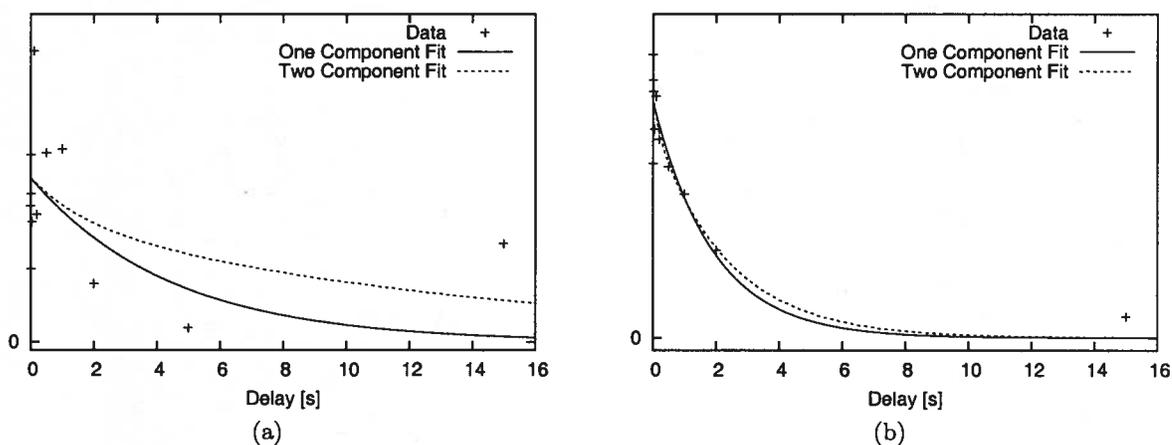


Figure 5.13: The Tyr ϕ -peak T_1 results for Temp = 20°C experiment.

The carbonyl carbon relaxation times are not dramatically affected by the temperature. This result is concordant with the notion that the peak is associated with only carbonyl carbons in the rigid components. The remaining results show that the rigid component signals are characterized by at least two, and possibly more, different relaxation rates for both backbone α -carbons and the side chain carbons. The distribution of relaxation rates suggests a distribution of different regimes in the rigid environments in the hagfish IF. These different regimes must possess varying degrees of flexibility to allow the variation in relaxation rates. For the backbone carbons, the temperature dependence of T_1 indicates that the sample is in the slow regime to the right of the T_1 minimum, (see Section 2.8). In this regime, a shorter T_1 value implies that the spins are rotating or vibrating faster, thus are incorporated in a part of the structure with greater flexibility. One interesting feature is that the α -carbon signal portrays a situation in which most of the spins are relaxing at a slower rate, whereas the side chain carbons have the minority of the population with a longer

T_1 . This is intuitive since the β - and other side chain carbons generally do have a greater degree of flexibility than those involved in the peptide bonding on the protein backbone. These T_1 results can be accounted for in the model proposed by Fudge et al [13, 14].

T_1 relaxation measurements proved difficult for the hagfish IFs due to long relaxation times and low sensitivity without cross-polarization. In an attempt to characterize the structure in mobile regions, the direct polarization results from the previous section (5.2) were further analysed. These results indirectly allowed estimates of correlation times for the mobile components, using techniques discussed in Section 2.8. Furthermore, because the single DP spectra was relatively quick to perform, data at a greater number of temperatures could be extracted. The T_1 experiments performed required a long time to complete (150hrs) and due to time constraints were only performed at the temperatures in Table 5.3.

The linewidth of a signal may be used to characterize the degree of motion of the nuclei associated with it. This is due to the fact that some of the spread in spin precession frequency stems from the random reorientation of spins in their immediate chemical environment. This reorientation causes relaxation as discussed in Section 2.8. A simulation of the line broadening caused by the modulation of the chemical shift anisotropy provides an expected lineshape based on a correlation time, τ_c . The contribution to the lineshape from a specific chemical shift frequency depends on the anisotropy Δ and the asymmetry, η :

$$\Delta\omega_{cs}(\theta, \phi) = \frac{\Delta}{2} (3 \cos^2 \theta - 1 + \eta \sin^2 \theta \cos 2\phi), \quad (5.3)$$

the angles θ and ϕ describe the orientation of the principle axis system of the nucleus with respect to the static field. These angles are assumed to change randomly and instantaneously at intervals of an exponentially distributed random duration T_D , on the order of τ_c ,

$$T_D = -\tau_c \log z, \quad (5.4)$$

where z is a random number with $0 \leq z \leq 1$. The spin is assumed to contribute to the signal at ω_{cs} for T_D at which point there is a jump to the next random orientation (θ, ϕ) . The phase accumulated over these time periods is numerically integrated using a “brute force” calculation. From the phase the FID signal is simulated for multiple initial orientations, and Fourier transformed to provide an estimate of the lineshape for the carbonyl carbons.

The simulated signal is processed slightly to account for field inhomogeneity. Broadening will occur due to the fact that identical spins may experience a slightly different field based on their position within the superconducting magnet, as it is not an infinite cylinder. The amount of broadening was estimated using the citrate peaks in the DP spectra. These peaks represent signal from an inviscid solution, which would be very sharp in a completely homogeneous field due to fast motional average in the liquid. The exponential broadening factor, BR, applied to the simulation was chosen to match the observed citrate linewidth. The broadening is implemented by a decaying exponential envelope on the signal S_n at point n in the FID:

$$S_n = S_n e^{-nBR}. \quad (5.5)$$

The source code for the simulation can be found in Appendix A.

The simulations were qualitatively compared to the direct polarization results from the experiments in Section 5.2, see Figure 5.14. The simulated lineshape was overlaid on the sharp carbonyl

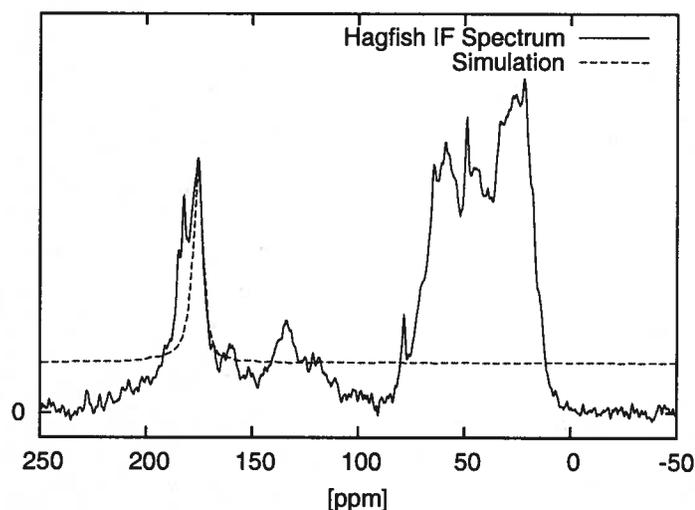


Figure 5.14: A simulated lineshape overlaid on a DP spectrum collected at Temp = 20°C. A correlation time of $\tau_c = 0.06\mu\text{s}$ achieves a reasonable fit.

peak of the DP spectra, and the simulation correlation time was adjusted until the simulated peak seemed to match the observed carbonyl peak. The broad part of the spectrum was ignored and the simulated peak was taken to stand atop the underlying broad signal due to rigid structure. The results of the simulations are presented in Table 5.4, and displayed graphically in Figure 5.15. It should be noted that this is not a direct measure or calculation of the correlation time, the simulation merely provides an order of magnitude estimation of τ_c . The simulation results were fit

Temp (°C)	τ_c (μ s)
20	0.06
15	0.08
10	0.10
5	0.15
-5	0.20
-10	0.30
-15	0.80
-25	1.00

Table 5.4: Simulated τ_c results for the mobile carbonyl carbons.

to the well known Arrhenius equation:

$$\ln(\tau_c) = -\frac{E_a}{RT}, \quad (5.6)$$

where R is the gas constant (8.3145 J/mol K), T is the temperature in Kelvin and E_a denotes the activation energy in J/mol. The activation energy in this case refers to a characteristic energy barrier between states that describe the motion of the protein backbone. For the simulation results, the value of E_a was found to be 40.3 kJ/mol.

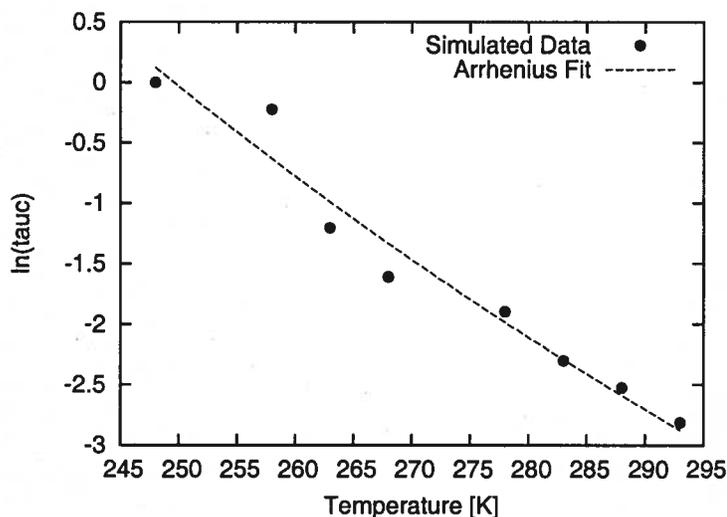


Figure 5.15: Simulated τ_c temperature dependence results displaying Arrhenius fit with $E_a = 40.3$ kJ/mol.

NMR studies of similar proteins have found correlation times ranging from ones to hundreds of nanoseconds. For example, a study of hydrated spider silk [38] reported rapidly moving backbone correlation times of $\tau_c \approx 10$ ns. Hydrated elastin filaments exhibited correlation times on the order of tens of ns [28]. The simulation results for the hagfish IFs are of a similar order of magnitude

for temperatures above 0°C, but rapidly become much larger as the temperature gets colder, see Figure 5.15. The results of the simulation provide a reasonable estimate of correlation times and correctly display an increase in τ_c as temperature is reduced. The simulation follows an Arrhenius law behaviour with a characteristic energy of 40.3 kJ/mol.

A feature of the hagfish biopolymer is that the flexible terminal ends contain a higher percentage of glycine than the central rod domain; approximately 30% in the ends compared to less than 2% in the central rod. It would be interesting to investigate the mobility of the glycine C α s compared to other α -carbons, as the glycines primarily represent the terminal regions. Unfortunately, the resolution in the spectrum was not adequate to facilitate such an investigation.

6 Resilin: A Resilient Rubber-Like Protein

Resilin provides a means of efficient energy storage in many insects [11]. It belongs to a family of elastic proteins including elastin and spider silks, and combines low stiffness and high strain. Resilin is normally found in parts of the cuticle of insects and is involved in roles from flight [36, 15] to generation of sound [5]. The basic function is mechanical storage of elastic potential energy, usually by means of a deflection, extension, compression or even twisting of an elastic structure containing the resilin protein. The strain is induced by a power muscle and is sometimes held by a latch system until release. Previous studies of naturally occurring resilin protein [9, 37], have established that the material consists of an amorphous network of randomly coiled polypeptide chains with stable chemical cross-links. A resilin-like protein has been tentatively identified in the gene of fruit flies (*Drosophila melanogaster*) [4]. The soluble recombinant protein, expressed in *E. coli* from the gene has been shown to cast into a solid rubber-like material via rapid photochemical crosslinking [10]. The study further shows that the resilience of the crosslinked recombinant resilin exceeds that of synthetic polybutadiene and chlorobutyl rubbers, high performance materials used in tire linings. The mechanical properties, along with the ability to rapidly cast samples with light may make it a prime candidate for *in situ* use in biomedical and industrial applications.

The recombinant resilin (rec-1) was sequenced from the first exon of the *Drosophila* CG15920 gene. The sequence contains 17 repeats of the putative elastic motif GGRPSDSYGAPGGGN. The ubiquitous 4 residue YGAP repeat is believed to be an important component in the resilin sequence, the site at which dityrosine crosslinks are formed. The rec-1 sequence is shown in Figure 6.1.

A second resilin sample was created from a consensus sequence from *Anopheles gambiae*,

MHHHHHPEPPVNSYLPPSDSYGAPGQSGPGRPSDSYGAPGGGNGGRPSDSYGAP
 GQQQQGQQGGYAGKPSDSYGAPGGGDGNGGRPSSSYGAPGGGNGGRPSDTYGAP
 GGGNGGRPSDTYGAPGGGGNGNGGRPSSSYGAPGQQGNGNGGRPSSSYGAPGGGN
 GGRPSDTYGAPGGGNGGRPSDTYGAPGGGNGGRPSSSYGAPGGGNGGRPSDTYGA
 PGGGNGNGSGGRPSSSYGAPGQQGGFGRPSDSYGAPGQNQKPSDSYGAPGSGNG
 NGRPSSSYGAPGSGPGRPSDSYGPPASG

Figure 6.1: Amino acid sequence of recombinant resilin, rec-1.

Residue	rec-1 (%)	an-16 (%)
Ala	6.1	18.2
Asp	4.2	0.0
Glu	0.3	0.0
Phe	0.3	0.0
Gly	34.8	9.1
His	1.9	0.0
Lys	0.6	0.0
Met	0.3	0.0
Asn	6.1	0.0
Pro	13.9	18.2
Gln	3.9	18.2
Arg	4.8	0.0
Ser	13.9	18.2
Thr	1.6	9.1
Val	0.3	0.0
Tyr	6.5	9.1
Leu	0.3	0.0

Table 6.1: Amino Acid content of resilin samples. Residues not listed are not present in either sequence.

AQTPSSQYGAP, repeated 16 times. This sequence (an-16) has a smaller separation of YGAP repeats, see Figure 6.2. The relative abundances of the amino acid residues in each sequence are provided in Table 6.1.

AQTPSSQYGAPAQTPSSQYGAPAQTPSSQYGAPAQTPSSQYGAP
 AQTPSSQYGAPAQTPSSQYGAPAQTPSSQYGAPAQTPSSQYGAP
 AQTPSSQYGAPAQTPSSQYGAPAQTPSSQYGAPAQTPSSQYGAP
 AQTPSSQYGAPAQTPSSQYGAPAQTPSSQYGAPAQTPSSQYGAP

Figure 6.2: Amino acid sequence of an-16 resilin.

The crosslinking in natural resilin occurs between adjacent tyrosine residues forming di- and trityrosine adducts [2]. In the recombinant resilin, roughly 21% of the tyrosines are converted to dityrosines (see Figure 6.3) under irradiation mediated by $\text{Ru}(\text{bpy})_3^{2+}$.

Samples of rec1-resilin and an-16 sequence were analysed on a Varian 400MHz with an HX probe and a homemade 5mm horizontal coil. DP spectra are shown in Figure 6.4.

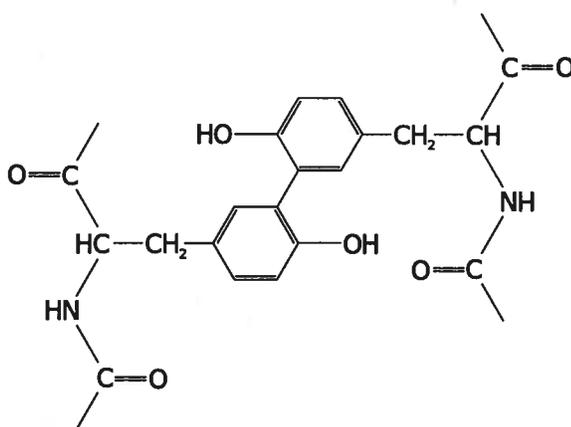


Figure 6.3: Dityrosine complex formed via photo-crosslinking. The crosslink site is shown here at the ϵ -carbon but may also occur at the δ -carbon.

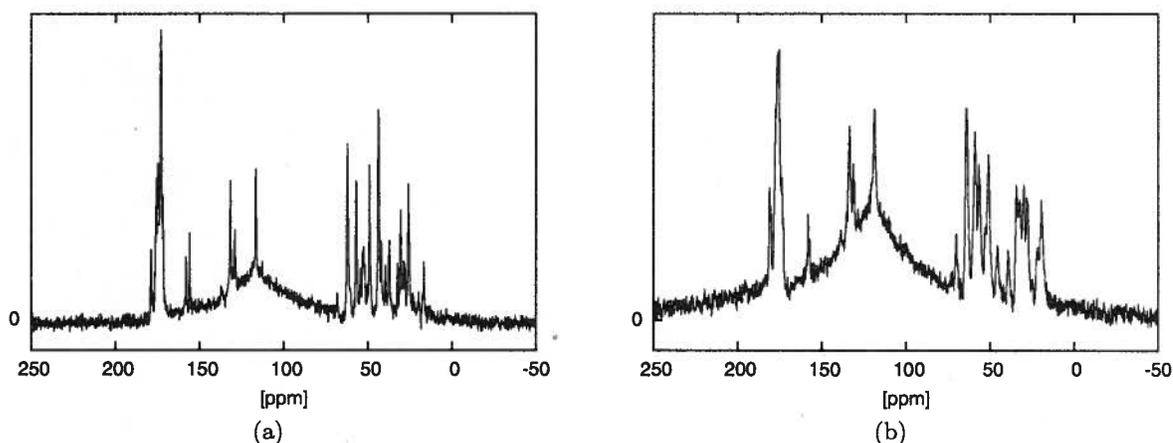


Figure 6.4: DP experimental spectra (a) rec-1 resilin (b) an-16 resilin. The broad peak centered at 120ppm is due to probe background.

6.1 Chemical Shift Assignment of Recombinant Resilins

In contrast to the hagfish IF biopolymer, the resilin ^{13}C spectra are well resolved and allow identification of different carbons in the amino acid residues. The chemical shifts for backbone and β -carbons of the residues in the amino acid sequence Figure 6.1 were predicted from a database of shifts. The prediction calculations were based on amino acid sequence and secondary structure, with corrections for nearest neighbour effects [39]. A value for each residue carbonyl, α - and β -carbon was calculated for each of 3 different conformations; α -helix, β -sheet and a structurally dynamic model denoted “random coil”. The predicted values were then compared to the observed chemical shifts and peaks were assigned based on agreement with calculated values as well as relative size of the peak corresponding to residue content in the chain (Table 6.1). A high degree of confidence was possible for some peaks, while other peak assignments could not be made positively.

Other chemical shifts (i.e. for γ -, δ -carbons etc.) were referenced from the Biological Magnetic Resonance Data Bank [31]. The assigned spectra are presented in Figure 6.5.

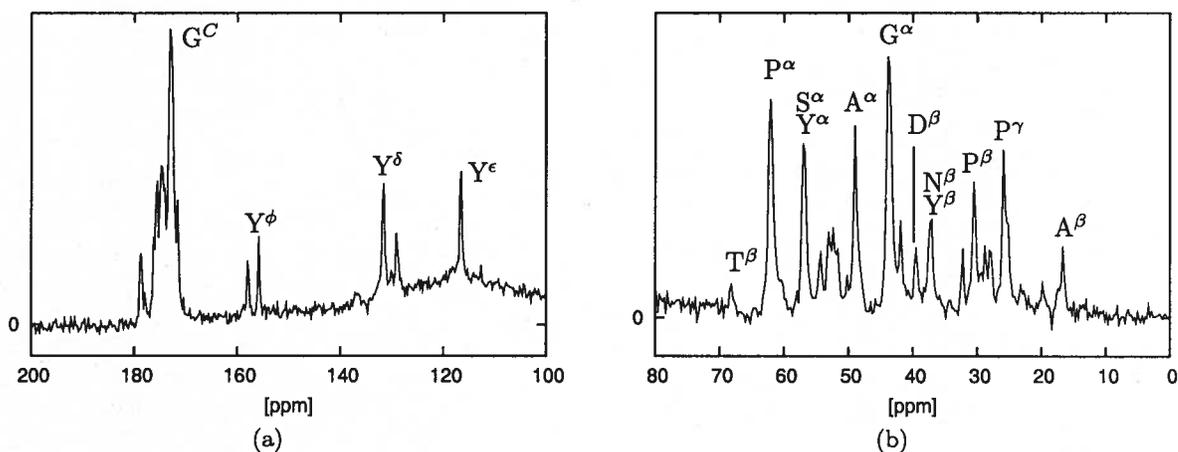


Figure 6.5: rec-1 resilin DP spectra peak assignment (a) carbonyl and aromatic region (b) α - and β -carbon region.

The peaks without labels were more difficult to assign owing to poor agreement with predicted values or overlap with other possible species. In Figure 6.5a, only the Gly carbonyl peak was positively assigned. It was easy to assign because of the relative high abundance of Gly in the sequence. Unfortunately, since the carbonyl peaks are all centered around a short range (~ 170 to 180 ppm) it was difficult to assign other peaks with any degree of confidence. The carbonyl peak with the highest shift at ~ 180 ppm could be Ala or Pro (or a combination of both). The other carbonyl peaks, having lower chemical shift predictions are not as probable.

Referring to Figure 6.5b, A ^{β} has a small shoulder on the downfield side. There is also a peak just downfield of A ^{β} at ~ 19.5 ppm. It is most likely these signals are alanine peaks as well, as no other β species appears lower than 20 ppm. There are three different occurrences of Ala in the sequence, GAP, YAG and PAS, appearing in a ratio 17:1:1. These varied sequences may give rise to slightly different shifts. Furthermore, Ala is one of the residues close to the putative cross-link site YGAP. It is possible that these factors provide for a second Ala peak at ~ 20 ppm.

There is suggestion of a peak just upfield of the P ^{γ} . This peak is not likely from Ala, since the shift is nearly 6 ppm higher than the assigned A ^{β} . It is possible that the peak is noise, however it does appear to be significant compared to the surrounding noise level. It is most likely that it is a second P ^{γ} peak. Pro appears in many different sets of neighbouring residues (see Figure 6.1) and is one of the residues in the YGAP repeat. Furthermore, the peptide bond *cis* conformation is more likely when the subsequent residue is Pro (1:4 probability of being *cis* instead of *trans* compared to

1:1000 [6]). These facts could possibly explain an extra peak from the Pro γ -carbon. The P^γ peak also has a shoulder on its upfield side; probably a P^γ conformer, as the only residues with remotely nearby chemical shifts are Gln and His β -carbons, which do not constitute an appropriate fraction of the protein, see Table 6.1.

There are two definite peaks downfield of the P^γ peak, which are presumed to be Gln and Arg β s, respectively, based on predicted chemical shift values and relative content. Between the Arg β peak and the P^β there is indication of a small peak which could be either a β -carbon from a Pro, Arg or Gln in a different residue sequence or combinations of the three. The peak just downfield of the P^β at ~ 32 ppm has many possible assignments. It could arise from the β -carbon of another conformer of Pro, the closest peak with a substantial number of residues in the chain. Alternatively, the β -carbons for Val, Met and Lys are predicted at this chemical shift, however the combined percentage of these residues in the chain is just 1.2%, which would not result in such a large peak.

In the α -carbon region, the peak between G^α and D^β is likely due to the α -carbon of Gly in an alternate sequence as there are no other nearby predicted shifts. The Gly α -carbon had two predicted shifts based on the adjacent residues. The sequence SGP accounted for roughly 3% of the Gly and resulted in a predicted shift of 43.1 ppm, while other sequences with Gly resulted in a 45.1 ppm predicted shift. These shifts are separated by 2 ppm, agreeing well with the separation observed in the spectrum. Much like the β -carbon, the A^α peak appears to have a small shoulder on the left side, possibly from an Ala in an alternate sequence. The other unknown peaks in the α region are in a conglomeration between the large A^α and S/Y^α . These peaks result from the α -carbons from Asn, Arg, Asp, Gln, Leu and His. The Leu and His contributions are small due to low concentrations in the protein. Based on predicted shifts, the 4 peaks from right to left are most likely Asn, Arg, Asp and Gln, respectively.

The an-16 sample had fewer peaks to assign thanks to the simpler amino sequence, but the resolution was not as fine as with the rec-1 resilin. Similar to the rec-1 carbonyl peaks, G^C was the only positively identified species, see Figure 6.6a. Predicted chemical shift, relative peak size and comparison with the rec-1 spectrum allowed identification of the Gly peak. Again, the peak at ~ 180 ppm is believed to be an Ala/Pro carbonyl.

Referring to Figure 6.6b the A^β has a shoulder that may be attributed to alternate sequences, PAQ and GAP, or the inclusion of the Ala in the YGAP repeat. The peak between A^β and Q^β is likely the Pro γ -carbon, as the chemical shift is in good agreement with reported values and lines up

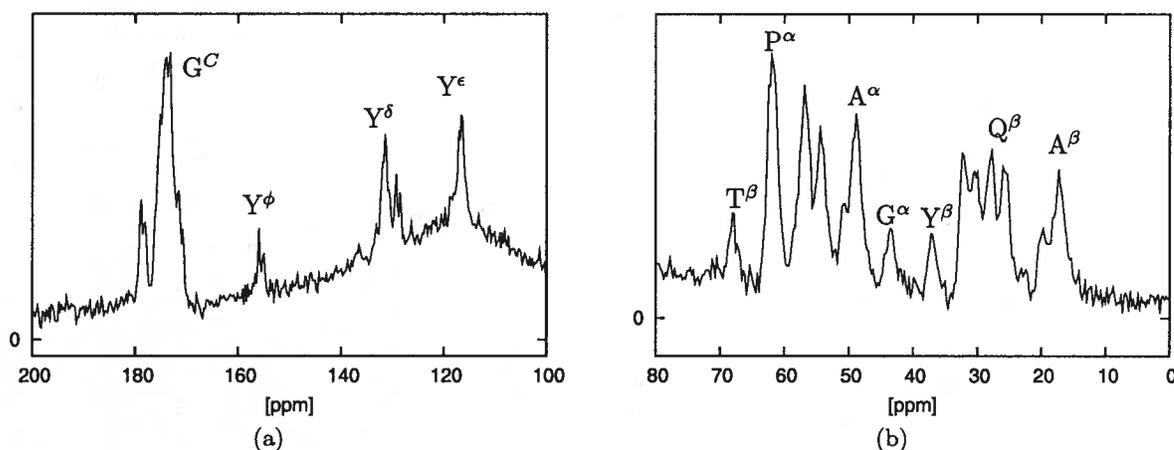


Figure 6.6: an-16 resilin DP spectra peak assignment (a) carbonyl and aromatic region (b) α - and β -carbon region.

well with the P^γ peak in the rec-1 spectrum, see Figure 6.7b. The two peaks between the Y^β and Q^β are probably due to Pro β -carbons in slightly different environments. There are two sequences involving Pro; TPS and APA, as well as the aforementioned *cis/trans* conformers, which may lead to different chemical shifts.

In the α -carbon region of an-16, there are two large peaks that are between the A^α and P^α peaks. The downfield peak is likely the combination of Tyr and Ser α -carbon signals, which aligns well in Figure 6.7b. The upfield peak must originate from the Gln α -carbons. The A^α peak appears to have a small downfield shoulder, possibly also attributed to alternate sequence environments.

The aromatic regions should provide some indication of the effects of the dityrosine bonds, as the aromatic carbons of the tyrosine are the closest to the bond location, which can occur at any one of the Tyr ϵ - or δ -carbons. However, Figures 6.5a and 6.6a do not exhibit any compelling evidence of dityrosine bonds in the sample. In the rec-1 sample, Figure 6.5a, the signal from the Tyr ϵ -carbons (the furthest upfield peak) appear in a single sharp peak. There is no apparent splitting or change in the isotropic chemical shift due to the presence of dityrosine bonding. The peaks in the center of the aromatic region are believed to be the δ -carbons (around 132 ppm) and the γ -carbon (around 129 ppm) signals. The apparent 2:1 ratio of the δ : γ peak sizes correctly accounts for the abundance of these species. There is also suggestion of two other small peaks, upfield from the δ and between the δ and γ peaks. These small signals may result from His ϵ -carbon (138 ppm) and γ -carbon (131 ppm), respectively. However, these peaks are still present in the an-16 spectrum, which does not contain any His, see Figure 6.7a. It is possible that these extra peaks arise from a changed chemical environment due to dityrosine bonding, or from alternate sequence

environments. It is strange however that a similar effect is not seen in the ϵ -carbon signal. The δ - and γ -carbon signals are very similar between the rec-1 and an-16 spectra.

In the rec-1 spectrum, signal from the Tyr ϕ -carbon (furthest downfield, next to the carbonyl region) appears as a single peak. The peak downfield of the ϕ peak in the rec-1 spectrum could possibly be the Arg ζ -carbon peak, occurring at 159 ppm with an appropriate magnitude relative to the Y^ϕ peak. The Arg ζ peak is absent from the an-16 spectrum, as expected. The Y^ϕ peak has a slight downfield shoulder that may be attributed to either dityrosine bonding or alternate sequences around the tyrosine.

The an-16 spectrum is more informative to look at due to the fact that the aromatic signals are provided solely by tyrosine. Referring to Figure 6.7a, the ϵ -carbon peak has a slight shoulder on it and the δ/γ -carbon peaks are broader but there are two peaks in close to the same positions as the rec-1 Tyr δ/γ peaks. The excess broadening may be a result of poorer SNR in the an-16 spectrum.

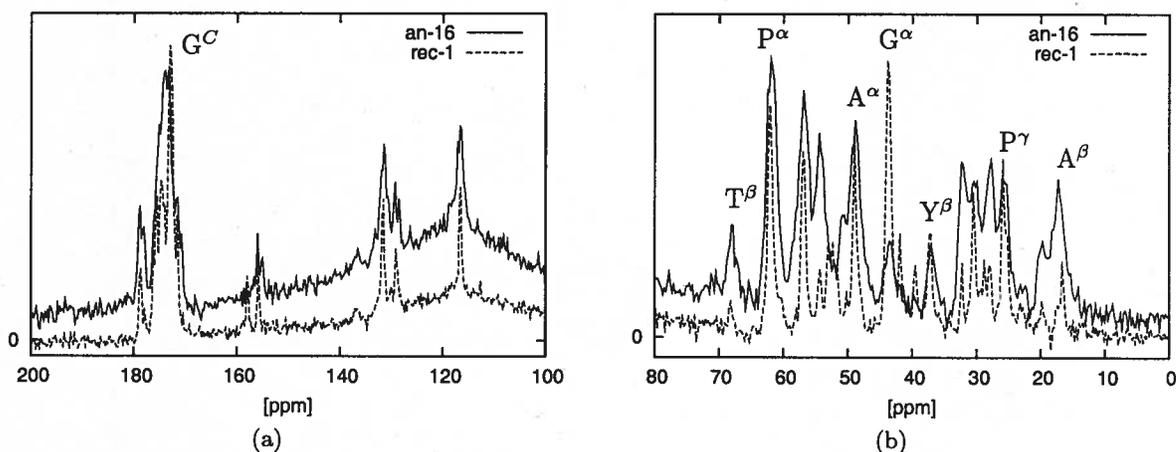


Figure 6.7: rec1 and an-16 DP spectra overlay (a) carbonyl and aromatic region (b) α - and β -carbon region.

6.2 Secondary Structure of Recombinant Resilin

The confidently assigned peaks were used to provide a quantitative measure of the possible secondary structure of the proteins. The difference between the observed and predicted chemical shifts,

$$\Delta\delta_i = \delta_i^{pre} - \delta_i^{obs}, \quad (6.1)$$

was calculated for each possible conformation. The deviations d_i , defined:

$$d_i = \Delta\delta_i - \langle\Delta\delta_i\rangle, \quad (6.2)$$

were plotted along the chemical shift scale for the residues under evaluation. $\langle\Delta\delta_i\rangle$ represents the average of all the differences for the definite peaks for a particular conformation. Figures 6.8a and 6.8b show the deviations from predicted chemical shifts for each of the 3 conformations. The deviations were used instead of the absolute differences $\Delta\delta_i$ as absolute referencing was not possible. This is simply because a chemical shift standard could not be inserted into the sample holder with the resilin samples.

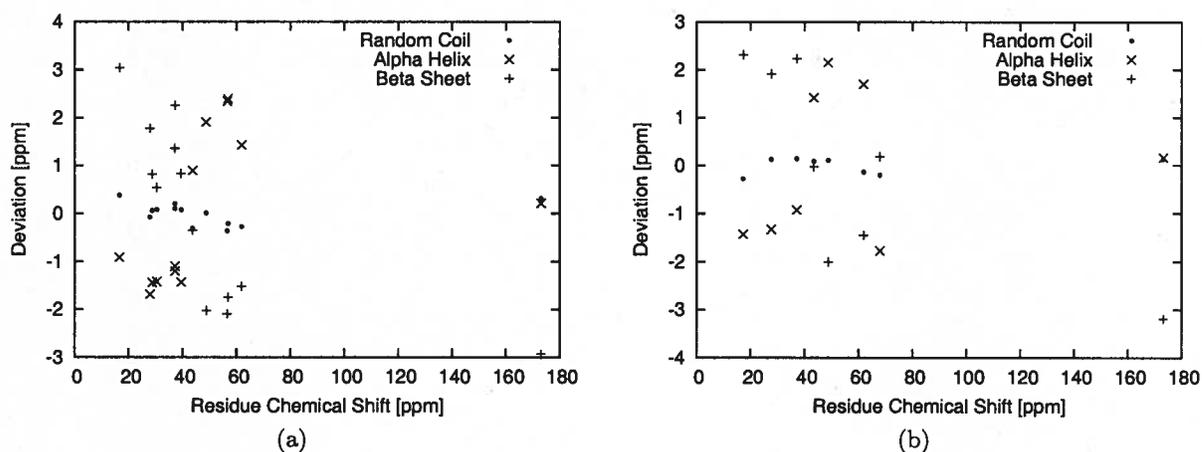


Figure 6.8: Deviation from observed chemical shift (a) rec-1 resilin, $N = 13$ (b) an-16 resilin, $N = 8$.

The standard deviation σ for each conformation was calculated from the N high confidence peak results:

$$\sigma = \sqrt{\frac{1}{N} \sum_i (\Delta\delta_i - \langle\Delta\delta_i\rangle)^2}. \quad (6.3)$$

The standard deviation results are summarized in Table 6.2. The standard deviation for the Random Coil configuration is shown to be almost an order of magnitude lower than either the α -helix or β -sheet, suggesting that the proteins do not form stable secondary structure, but rather the protein backbone averages rapidly over many different conformations. These results suggest that the recombinant resilin samples exhibit random coil structure, similar to the conformations discovered by Weis-Fogh et al [37].

Conformation	rec-1 Variance [ppm]	an16 Variance [ppm]
Random Coil	0.23	0.16
Alpha Helix	1.52	1.47
Beta Sheet	1.83	1.95

Table 6.2: Simulated chemical shift variance.

6.3 Relaxation Measurements

The T_1 relaxation times were also measured in the resilin samples. The results further indicate a material with a high degree of molecular mobility. The experiments were carried out on a Varian 400MHz spectrometer with a HX probe and a homemade 0.5cm (inner diameter) copper coil. The pulse sequence used employed selective phase cycling to eliminate the background artifacts (Figure 6.9). A side effect of the background elimination combined with delay times was the loss of some SNR, as seen in Figure 6.10.

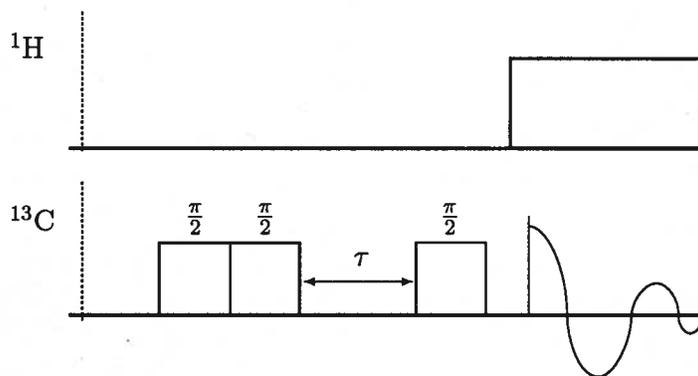


Figure 6.9: T_1 experiment employing background suppression via phase modulation.

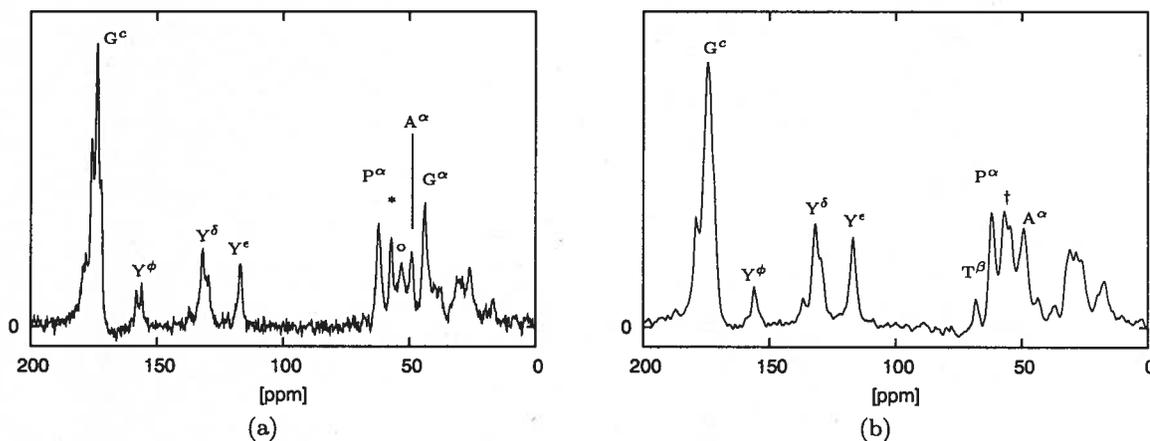


Figure 6.10: T_1 spectra with $\tau = 0.01s$ and Temp = $4^\circ C$ a) rec-1 resilin b) an-16 resilin.

Peak	Species	Temp [°C]	T_1 [s]	τ_c [ns]	
				Fast	Slow
C=O	all C=O	4	1.72 ± 0.05	0.96	2.6
		-8	1.87 ± 0.08	0.82	3.7
Y $^\delta$	Tyr δ	4	0.47 ± 0.02	0.106	8.2
		-8	0.48 ± 0.05	0.105	8.3
Y $^\epsilon$	Tyr ϵ	4	0.37 ± 0.04	0.14	6.4
		-8	0.35 ± 0.02	0.15	6.0
P $^\alpha$	Pro α	4	0.24 ± 0.02	0.28	3.7
		-8	0.25 ± 0.02	0.25	3.9
*	Tyr, Ser α	4	0.31 ± 0.01	0.18	5.2
		-8	0.38 ± 0.02	0.14	6.6
o	Asn, Arg Asp, Gln α	4	0.30 ± 0.01	0.19	4.9
		-8	0.38 ± 0.03	0.14	6.5
A $^\alpha$	Ala α	4	0.28 ± 0.02	0.20	4.7
		-8	0.30 ± 0.02	0.18	5.0
G $^\alpha$	Gly α	4	0.21 ± 0.01	0.12	7.4
		-8	0.22 ± 0.01	0.11	7.8

Table 6.3: rec-1 relaxation times and corresponding correlation times τ_c . Temperature dependence trend indicates that backbone and aromatic carbons are in the slow regime.

The identifiable peaks, those denoted in Figure 6.10, were used in the T_1 calculations. All of the peaks fit a single exponential decay well, though some of the smaller peaks had larger uncertainties due to poorer SNR. The T_1 values were subsequently used to determine the correlation times τ_c , using the relations in Section 2.8. Again, the carbonyl carbons were assumed to relax primarily under the chemical shift anisotropy (Equation 2.71). The other carbons were assumed to relax due to dipolar interaction with nearby protons (Equation 2.66). The results are summarized in Tables 6.3 and 6.4. Consideration of the amide proton (NH) in the relaxation of ^{13}C nuclei had negligible effect on the resulting correlation time. The glycine α -carbons were assumed to relax under the influence of the dipolar interaction with 2 nearby protons, instead of just 1. The experiments were performed at 4 and -8°C to identify whether the protein backbone carbons were in the fast or slow regime. The backbone carbons should move slower as temperature is reduced, resulting in a longer τ_c . Thus, it is apparent from Tables 6.3 and 6.4, that all of the nuclei observed are in the slow regime, which is common for large molecules. The fast regime would see the various parts of the molecule tumbling faster at lower temperature.

The β -carbons were excluded from T_1 analysis as the correlation times from the more freely mobile side-chain carbons do not provide much insight into the dynamics of the protein backbone. The backbone carbon correlation times are all quite fast (2-8 ns). This further supports the notion

Peak	Species	Temp [°C]	T_1 [s]	τ_c [ns]	
				Fast	Slow
C=O	All C=O	4	1.81 ± 0.07	0.86	2.9
		-8	2.11 ± 0.05	0.67	3.7
Y^δ	Tyr δ	4	0.46 ± 0.04	0.11	8.0
		-8	0.51 ± 0.04	0.097	9.0
Y^ϵ	Tyr ϵ	4	0.29 ± 0.02	0.20	4.7
		-8	0.31 ± 0.01	0.18	5.2
T^α	Thr α	4	0.17 ± 0.03	0.73	2.0
		-8	0.20 ± 0.03	0.43	2.7
†	Tyr, Ser α	4	0.22 ± 0.02	0.34	3.2
		-8	0.22 ± 0.01	0.26	3.4
Q^α	Gln α	4	0.33 ± 0.01	0.16	5.5
		-8	0.36 ± 0.01	0.15	6.0
A^α	Ala α	4	0.30 ± 0.02	0.19	4.9
		-8	0.35 ± 0.01	0.15	6.0

Table 6.4: an-16 relaxation times and corresponding correlation times τ_c .

that the resilin is composed of a highly flexible protein chain, providing the elastic properties of the material. The correlation times found for the aromatic Y^δ carbons were the longest reported. This is interesting because these are carbons close to, or involved in the dityrosine crosslinks. It is possible that the crosslinks provide extra stability that restricts the motion of these aromatic tyrosine carbons. However, the ϵ -carbons did not prove to have significantly longer τ_c s. It might also be expected that the aromatic carbons would display two-component behaviour due to 20% involvement in dityrosine bonding, but this is not seen either. The correlation times of the backbone carbons compare well between the two samples. One might expect the larger separation of YGAP binding sites would allow for greater flexibility, and therefore faster correlation times in the rec-1 molecule. However this is not the case, the rec-1 backbone correlation times report a slightly slower upper limit; τ_c s from 2-8 ns vs. 2-6 ns. This difference between samples is small and may be partially attributed to experimental error.

7 Incorporation of Metal Ions into Hagfish IF: Preliminary Work

Many proteins bind metal ions in a variety of ways, for a multitude of purposes. It has been demonstrated that formation of IFs is affected by ligand metal concentrations [20]. The cohesive properties of the slime in water have been shown to be dependent on the presence of calcium [12]. It is possible that the IF component properties are dependent on Ca^{2+} , and a calcium binding mechanism has been purported (P. Guerette, private communication). It was suggested

that the hagfish biopolymer may contain 3,4-dihydroxy-L-phenolalanine (L-DOPA), frequently seen in marine bioadhesion proteins [26, 33, 40]. DOPA is converted from the amino acid tyrosine and provides a binding site for divalent metals such as Al^{2+} , Cu^{2+} or Ca^{2+} , see Figure 7.1. In an attempt to support or refute this claim, the ability of the protein to bind Ca^{2+} was investigated.

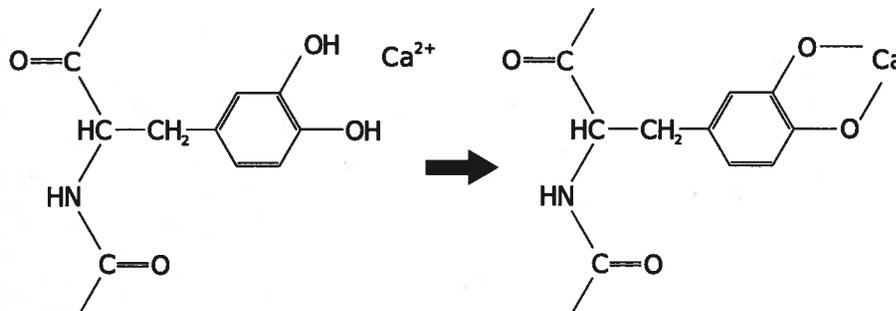


Figure 7.1: 3,4-dihydroxy-L-phenolalanine (L-DOPA). (a) DOPA is much like the amino acid tyrosine but has a second hydroxyl group on the aromatic ring; (b) a divalent metal (in this case Ca^{2+}) may form an ionic bond with the oxygen atoms.

To accomplish this, the hagfish IFs were treated with an ionic solution. The procedure for this is outlined in Section 7.1. The metal ion incorporation, if present, could be characterized by an NMR experiment which is sensitive to nuclear separation. Rotational-echo double-resonance (REDOR) and REDOR-like experiments (REAPDOR [17], SEDOR [22]) are techniques used to measure the dipolar coupling and therefore could determine the internuclear distance and indicate the location of binding. Such an experiment is discussed in Section 7.2. A simulation of the experiment given the purported binding site was also performed for comparison with results. The success of the experiment depends on the ability to maintain the signal for an adequate amount of time, discussed in Section 7.2.1, as well as the ability of the hagfish IFs to actually bind the ions. The latter point is addressed by EDX in Section 7.3.

7.1 Sample Preparation

Samples of the hagfish IF bundles were subjected to a bath in a concentrated solution of Ca^{2+} in an attempt to promote ligand binding. Care was taken to provide a setting in which exposure to pollutant ions was minimal.

The IFs were taken from the citrate buffer and stored at -4°C when not in use. A solution of 10 mM Ca^{2+} was produced by adding 5 mg of CaCO_3 to 2 mL of 0.12 M HCl solution. The solution was then neutralized using 0.12 M NaOH solution, bringing the pH to approximately 8 and the final volume to 5 mL. The concentration of dissolved CaCO_3 is dependent on the solvent pH.

Adding CaCO_3 to a strong acid allowed for improved dissolution of the compound and the release of CO_2 gas, thereby mitigating precipitation of CaCO_3 as the solution was brought back to a more neutral pH. Excess solid CaCO_3 was removed by centrifugation and pipetting. Approximately 2 mg of IFs were added to 5 mL of the 10 mM Ca^{2+} solution and allowed to sit for 24 hrs, periodically agitating the sample to ensure adequate mixing. After exposure, the sample was rinsed by repeatedly pipetting off solution, adding nanopure H_2O to dilute the concentration of Ca^{2+} and centrifuging the new dilute solution. This was done 10-20 times to reduce the free Ca^{2+} ion concentration to less than 10^{-12} times the amount expected to bind assuming all of the tyrosines were converted to DOPAs.

The concentration of Ca^{2+} used was chosen to be a 10-fold excess of the available binding sites in the hagfish IF. Since the purported binding site is DOPA converted from tyrosine, this is simply the number of tyrosine residues in a given mass of the hagfish IF.

Standards were prepared by adding IF samples to the nanopure H_2O and subjecting them to the same rinsing procedure. 5 mM Ca^{2+} solutions were also prepared with CaCl_2 (much more soluble in water) in nanopure H_2O and additional samples were treated in this solution. The samples were dried overnight in a vacuum chamber on glass microscope slides before being subjected to energy dispersive x-ray spectroscopy to determine elemental content.

7.2 Using REAPDOR to Determine Metal Binding Site

If the samples could be prepared reliably, the next step would be to determine the location that the metals were binding. This could be done using rotational echo, adiabatic passage double resonance (REAPDOR) NMR, an experiment developed by Gullion [16]. The REAPDOR experiment isolates the dipolar coupling interaction between a spin- $\frac{1}{2}$ nucleus and a spin ≥ 1 nucleus with a large quadrupolar coupling. By comparing the estimated nuclear separation at the possible binding sites in the protein with that obtained experimentally, the location or locations binding the metal could be uncovered.

The REAPDOR pulse sequence (Figure 7.2) essentially allows the observation of magnetization dephased solely under the dipolar interaction, using selective π pulses every half rotor cycle, $T_R/2$. In the absence of pulses, the anisotropic interactions are averaged to zero by MAS. The addition of perfect π pulses on the I channel every half rotor cycle ensures there is no contribution from the isotropic chemical shift Hamiltonian. The I channel pulses, however, result in a non-zero average

for the anisotropic terms over a single rotor cycle, T_R . This undesired effect is corrected by omitting a π pulse at the center of the pulse sequence, refocusing the dephasing from these interactions at the time of signal acquisition. Finally, since the dipolar interaction is dependent on both the I and S spins, an inverting pulse on the S channel causes the refocusing of the dipolar dephasing to be undone. The pulse is termed the *recoupling* pulse and allows the observation of the dipolar dephased signal. Application of the sequence without the S channel recoupling pulse provides a reference signal which accounts for T_2 relaxation effects. By comparing a dipolar dephased signal to the reference signal, a measure of the dipolar coupling can be obtained. To understand this, the

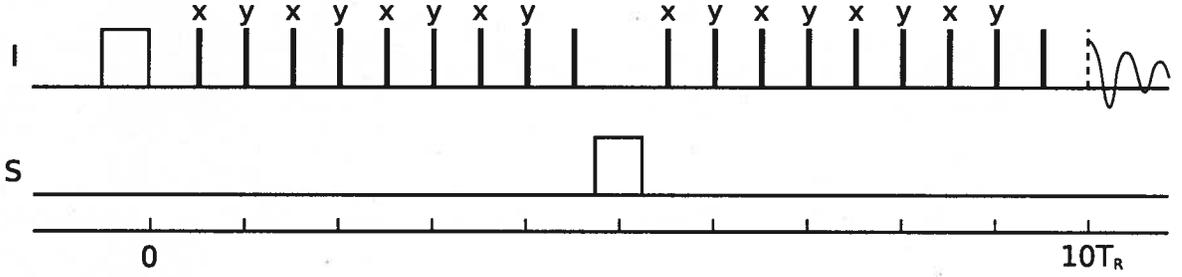


Figure 7.2: A REAPDOR pulse sequence. The x and y are hard π pulses at $\frac{T_R}{2}$ intervals. The pulse on the S channel is an adiabatic passage recoupling pulse.

pertinent interactions must be traced over the course of the pulse sequence.

First, the Hamiltonians must be written in terms of PAF components. The Hamiltonian of the I - S system with $I = \frac{1}{2}$ and $S \geq 1$ has the terms:

$$\hat{H} = \hat{H}_{csa}^I + \hat{H}_{csa}^S + \hat{H}_d + \hat{H}_Q + \hat{H}_{rf}^I + \hat{H}_{rf}^S, \quad (7.1)$$

respectively denoting the chemical shift for the I and S spins, the dipolar coupling, the quadrupolar coupling and the rf pulses on the I and S spins. In the rotating frame, the terms have the form:

$$\begin{aligned} \hat{H}_{csa}^I &= -\frac{1}{2}\omega_I\delta^I (3\cos^2\theta_c^I - 1 - \eta_c^I \sin^2\theta_c^I \cos 2\phi_c^I) \hat{I}_z \\ \hat{H}_{csa}^S &= -\frac{1}{2}\omega_S\delta^S (3\cos^2\theta_c^S - 1 - \eta_c^S \sin^2\theta_c^S \cos 2\phi_c^S) \hat{S}_z \\ \hat{H}_d &= \frac{\gamma_I\gamma_S\hbar}{r_{IS}^3} (1 - 3\cos^2\theta_d) \hat{I}_z \hat{S}_z \\ \hat{H}_Q &= \frac{1}{2}\omega_Q (3\cos^2\theta_Q - 1 - \eta_Q \sin^2\theta_Q \cos 2\phi_Q) \left[\hat{S}_z^2 - \frac{1}{3}S(S+1)\hat{1} \right]. \end{aligned} \quad (7.2)$$

The quadrupolar frequency $\nu_Q = \omega_Q/2\pi$, can be represented in terms of the quadrupolar coupling

constant $\chi = e^2qQ/h$:

$$\nu_Q = \frac{3\chi}{2S(2S-1)} \quad (7.3)$$

Each term in the Hamiltonians (Equation 7.2) has some complicated time dependence of the orientation angles θ_j and ϕ_j , brought on by MAS of the sample, see Section 3.2. The time dependence of the coefficients for each interaction can be represented as a function with terms that oscillate at the MAS rotor frequency ω_R and $2\omega_R$:

$$g(t) = g_1 \cos(\omega_R t + \gamma_1) + g_2 \cos(2\omega_R t + \gamma_2), \quad (7.4)$$

where the γ s are angles representing initial molecular orientations, and the g_α s contain the interaction constants displayed in Equation 7.2.

The $g(t)$ functions are oscillatory with a period of T_R . This periodicity results in rotational echoes which are vital to the function of the sequence, noting:

$$\begin{aligned} \int_0^{T_R} g(t) dt &= 0 \\ \int_0^{\frac{T_R}{2}} g(t) dt &= - \int_{\frac{T_R}{2}}^{T_R} g(t) dt. \end{aligned} \quad (7.5)$$

The functions $g(t)$ must be compared with the tensor form of the Hamiltonians to determine the coefficients g_α in terms of the PAF components. The dipolar Hamiltonian for example,

$$\hat{H}_d = \frac{\gamma_I \gamma_S \hbar}{r_{IS}^3} (1 - 3 \cos^2 \theta_d) \hat{I}_z \hat{S}_z = -2\hat{\mathbf{I}} \cdot \mathbf{D}(t) \cdot \hat{\mathbf{S}}, \quad (7.6)$$

where $\mathbf{D}(t)$ is dependent on $g(t)$ and the initial orientation of the tensor PAF with respect to the static field. To determine $\mathbf{D}(t)$, the rotor frame tensor, $\mathbf{D}^R(t)$ must be projected into the lab frame. The rotor frame tensor is related to the PAF tensor by the appropriate Euler transformation

$$\mathbf{D}^R = R^{-1}(\alpha, \beta, \gamma) \mathbf{D}^{PAF} R(\alpha, \beta, \gamma). \quad (7.7)$$

$$\mathbf{D}^{PAF} = \begin{pmatrix} -\frac{d}{2} & 0 & 0 \\ 0 & -\frac{d}{2} & 0 \\ 0 & 0 & d \end{pmatrix}, \quad (7.8)$$

where d is the dipolar coupling constant for the IS pair,

$$d = \frac{\gamma_I \gamma_S \hbar}{r_{IS}^3}. \quad (7.9)$$

The important component is the lab frame z -component, obtained from the projection of \mathbf{D}^R onto the lab frame z -axis:

$$D_{zz} = \mathbf{b}_0^R \cdot \mathbf{D}^R \cdot \mathbf{b}_0^R, \quad (7.10)$$

where $\mathbf{b}_0^R = (\sin \theta_m \cos \omega_R t, \sin \theta_m \sin \omega_R t, \cos \theta_m)$ is parallel to the static field. Comparing the left and right sides of Equation 7.6:

$$-2D_{zz} = d(1 - 3 \cos^2 \theta_d). \quad (7.11)$$

Finally, the lab frame Hamiltonian can be written in terms of the dipolar tensor PAF components and the initial orientation angles α , β and γ :

$$\hat{H}_d = \left[\sqrt{2}d \sin 2\beta \cos(\omega_R t + \gamma) + d \sin^2 \beta \cos(2\omega_R t + 2\gamma) \right] \hat{I}_z \hat{S}_z. \quad (7.12)$$

Average Hamiltonian theory can be applied to the pulse sequence, noting that $g(t)$ changes sign every half rotor period and the π pulses induce a sign change from the affected spin operator. The first sequence, to obtain the control signal, is run without the recoupling pulse. The resulting average Hamiltonians are \hat{H}_d , \hat{H}_{csa} and \hat{H}_Q . When the recoupling pulse is applied to obtain the dipolar dephased signal, the chemical shift and quadrupolar Hamiltonians are averaged to 0 by MAS, while the dipolar Hamiltonian average is unaffected. This is due to the fact that pulses on both channels affect the sign of the dipolar Hamiltonian, which has an $\hat{I}_z \hat{S}_z$ dependence, while the other interactions depend only on \hat{I}_z , \hat{S}_z or \hat{S}_z^2 .

Ideally the recoupling pulse will invert all of the S spins in the sample. However, when the S spin has a large quadrupolar coupling this is not the case, and the recoupling pulse must be the adiabatic passage pulse. With the application of a hard π pulse, only a fraction of the spins with frequencies near the rf pulse will be allowed to undergo state transition, see Figure 7.3. As the rotor spins during the recoupling pulse, the spins experience a range of frequencies due to the time dependence on the quadrupolar interaction. If the frequency of the spin passes through the Larmor frequency ω_0 , then it is allowed to undergo a state transition. The duration of the pulse is

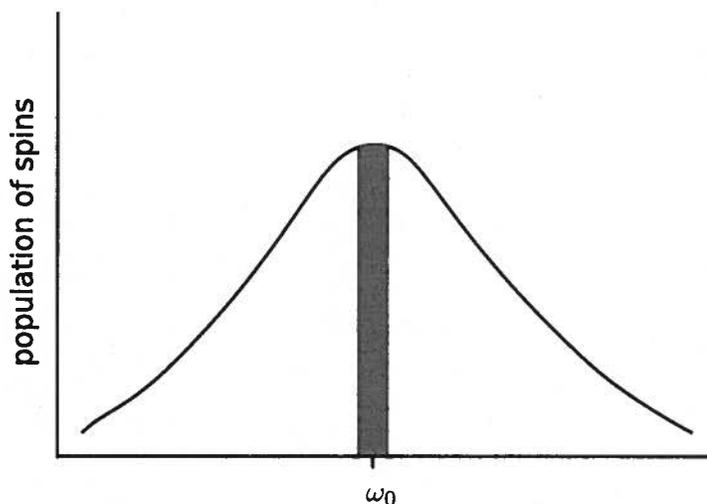


Figure 7.3: The adiabatic passage pulse is required to invert an adequate population of S spins.

selected such that an optimal number of spins precess at the Larmor frequency only once during the recoupling pulse. This optimal duration has been shown to be $T_R/3$ [18].

The adiabatic passage pulse has further power requirements. In order for state transitions to occur, the passage of the nuclear interaction frequency through the Larmor frequency must happen slowly enough to allow the transitions to happen. This requirement is described by the adiabaticity parameter α . The adiabatic passage requirement depends on the rotor frequency and the magnitude of the quadrupolar coupling:

$$\alpha = \frac{\nu_1^2}{\nu_R \nu_Q} > 1. \quad (7.13)$$

A strong rf pulse may have a power of $\nu_1 = 100\text{kHz}$ and a typical spinning rate for suitable resolution is $\nu_R = 5\text{kHz}$. With these conditions, the REAPDOR sequence could be successfully applied to a system which has a quadrupolar coupling frequency ν_Q up to 2MHz. Recently reported values for ν_Q [3, 23, 29] ranged from 0.1 to 0.7 MHz, well below the required limit.

The REAPDOR experiment is repeated numerous times, increasing the number of cycles between time 0 (at the initial $\frac{\pi}{2}$ pulse) and acquisition. For each number of cycles the reference signal, S is compared to the dipolar dephased signal, S_r obtained with the recoupling pulse present. The resulting signal ratio (conventionally displayed as $\Delta S/S = 1 - S_r/S$) is plotted against the experiment duration, which is equivalent to the number of rotor cycles times the rotor period; $\tau = NT_R$. The modulation of the signal ratio with respect to τ is dependent upon the magnitude of the dipolar coupling d . An example of a REAPDOR curve is shown in Figure 7.4, where the abscissa

is plotted as the dimensionless parameter $\lambda = \tau d$. The success of the REAPDOR experiment is

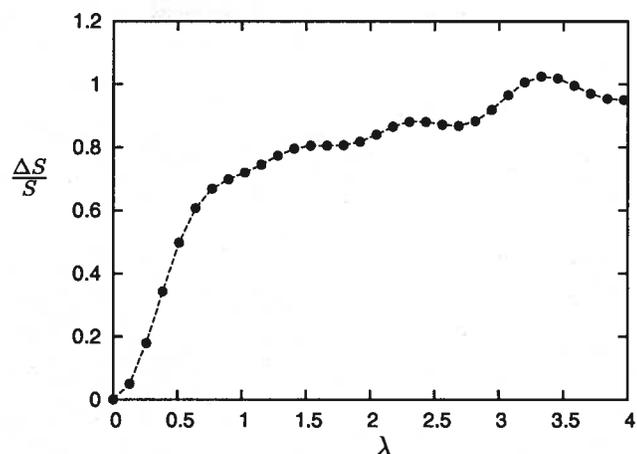


Figure 7.4: An example of a REAPDOR experimental curve.

contingent on the presence of adequate signal at the end of the rotor cycles. Here, dephasing of transverse magnetization is a limiting factor. T_2 relaxation measurements provide a timescale on which the transverse magnetization decays.

7.2.1 T_2 Relaxation in Hagfish IFs

A requirement for the REAPDOR experiment is that the number of cycles be adequate to see a modulation due to the dipolar coupling. To determine feasibility of the REAPDOR experiments, a measure of T_2 relaxation times in the hagfish IFs was performed. The T_2 experiment was a spin echo experiment (Figure 7.5) first developed by Hahn [19]. The T_2 experiment consists of an

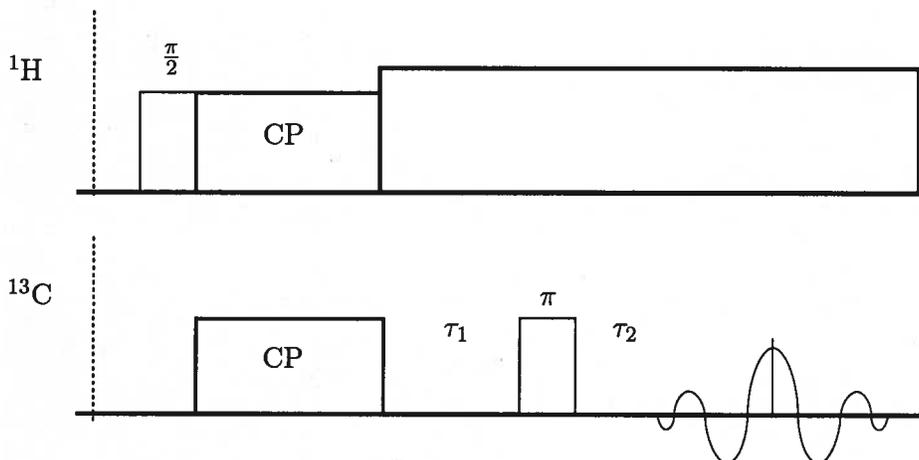


Figure 7.5: Hahn spin echo experiment with cross-polarization. The delays τ_1 and τ_2 are very similar, but the acquisition may start sooner than τ_1 after the π pulse to ensure the full echo is captured.

Peak	Y ϕ	C=O	α	Other
T_2 [ms]	4.72 ± 0.27	6.11 ± 0.09	3.96 ± 0.14	3.72 ± 0.14

Table 7.1: T_2 relaxation times.

initial cross-polarization pulse to put the ^{13}C magnetization in the xy -plane followed by a period of free evolution (decoupling protons to enhance SNR and resolution), τ_1 . During the evolution, the transverse magnetization will begin to dephase, owing to the difference in precession frequencies due to chemical shift and dipolar interactions. After the delay, there is a π pulse on the ^{13}C channel flipping the transverse magnetization. This does not affect the precession frequency of the spins and as a result, the “dephasing” effect is reversed and the transverse magnetization refocuses, see Figure 7.6. This is known as the *spin echo* and it peaks at a time equal to τ_1 after the π pulse.

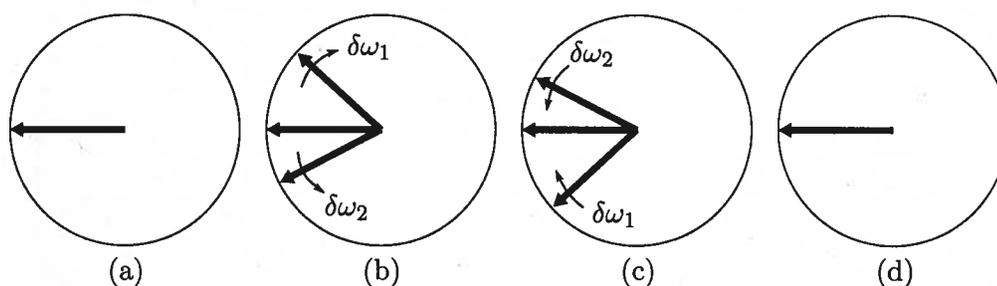


Figure 7.6: The spin echo. Vectors represent transverse magnetization in a frame rotating at the Larmor frequency ω_0 ; (a) Initial transverse magnetization following $\frac{\pi}{2}$ pulse; (b) after time τ the magnetization has dephased due to spins precessing at frequencies slightly different from the Larmor frequency, $\omega_i = \omega_0 + \delta\omega_i$; (c) a π pulse is applied, inverting magnetization; (d) magnetization refocuses at time τ after the π pulse.

Over longer and longer delays, $\tau = 2\tau_1$, the spin echo magnitude will eventually die out. The rate at which it does so is characterized by T_2 , measured in the same manner as T_1 relaxation times in Section 5.3:

$$I(\tau) = I_0 e^{-\frac{\tau}{T_2}}. \quad (7.14)$$

The T_2 values were measured for the hagfish IFs on a Varian 400MHz spectrometer with an HXY 4mm MAS probe spinning at 5kHz. A sample spectrum from the T_2 experiment is shown in Figure 7.7. The resulting T_2 values for peaks identified in the spectrum are summarized in Table 7.1.

The tyrosine ϕ -carbon is the putative binding site. To determine if the T_2 relaxation of 4.72ms for this species is adequate, a computer simulation of the REAPDOR program was written.

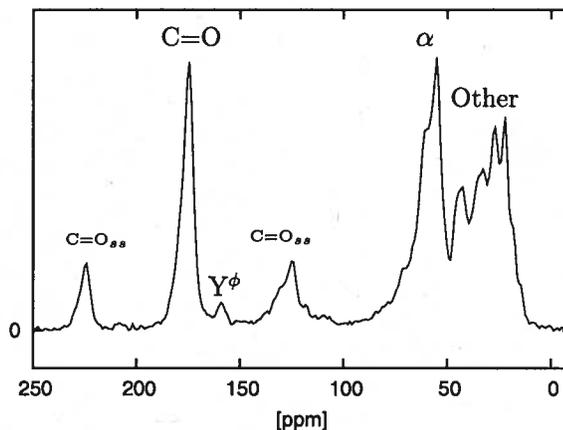


Figure 7.7: Spin Echo experimental spectrum. $\nu_R = 5\text{kHz}$

7.2.2 REAPDOR Simulation

A simulation of the REAPDOR experiment was written, following the algorithm put forth by Hughes and Gullion [21]. The simulation evolves the density operator, ρ according to the Liouville von Neumann equation (Equation 2.21) throughout the pulse sequence and obtains the resulting REAPDOR signal, $\Delta S/S$. The code was written to simulate the dipolar dephasing curve obtained from the DOPA-Ca ligand, Figure 7.1. The dipolar evolution of the I spin magnetization is

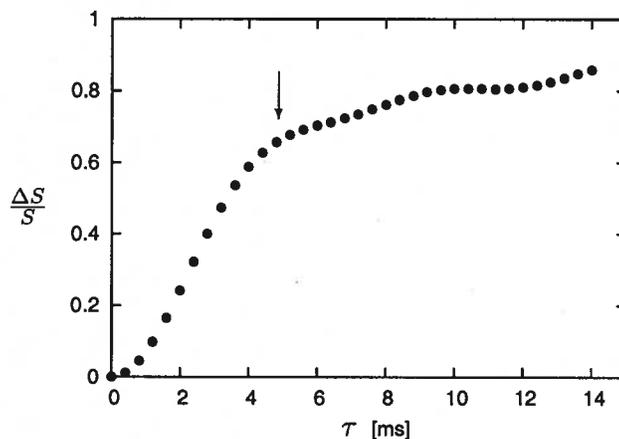


Figure 7.8: REAPDOR Simulation for DOPA-Ca binding: $\nu_R = 5\text{kHz}$, $\nu_d = 57\text{Hz}$, $\chi = 50\text{kHz}$, $S = \frac{7}{2}$. The arrow indicates the T_2 limit for the tyrosine ϕ -carbon.

calculated by evaluating $\text{Tr}(\hat{I}_+ \rho(\tau))$. Each I - S spin pair, at different orientations contribute a dipolar dephased signal $s_d(\alpha, \beta, \tau)$. The ensemble signal is computed as a powder sum of signals over all angles α, β :

$$\frac{S_r}{S} = \frac{\int_{\beta} \int_{\alpha} s_d(\alpha, \beta, \tau) \sin \beta d\alpha d\beta}{\int_{\beta} \int_{\alpha} \sin \beta d\alpha d\beta}. \quad (7.15)$$

The form of s_d depends on the spin number S . For a spin $S = \frac{7}{2}$, the dipolar dephased signal is [18]

$$s_d = \frac{1}{4} [\cos(\omega_d \tau) + \cos(3\omega_d \tau) + \cos(5\omega_d \tau) + \cos(7\omega_d \tau)], \quad (7.16)$$

where $\omega_d = 2\sqrt{2}d \sin 2\beta \sin \alpha$. The solution is quite involved and will be omitted in this work, suffice to say the resulting integral contains Bessel functions [18, 30]:

$$\frac{S_r}{S} = \frac{\sqrt{2}\pi}{16} \left[J_{1/4}(\sqrt{2}\lambda)J_{-1/4}(\sqrt{2}\lambda) + J_{1/4}(3\sqrt{2}\lambda)J_{-1/4}(3\sqrt{2}\lambda) \right. \\ \left. + J_{1/4}(5\sqrt{2}\lambda)J_{-1/4}(5\sqrt{2}\lambda) + J_{1/4}(7\sqrt{2}\lambda)J_{-1/4}(7\sqrt{2}\lambda) \right]. \quad (7.17)$$

The experimental REAPDOR curve can be compared to a curve using Equation 7.17 to determine the dipolar coupling constant and thus confirm the binding location of the calcium ions.

The T_2 result for the proposed binding site, 4.72ms, is indicated on the simulation curve 7.8. Experimental data would be available for the tyrosine ϕ -carbon up to this point. After the T_2 cutoff, the SNR would start to become poor. Although more data would be better, the cutoff is high enough to allow comparison between simulated and experimental results.

7.3 Energy Dispersive X-Ray Spectroscopy

Energy dispersive x-ray spectroscopy (EDX) was employed to determine if calcium could be incorporated into the hagfish IFs. The EDX experiments were carried out on a variable pressure Hitachi S-3000N Scanning Electron Microscope (SEM) with light element EDX at 20 Pa. Electron micrographs of the samples (Figure 7.9) were analysed to identify regions of the protein free of visible impurities. Suitable surfaces were selected for point focused EDX measurements with a beam exposure duration of 200s. Because the SEM did not have vacuum conditions inside, there was some electron beam deviation due to deflections by atmospheric ions inside the chamber. These deviations are infrequent and errant electrons should strike within a 20 μm diameter of the desired focus point.

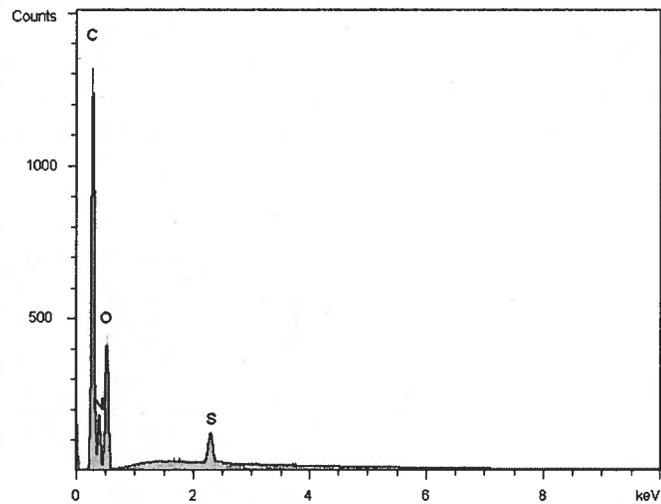
The EDX experiments provided some positive results although they were not easily reproduced due to difficulty in controlling variables in sample preparation. The control samples repeatedly showed a negative result, lacking any sign of Ca in the EDX spectra. However, some control samples were contaminated with Ca and other heavy metals (see Figure 7.10b), illustrating the need for extremely clean handling conditions of the samples during preparation.



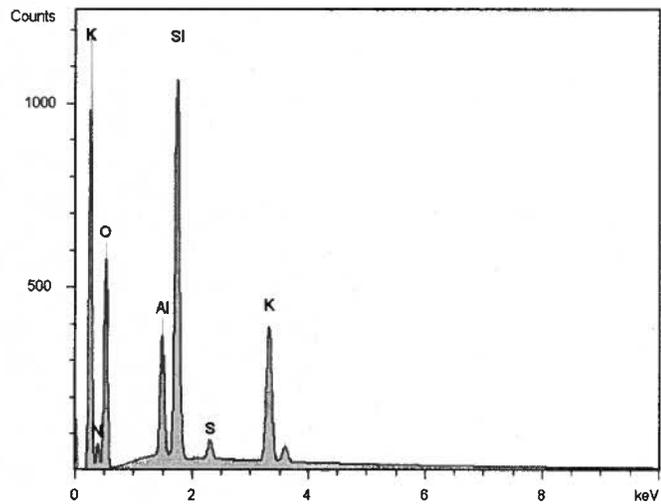
Figure 7.9: Trial SEM micrograph of GTC bundle with nearby CaCO_3 contaminant.

Both the CaCO_3 and CaCl_2 exposed samples exhibited positive and negative results. The positive results were not reliably reproducible. It appears that the ability for the protein to bind metals is dependent on solution pH, but the degree to which is not known. Unfortunately, the NMR active isotope ^{43}Ca poses a potential problem for the experiment due to low natural abundance (0.135%). Because of this, a costly, isotopically enriched treatment compound would be needed and in order to carry out the REAPDOR experiment with limited supply, the sample preparation must result in successful incorporation of Ca into the protein every time. This was not achieved in the course of this study. It is clearly evident that despite careful preparation methods, contamination with heavy atoms is frequent, as shown in Figures 7.10b, 7.11b and 7.12b. This contamination could arise at any step in the preparation or handling of the samples. Exposure to open air was kept to a minimum, but impossible to avoid completely. The use of metal tools was avoided and glass test tubes were cleaned with methanol and rinsed thoroughly with nanopure water before containing samples or solutions. With the CaCO_3 preparation there is some possibility of contamination with Cl^- and Na^+ ions from the HCl/NaOH solutions. These ions should not interfere with the purported binding site as they have incorrect valency.

The spectra displayed in this Section are just a few of the many attempts, but summarize the breadth of the results obtained during the study. Initial preparations were made without monitoring the solution pH carefully and the procedure was revised many times for subsequent trials. Unfortunately, the techniques used were not adequate to acquire the desired confidence levels. The lack of Ca in the treated samples Figures 7.11c and 7.12c must be addressed before



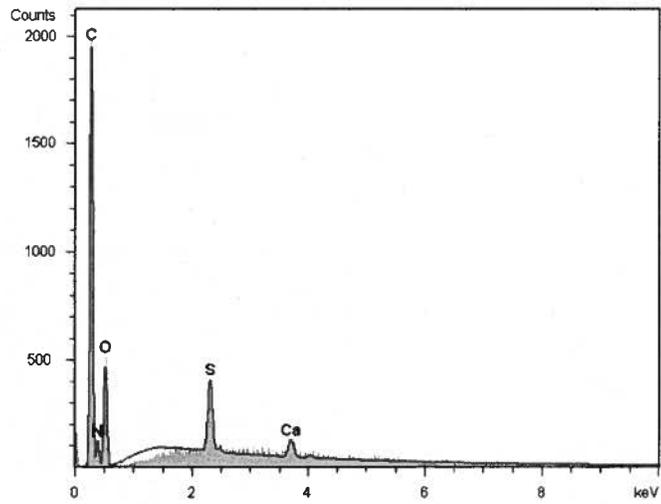
(a)



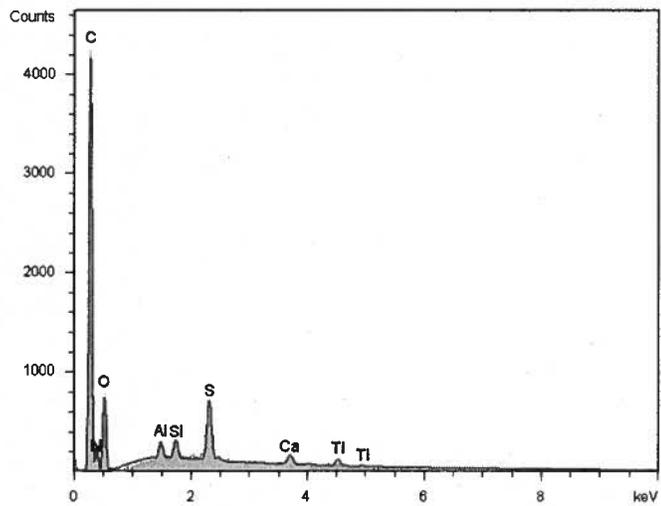
(b)

Figure 7.10: EDX spectra for control samples a) pure b) contaminated.

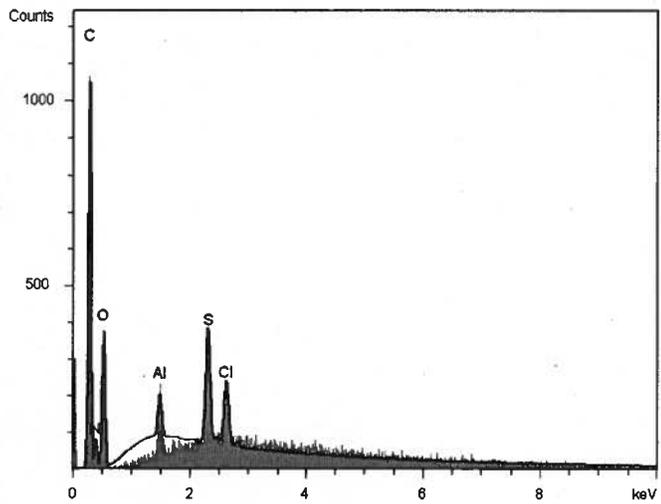
attempting the REAPDOR experiment. Solution pH and temperature or UV exposure of the sample before treatment may affect the ability of the sample to bind calcium. Before a reasonable result can be obtained in this line of research, the effect of these variables on the Ca-binding potential must be carefully analysed.



(a)

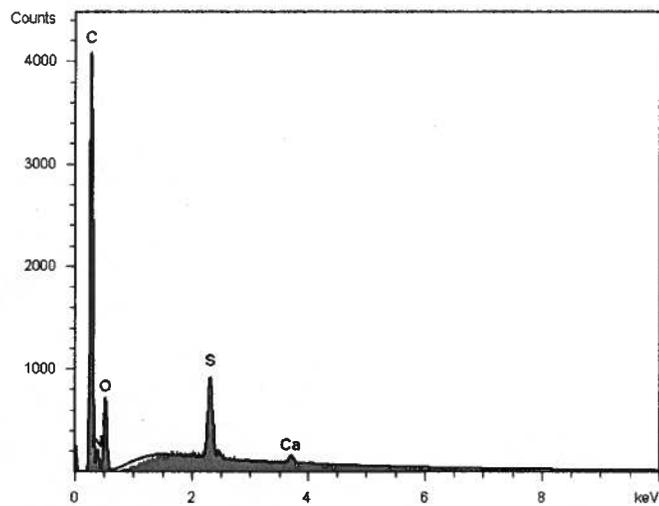


(b)

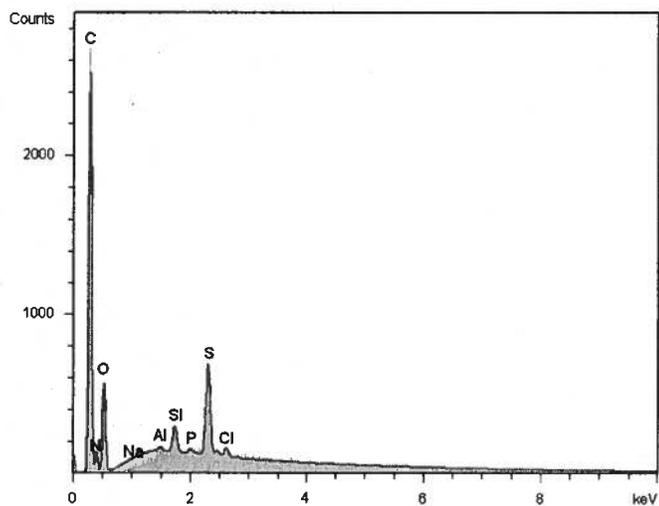


(c)

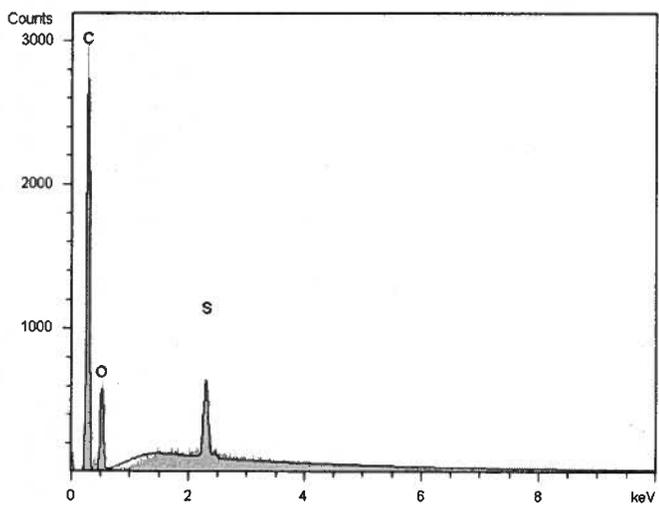
Figure 7.11: EDX spectra for samples treated with CaCO_3 a) positive results b) contaminated c) no apparent Ca presence.



(a)



(b)



(c)

Figure 7.12: EDX spectra for samples treated with CaCl_2 a) positive results b) contaminated c) no apparent Ca presence.

8 Concluding Remarks

A variety of NMR techniques were employed to probe the structure and mechanics of two potentially useful biological materials. Proteins forming materials with mechanical properties rivalling and exceeding those of synthetic materials may eventually replace them. Understanding the facets of the protein structure is an imperative step along the road to achieving this goal.

Hagfish gland thread cell protein forms intermediate filament threads that are strong and highly extensible when hydrated. Direct polarization and cross-polarization experiments were used to expose the underlying two-component structure of the protein backbone. This dichotomous composition agrees with the predicted model of the IF structure. Signal from a rigid central rod domain appears in cross-polarized spectra while the dynamic terminal regions fail to cross-polarize well. The next step would be to definitively assign the dynamic behaviour in the chain to the terminal regions, perhaps using the high concentration of glycine in these domains. It would also be of interest if one could extend these measurements to observe the filaments in an orientationally dependent state, or to look at the effects of the protein dynamics while the filaments are under strain. To accomplish this, the problem of naturally dilute ^{13}C nuclei must be addressed, either by isotopic labelling (very difficult given the size and maturation rate of the organism) or the employment of smaller coils in the NMR probe. Hagfish threads also provide a convenient model of intermediate filaments, normally difficult to study *in vitro*. There is evidence of interaction with metal ions that seem to play a role in the formation and structural properties of the slime as a whole, and may be important to the IFs themselves.

Lineshape simulations were performed to characterize the degree of the molecular motion in the flexible parts of the chain. These measurements provide reasonable estimates of the dynamic nature of the elastic regions, compared with results from similar protein structures.

Recombinant resilin casts were identified as highly elastic materials, composed of protein with an extremely mobile backbone adopting random coil conformational properties. The molecular dynamics reflect the observed macroscopic mechanics of the material. These results suggest that the recombinant resilin has a structural composition similar to the natural resilin studied in the 1960s by Weis-Fogh. The notion of an amorphous network of flexible random coil polypeptides is entirely congruent with the well resolved spectra and the remarkably short correlation times displayed by the resilin samples. Because this protein is rapidly cast from a solution into a solid-state under photo-crosslinking, it may find a wealth of application in biomedical or commercial

pursuits.

The recombinant resilins exhibited a highly mobile backbone structure compared to the hagfish IFs. Backbone carbon correlation times taken from experimental relaxation values were always much faster in the resilin (2 to 8 ns), compared the hagfish IFs (60 ns for the most mobile components). This fact is intuitive as the hagfish IF is a much tougher material and the primary feature of the resilin is extreme elasticity. Elastin, a highly elastic natural protein is a similar mechanically robust material. Previous studies have identified elastin as a network of mobile polypeptide chains with cross-link regions, similar to the resilin model. The relative mobility and structural order of all three materials can be seen in Figure 8.1. From these results it is immediately evident that

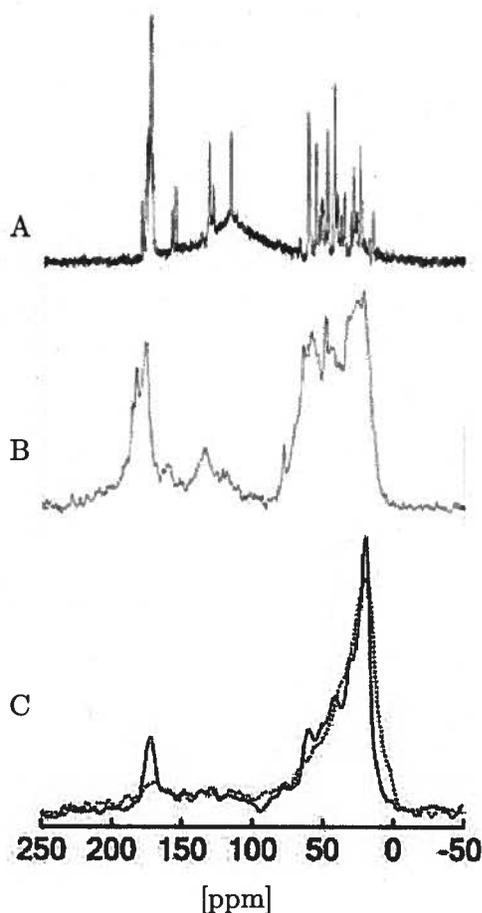


Figure 8.1: DP spectra comparison of elastin (C) with rec-1 resilin (A) and hagfish IF (B) spectra. (C) was taken from [32] and displays spectra for strained (dotted) and unstrained (solid) elastin.

the resilin material has a much more mobile nature in the protein chain than both the hagfish IF and the elastin proteins. The much sharper lineshapes in the resilin spectra reflect the greater degree of molecular motion. The elastin and hagfish IFs exhibit similar linewidths owing to regions

with similar correlation times. Table 8.1 summarizes the dynamic quantities of the three different materials. NMR of the dynamics in elastin networks show a correlation time of approximately

Material	Correlation Time [ns]
Resilin	≤ 8
Elastin	40
Hagfish IF	60

Table 8.1: Summary of dynamic characteristics for structural proteins.

40 ns [28], putting elastin at an intermediate flexibility between the highly mobile resilin and the tougher hagfish IF. The acquired correlation time from the mobile region of the hagfish IF is close to that of the elastin measurements, suggesting that the flexible terminal ends may have a similar, randomly oriented, dynamic nature. This information could also be elucidated by an amino acid specific investigation of mobility in the different regions of the hagfish protein chain.

References

- [1] A. Abragam. *The Principles of Nuclear Magnetism*. Oxford, 1961.
- [2] S. O. Andersen. The crosslinks in resilin identified as dityrosine and trityrosine. *Biochim. Biophys. Acta*, 93:213–215, 1964.
- [3] F. Angeli, M. Gaillard, P. Jollivet, and T. Charpentier. Contribution of ^{43}Ca MAS NMR for probing the structural configuration of calcium in glass. *Chem. Phys. Lett.*, 440(4-6):324–328, 2007.
- [4] D. H. Ardell and S. O. Andersen. Tentative identification of resilin gene in *Drosophila melanogaster*. *Insect Biochem. Mol. Biol.*, 31:965–970, 2001.
- [5] H. Bennet-Clark. Tymbal mechanics and the control of song frequency in the cicada *Cyclochila australasiae*. *Journal of Experimental Biology*, 200:1681–1694, 1997.
- [6] Thomas E. Creighton. *Proteins*. W.H. Freeman and Company, 2nd edition, 1993.
- [7] S. W. Downing, R. H. Spitzer, E. A. Koch, and W. L. Salo. The hagfish slime gland thread cell: I. a unique cellular system for the study of intermediate filaments and intermediate filament-microtubule interactions. *J. Cell Biol.*, 98:653–669, 1984.
- [8] P.T. Eles. *A Nuclear Magnetic Resonance Study of Peptide Backbone Orientation and Dynamics in Spider Dragline Silk and Two-Photon Excitation in Nuclear Magnetic and Quadrupole Resonance*. PhD thesis, University of British Columbia, 2005.
- [9] G.F. Elliott, A.F. Huxley, and T. Weis-Fogh. On the structure of resilin. *J. Mol. Biol.*, 13:791–795, 1965.
- [10] C. M. Elvin, A. G. Carr, M. G. Huson, R. D. Pearson, T. Vuocolo, N. E. Liyou, D. C. C. Wong, D. J. Merritt, and N. E. Dixon. Synthesis and properties of crosslinked recombinant pro-resilin. *Nature*, 437(13):999–1002, 2005.
- [11] J. Gosline et al. Elastic proteins: biological roles and mechanical properties. *Phil. Trans. R. Soc. Lond. B*, 357:121–132, 2002.
- [12] D. S. Fudge. *Biomechanics of Intermediate Filament Based Materials: Insights from Hagfish Slime Threads*. PhD thesis, University of British Columbia, 2002.

- [13] Douglas S. Fudge, Kenn H. Gardner, V. Trevor Forsyth, Christian Riekel, and John M. Gosline. The mechanical properties of hydrated intermediate filaments: Insights from hagfish slime threads. *Biophysical Journal*, 85:2015–2027, September 2003.
- [14] Douglas S. Fudge and John M. Gosline. Molecular design of the α -keratin composite: insights from a matrix-free model, hagfish slime threads. *Proc. R. Soc. Lond. B*, 271:291–299, 2004.
- [15] S. N. Gorb. Serial elastic elements in the damselfly wing: mobile vein joints contain resilin. *Naturwissenschaften*, 86:552–555, 1999.
- [16] T. Gullion. Measurement of dipolar interactions between spin-1/2 and quadrupolar nuclei by rotational-echo, adiabatic-passage, double-resonance nmr. *Chem. Phys. Lett.*, 246:325–330, 1995.
- [17] T. Gullion and J. Schaeffer. Rotational-echo, double-resonance nmr. *J. Magn. Reson.*, 81:196–200, 1989.
- [18] T. Gullion and A. J. Vega. Measuring heteronuclear dipolar couplings for $i=1/2$, $s_j=1/2$ spin pairs by redor and reapior nmr. *Prog. in Nucl. Magn. Reson. Spec.*, 47, 2005.
- [19] E. L. Hahn. Spin echoes. *Phys. Rev.*, 80(4):580–594, 1950.
- [20] I. Hofmann, H. Herrmann, and W. W. Franke. Assembly and structure of calcium-induced thick vimentin filaments. *Eur. J. Cell Biol.*, 56:328–341, 1991.
- [21] E. Hughes, T. Gullion, A. Goldbourt, S. Vega, and A. J. Vega. Internuclear distance determination of $s = 1$, $i = 1/2$ spin pairs using reapior nmr. *J. Magn. Reson.*, 156:230–241, 2002.
- [22] E. Kaplan and E. L. Hahn. *J. Phys. Radium*, 19:821–825, 1958.
- [23] S.H. Kim, W. Gregor, J.M. Peloquin, M. Brynda, and R.D. Britt. Investigation of the calcium-binding site of the oxygen evolving complex of photosystem ii using ^{87}Sr esem spectroscopy. *Journal of the American Chemical Society*, 126(23):7228–7237, 2004.
- [24] Elizabeth A. Koch, Robert H. Spitzer, Ron B. Pithawalla, Francisco A. Castillos III, and David A. D. Parry. Hagfish biopolymer: a type i/type ii homologue of epidermal keratin intermediate filaments. *Int. J. Biol. Macromol.*, 17(5):283–292, 1995.

- [25] Elizabeth A. Koch, Robert H. Spitzer, Ron B. Pithawalla, and David A. D. Parry. An unusual intermediate filament subunit from the cytoskeletal biopolymer released extracellularly into seawater by the primitive hagfish (*eptatretus stouti*). *Journal of Cell Science*, 107:3133–3144, 1994.
- [26] H. Lee, N. F. Scherer, and P. B. Messersmith. Single-molecule mechanics of mussel adhesion. *Proc. Natl. Acad. Sci. USA*, 103(35):12999–13003, 2006.
- [27] M. H. Levitt. *Spin Dynamics: Basics of Nuclear Magnetic Resonance*. John Wiley and Sons, 2001.
- [28] J.R. Lyerla and D.A. Torchia. Molecular mobility and structure of elastin deduced from the solvent and temperature dependence of ^{13}C magnetic resonance relaxation data. *Biochemistry*, 14(23):5175–5182, 1975.
- [29] K. J. D. MacKenzie, M. E. Smith, and A. Wong. A multinuclear mas nmr study of calcium-containing aluminosilicate inorganic polymers. *J. Mat. Chem.*, 17:5090–5096, 2007.
- [30] K.T. Mueller, T.P. Jarvie, D.J. Aurentz, and B.W. Roberts. The redor transform: direct calculation of internuclear couplings from dipolar-dephasing nmr data. *Chem. Phys. Lett.*, 242:535, 1995.
- [31] University of Wisconsin. Biological magnetic resonance data bank. http://www.bmrb.wisc.edu/ref_info/statsel.htm, 2008.
- [32] M.S. Pometun, E.Y. Chekmenev, and R.J. Wittebort. Quantitative observation of backbone disorder in native elastin. *J. Biol. Chem.*, 279(9):7982–7987, 2004.
- [33] X. Qin and J. H. Waite. A potential mediator of collagenous block copolymer gradients in mussel byssal threads. *Proc. Natl. Acad. Sci. USA*, 95:10517–10522, 1998.
- [34] K. Schmidt-Rohr and H.W. Spiess. *Multidimensional Solid-State NMR and Polymers*. Harcourt Brace and Co., 1994.
- [35] D. A. Torchia. The measurement of proton-enhanced carbon- ^{13}C t_1 values by a method which suppresses artifacts. *J. Magn. Reson.*, 30:613–616, 1978.
- [36] T. Weis-Fogh. A rubber-like protein in insect cuticle. *J. Exp. Biol.*, 37:887–907, 1960.

- [37] T. Weis-Fogh. Molecular interpretation of the elasticity of resilin, a rubber-like protein. *J. Mol. Biol.*, 3:648–667, 1961.
- [38] Z. Yang, O. Liivak, A. Seidel, G. LaVerde, D. Zax, and L.W. Jelinski. Supercontraction and backbone dynamics in spider silk: ^{13}C and ^2H nmr studies. *J. Am. Chem. Soc.*, 122:9019–9025, 2000.
- [39] H. Zhang, S. Neal, and D. S. Wishart. Refdb: a database of uniformly referenced protein chemical shifts. *J. Biomol. NMR*, 25(3):173–95, 2003.
- [40] H. Zhao, N. B. Robertson, S. A. Jewhurst, and J. H. Waite. Probing the adhesive footprints of *mytilus californianus* byssus. *J. Biol. Chem.*, 281(16):11090–11096, 2006.

A Lineshape Simulation Code

```
/* program to calculate spectrum of a CSA undergoing random flips, with a s
   correlation time and an average angular jump
*/

#include <math.h>
#include <stdio.h>
#include <stdlib.h>

/* for .01us did NPC=1024, BR=400, NSIM=100
   for .1us did NPC=1024, BR=400, NSIM=1000
   for 1us did NPC=1024, BR=400, NSIM=10000
   for 2us did NPC=1024, BR=400, NSIM=10000
   for 5us did NPC=1024, NR=400, NSIM=100000 */

// number of sites to simulate 100000 for static, try 1000 for rapid flipping
#define NSIM 10000

// correlation time in sec
#define TAUC 0.8e-6

// jump size has an exponential distribution, this is correlation angle
#define DTH 15

#define DT 2e-6 // dwell time of data
#define NP 2048 // number of points in fid.
#define NPC 2048 // number of points to actually calc in fid
#define BR 400.*DT*M_PI // broadening in points

// CSA parameters based on polyglycine data
#define DELTA -75.66*100.523*2*M_PI
#define ETA -0.8267
#define ISO 0.0

void evolve_to_next();
float dur,curr_freq,phase,thi,phi;

main(){
  float spec[2*NP],ct,sti,fid[2*NP];
  int i,j,next=0,last=0;

  for (i=0;i<2*NP;i++)
    spec[i]=0.;

  for(j=0;j<NSIM;j++){

    // get the initial angle
    thi = (float) random()*M_PI/RAND_MAX;
    phi = (float) random()*M_PI/RAND_MAX;

    sti=sin(thi);
    ct=cos(thi);
    ct=ct*ct;
    curr_freq = DELTA/2.*(3*ct-1+ETA*(1.-ct)*cos(2*phi));
```

```

phase = 0;
dur = -log((float) random()/RAND_MAX)*TAUC;

for( i=0 ; i<NPC ; i++ ){
    spec[2*i] += cos(phase)*sti;
    spec[2*i+1] += sin(phase)*sti;

    evolve_to_next();
}

// keep the user informed:
next = (int) j/(NSIM/101.);
if (next > last) {
    fprintf(stderr, "\r%i/100", next);
    //      fflush(stdout);
    last=next;
}

}
fprintf(stderr, "\n");
// now do ft and output.

// store fid before broaden.
for(i=0; i<2*NP; i++)
    fid[i]=spec[i];

// broaden:
for (i=0; i<NP; i++){
    spec[2*i] *= exp(-(float)i*BR);
    spec[2*i+1] *= exp(-(float)i*BR);
}

spec[0] /=2.;

four1(spec-1, NP, -1);
// unscramble
for(i=0; i<NP; i++){
    thi=spec[i];
    spec[i]=spec[i+NP];
    spec[i+NP]=thi;
}

for(i=0; i<NP; i++)
    printf("%f %f %f %f %f\n", i/DT/NP/100.523, spec[2*i],
        spec[2*i+1], fid[2*i], fid[2*i+1]);
}

void evolve_to_next(){
    float time_remain=DT;
    float thn, phn, gam, squig;
    float cg, st, ct, sg, sp, cp, cs, ss;

    do{
        if (dur > time_remain){
            phase += curr_freq*DT;

```

```

    dur -= DT;
    return;
}
else{
    phase += curr_freq*dur;
    time_remain -= dur;
    dur = -log((float) random()/RAND_MAX)*TAUC;
    // do the jump, set the new curr_freq
    squig = (float)random()/RAND_MAX*M_PI*2; // could be in any direction

    // exponentially distributed gamma
    gam = -log((float) random()/RAND_MAX)*DTH*M_PI/180.;
    // get the new angles
    cg=cos(gam);
    sg=sin(gam);
    st=sin(thi);
    ct=cos(thi);
    sp=sin(phi);
    cp=cos(phi);
    ss=sin(squig);
    cs=cos(squig);
    phn = atan2(cg*st*sp-sg*(cs*ct*sp+ss*cp),
cg*st*cp-sg*(cs*ct*cp-ss*sp));
    ct=sg*cs*st+cg*ct;
    if (ct > 1 || ct < -1) fprintf(stderr,"ct is: %f\n",ct);
    thn = acos(ct);

    thi=thn;
    phi=phn;
    //      fprintf(stderr,"\n%f %f %f %f\n",thn,phn,squig,gam);

    ct=ct*ct;
    curr_freq = DELTA/2.*(3*ct-1.+ETA*(1.-ct)*cos(2*phi));
}
} while(0==0);
}

```