Site-Specific Solid-State NMR Studies of the Protein-Water Interface of
Anabaena Sensory Rhodopsin

by
Emily Ritz

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Physics

Guelph, Ontario, Canada

© Emily Ritz, September, 2012
ABSTRACT

SITE-SPECIFIC SOLID-STATE NMR STUDIES OF THE PROTEIN-WATER INTERFACE IN ANABAENA SENSORY RHODOPSIN

Emily Ritz  
University of Guelph, 2012

Advisor:  
Professor V. Ladizhansky

Solid-state NMR spectroscopy was used to site-specifically investigate the protein-water interface of a seven alpha-helical transmembrane protein, *Anabaena* sensory rhodopsin (ASR). Water-edited experiments, which employ a T2-filter to select for mobile protons, provided a means to detect residues which appear to be in close contact to water molecules, and to gain insights about the water-protein interface of ASR. First, through the application of Lee-Goldburg homonuclear decoupling, it was determined that polarization transfer across this interface is dominated by through-space interaction mechanisms, as opposed to chemical exchange. A series of two-dimensional experiments were also performed to detect polarization transfer along the backbone and to the sidechains of the protein. Residues located in solvent-accessible regions of the protein, such as the B-C loop, were found to obtain polarization quickly, as expected, and in agreement with previous H/D exchange data. Residues known to be in contact with bound crystal water molecules were also detected. In addition to these, we found new residues which appear to be in contact with water, indicating additional H\(_N\)-H\(_2\)O interactions, or additional contacts with bound water molecules. Most of these residues were located beside exchangeable regions of ASR. Sidechains of residues located in the cytoplasmic side of helix F were seen to be in close contact with mobile water molecules, supporting evidence of a hydrophilic chain along the cytoplasmic half of the protein, which is suggested to cause a functional outward tilt of the cytoplasmic half of helix F upon light-activation.
Acknowledgements

First, I would like to thank my advisor, Dr. Vladimir Ladizhansky, for his guidance over the past few years. I am encouraged and inspired by his patience, ability to explain and teach difficult concepts, pursuit of knowledge, love of learning, and passion for this field. I thank him for affirming my abilities, and my research, even when I doubted them. I am thankful to Dr. Leonid Brown for his support, advice, and new perspectives in the research as well. I also thank Dr. Jim Davis for being on my advisory and examination committees. I am grateful for the staff of the NMR Centre, Valerie Robertson, Peter Schaeffer, and Dr. Andy Lo, for assisting me with technical problems, and for always being willing to help in any way they could.

Thank you to the past members of our research group, Dr. Shenlin Wang, Dr. Lichi Shi, Dr. Mumdooh Ahmed, and Andrew Gravelle for being both teachers and mentors to me throughout this learning process. To the current members of the group, Daryl Good, Meaghan Ward, and Matthew O’Halloran, I have learned so much from each of you – thank you for teaching and learning alongside me. I am thankful for my officemates, Ivana Komljenovic, Miranda Schmidt, Sanaz Emami, and Iain Braithwaite who have made the office an enjoyable place to work over the years. And to James – my depth of gratitude goes beyond words.

Finally, thank you to my amazing community of friends and family, for your grace and understanding in all the times I put this work ahead of you, your encouragement and support through the frustrations and difficulties, and for the love you show as you celebrate in my physics joys.
O Lord my God! When I in awesome wonder
Consider all the works Thy hands have made
I see the stars, I hear the rolling thunder,
Thy power throughout the universe displayed.

Then sings my soul, my Saviour God, to Thee:
How great Thou art, how great Thou art!

-Carl Boberg & Stuart Hine
Table of Contents

CHAPTER 1 INTRODUCTION ............................................................................................................. 1
  1.1 MEMBRANE PROTEINS .................................................................................................................. 1
  1.2 PREVIOUS PROTEIN-WATER INTERACTION STUDIES ............................................................... 3
  1.3 ANABAENA SENSORY RHODOPSIN ............................................................................................ 7
  1.4 EXPERIMENTAL STRATEGIES .................................................................................................... 9
  1.5 SUMMARY OF GOALS .................................................................................................................. 11

CHAPTER 2 NMR THEORY .................................................................................................................. 13
  2.1 STATIC NMR BACKGROUND ....................................................................................................... 13
     2.1.1 Zeeman Interaction ............................................................................................................... 14
     2.1.2 Chemical Shift Interaction ................................................................................................ 16
     2.1.3 Dipolar Interaction ............................................................................................................. 17
     2.1.4 J-Coupling ........................................................................................................................... 18
     2.1.5 Applied Radiofrequency Field ............................................................................................ 18
  2.2 MAGIC ANGLE SPINNING NMR ................................................................................................. 19
  2.3 NMR EXPERIMENTS .................................................................................................................... 25
     2.3.1 Signal Detection .................................................................................................................... 25
     2.3.2 Cross-Polarization and Multidimensional Spectroscopy ..................................................... 26
  2.4 RELAXATION RATES ................................................................................................................... 30
     2.4.1 $T_1$ – Longitudinal Relaxation .............................................................................................. 30
     2.4.2 $T_2$ and $T_2^*$ – Transverse Relaxation ................................................................................. 30
     2.4.3 $T_{1p}$ – Transverse Relaxation in the Rotating Frame ............................................................ 31

CHAPTER 3 EXPERIMENTAL METHODS ......................................................................................... 32
  3.1 SAMPLE PREPARATION ............................................................................................................... 32
     3.1.1 Materials ............................................................................................................................... 32
     3.1.2 Preparation of GB3 .............................................................................................................. 32
     3.1.3 Preparation of ASR .............................................................................................................. 32
  3.2 SSNMR EXPERIMENTAL SETUP ................................................................................................ 33
     3.2.1 $T_2$ – Filter ............................................................................................................................. 33
     3.2.2 N-CA Chemical Shift Correlation Experiments ................................................................. 34
     3.2.3 C-C Chemical Shift Correlation Experiments ....................................................................... 35
     3.2.4 Spin Lock at the Magic Angle ............................................................................................... 36
     3.2.5 H-(N)-CA Correlation Experiments .................................................................................... 39

CHAPTER 4 RESULTS AND DISCUSSION ....................................................................................... 42
  4.1 INVESTIGATION OF THE PROTEIN-WATER INTERFACE USING 1D SPECTROSCOPY .................. 42
     4.1.1 Nitrogen Spectra ..................................................................................................................... 45
     4.1.2 Carbon Spectra ...................................................................................................................... 50
  4.2 WATER-PROTEIN CORRELATION EXPERIMENTS .................................................................... 55
  4.3 METHODOLOGY OF SITE-SPECIFIC DETECTION .................................................................... 62
  4.4 SITE-SPECIFIC DETECTION RESULTS ......................................................................................... 66
     4.4.1 Investigation of Backbone Polarization Transfer ................................................................. 66
     4.4.2 Investigation of Sidechain Polarization Transfer ................................................................. 74

CHAPTER 5 FINAL SUMMARY AND FUTURE DIRECTIONS ............................................................ 81
1.1 Membrane Proteins

Membrane proteins constitute about one third of all known proteins\(^1\), account for more than half of all existing drug targets, and include a variety of important functional classes, such as enzymes, ion channels, transporters and receptors. They control many essential biological and chemical functions, including protein synthesis, signal transduction, and ion transport.\(^1,2\) Water is ubiquitously present in biological systems, and the interactions occurring between water and proteins are the major determinants of protein stability.\(^3\) These interactions are important in determining, and characterize, the structure, folding, and dynamics of a membrane protein. The function of biological systems, specifically membrane proteins, depends on the interactions between the protein and solvent water. Water molecules are essential to the functionality of membrane proteins, especially in the case of ion channels and proton pumps.\(^4,5\) The interactions between these molecules and the protein can be an indicator of solvent accessibility, and of the implication of protein protons (or other nuclei like nitrogen and oxygen) in hydrogen-bonds, which gives rise to valuable structural and dynamical information, so their study is of great interest and importance.

Recently, solid-state nuclear magnetic resonance (SSNMR) has rapidly evolved into a powerful technique suitable for determining high resolution structures and studying dynamics of membrane proteins in the lipid environment in their native-like functional states.\(^1,2,6-11\) This technique is capable of providing site-specific information, and therefore valuable structural
insight, about a solid biological system. In this work, we use this site-specific spectroscopy to investigate the interactions between protein and solvent.

There are a variety of ways in which water can interact with a membrane protein. First, many polar residues on the solvent-exposed protein surface can directly interact with water. Protons in their exchangeable groups (OH, for example) can undergo chemical exchange with water protons. If the residence time of water molecules on the exposed protein surface (which has been shown to be on the order of milliseconds\textsuperscript{12}) is sufficiently long on the timescale of proton-proton dipolar interactions, which are typically \(\sim 120\) kHz (for two protons 1Å apart), there can be through-space polarization exchange across the interface between the protein and water proton species. Polarization can also transfer through cross-relaxation effects.\textsuperscript{13,14}

![Figure 1-1: Schematic of various polarization transfer mechanisms across the protein-water interface. Solid arrow indicates chemical exchange between protons in exchangeable groups and water protons. The dotted arrows indicate through-space transfer, which occurs between water molecules in close proximity to protein protons (dipolar transfer and cross-relaxation effects), and also gets redistributed through intraprotein spin diffusion. Blue spheres represent bound waters which can exchange with bulk water and transfer polarization to nearby protein protons.](image)

In addition to having exposed regions, internal residues can have bound water in their proximity. Many X-ray derived structures of both soluble and membrane proteins contain bound
water molecules, which typically form hydrogen bonds with protein polar groups, and play an essential role in protein function. Specific examples among membrane proteins include proton pumps (e.g., bacteriorhodopsin), water channels (e.g., aquaporins), ion channels and many other proteins. Identification of these internal bound water molecules is essential to understating protein functions.

1.2 Previous Protein-Water Interaction Studies

A variety of techniques can be used to study protein-solvent interactions. X-ray diffraction techniques have been used to detect hydration waters in crystallized protein. Neutron diffraction studies from Kossiakoff et al of trypsin protein that had been placed in D$_2$O for about a year produced qualitative proton-deuterium chemical exchange data. This data revealed that most proton sites were fully exchanged, except where protons were well-protected by secondary structure, and that free solvent exchange occurs between both bulk and interior solvent regions. According to Ref [17], solvent accessibility alone is not the only condition required for exchange to occur, since exchangeable regions of the protein were found even in the interior of the protein, which is not exposed to bulk solvent. However, these diffraction techniques are limited by the requirement that the protein be crystallized, so information about interaction behaviour can only be gained once the structure is already determined and contains coordinates of bound waters. While X-ray and neutron studies provide static, qualitative information about the protein before and after its interactions occur, additional studies are required for quantitative, “time-resolved” information on a site-specific level.

Solution NMR has been proven as a suitable technique to study the structure of membrane proteins in detergent micelles, but the large size of these systems often prevents
them from tumbling rapidly enough to average anisotropic interactions. However, solution NMR has remained as a suitable technique to study interactions between protein and water in soluble systems.\textsuperscript{21-24} It has also been used for observing quantitative proton-deuterium (H/D) exchange rates in soluble proteins.\textsuperscript{25} The work of Wagner and Wuthrich (Ref. [5]) used a 2D NMR experiment to obtain a complete, site-specific exchange data set, in which individual exchange rates of all the amide protons in a soluble protein, Basic Pancreatic Trypsin Inhibitor (BPTI), were found. Since that experiment in 1982, soluble NMR has been combined with H/D exchange to further study the interactions between a protein and its solvent in a site-specific manner.\textsuperscript{26,27} This provides valuable structural insight about the solvent accessibility, or about the strength of hydrogen bonding, at various sites on a protein.

While H/D exchange experiments are straightforward in the case of soluble proteins, the observation of proton exchange and other protein-solvent interactions is more complicated in the solid environment. These complications are generally associated with difficulties of spectral sensitivity in SSNMR spectra. In solids, the dipolar couplings are not averaged out through molecular tumbling. The increased number of coupled protons results in crowded and broad spectra, making direct detection of protons difficult. A common remedy for this problem is the deuteration of solid-state systems, which decreases the number of proton-proton couplings by diluting the proton spin bath.\textsuperscript{28-30} As a result, the rate of proton spin diffusion is decreased, leading to improved proton resolution (~0.2 ppm).\textsuperscript{29-31} This, along with homonuclear decoupling methods,\textsuperscript{32} allows for high-resolution proton detection, and thus for the direct observation of protein-solvent interactions through the detection of interacting protons.\textsuperscript{33} However, some solid-state protein systems are not readily amenable to proton back-exchange\textsuperscript{34,35}, so a more universally-applicable method is required in these cases.
One alternative is to work with fully protonated samples under fast magic angle spinning (MAS) conditions to reduce proton-proton dipolar couplings and reduce linewidth. Although fast MAS proton spectra are very complex, they provide a more complete and functionally relevant picture of protein-solvent interactions than in the case of deuterated samples, because membrane proteins are naturally in an aqueous environment.\textsuperscript{36} Since the properties of the protein, such as the dynamics and exchange rates, could differ under deuteration conditions, an investigation of these interactions needs to be performed in this more realistic (fully protonated) membrane protein environment.

SSNMR techniques in fully protonated samples provide enough sensitivity to study the interactions, such as proton exchange, between protein and solvent in a lipid membrane.\textsuperscript{14,37–40} However, in a fully protonated sample, the proton exchange process across the protein-solvent interface competes with intraprotein spin diffusion between protons. In the first attempt to characterize protein-water interactions, Harbison \textit{et al} used a so-called T\	extsubscript{2}–filter approach (to be described in more detail later) to completely remove proton signals of bacteriorhodopsin, leaving only the solvent protons polarized.\textsuperscript{41} They monitored the kinetic transfer of magnetization from solvent water to a single site in the protein. This provided evidence that the Schiff base proton of bacteriorhodopsin undergoes direct chemical exchange with the bulk water. This experimental approach allows for the use of fully protonated samples, and has been employed in the study of protein-solvent interactions since its introduction.\textsuperscript{12,42,43} The T\	extsubscript{2}–filter method has been used towards detecting interactions between protein and mobile-water in the solid-state.

In order to fully characterize the protein-solvent interface, surface exposed residues must be distinguished from residues buried within the core of the protein. Distinguishing these residues would also allow for the differentiation of polarization transfer mechanisms, and for
unambiguous identification of the solvent-protein polarization transfer mechanisms, either of chemical exchange or through-space dipolar nature. Lesage, Bockmann and co-workers detected polarization transfer arising from chemical exchange in a microcrystalline catabolite repression histidine-containing phosphocarrier (Crh) protein.\textsuperscript{33} They performed a series of studies to differentiate between possible transfer mechanisms, including varying temperature\textsuperscript{44} and applying various homonuclear decoupling techniques\textsuperscript{32}, and concluded that chemical exchange remained the dominant transfer mechanism in their sample. Differentiating the contributions of chemical exchange and dipolar interactions to polarization transfer could lead to a quantification of chemical exchange rates on a site-specific level.

Understanding the interactions between protein and mobile water molecules can be used for the identification of structural rearrangements during the functional cycle of proteins. This was seen in Ader \textit{et al} in their work in Ref.[18], where changes in the solvent accessibility in response to the functional cycle of the chimeric potassium channel KcsA-Kv1.3 led to an estimate of the change in diameter of the channel.\textsuperscript{45} In this work, diffusion coefficients were estimated for the protein-water interface and protein-protein polarization transfer. Water surface to protein-volume ratios were calculated for various membrane-proteins, and thus the global protein-water surface could be monitored. While this provided a means to observe structural changes in response to function, these were only seen on a global scale, and site-specific information is lacking. Site-specific structural changes to the membrane protein \textit{Anabaena} sensory rhodopsin (ASR) during its functional cycle were detected using H/D exchange experiments performed by Wang \textit{et al}.\textsuperscript{46} H/D exchange experiments revealed site-specific changes in the solvent accessibility upon light-activation of ASR, suggesting that a cleft opens on the cytoplasmic side between helices, allowing for water to penetrate into the protein core.
Despite these advances, there are still areas of research to be explored, especially towards identifying protein and mobile-water interactions on a site-specific level. Such an understanding would allow for the quantification of site-specific chemical exchange rates. This, and further work in the resolution of bound waters in order to gain a complete picture of the position of these bound water molecules and the hydrogen bonds involved in their interactions, is needed in order to improve understanding of membrane protein structure and function. This thesis aims to develop methods for site-specific detection of protein-water interactions in membrane proteins. We also aim to quantify these interaction rates and polarization transfer kinetics. To accomplish these goals, we use the membrane protein \textit{Anabaena} sensory rhodopsin.

1.3 \textbf{Anabaena Sensory Rhodopsin}

Rhodopsins are a class of membrane proteins found in archaeal, eubacterial and eukaryotic domains which bind vitamin-A aldehyde (retinal) as their chromophore. All rhodopsin proteins share a common structural architecture of seven membrane-spanning \(\alpha\)-helices which form a pocket for the retinal molecule, which is attached to a lysine residue in the 7th helix (helix G) through a protonated Schiff base. Thus, there is a positive charge buried within the helices of these proteins, often transported upon light-activation, which is related to the function of rhodopsin proteins.\(^{47}\) There are two main families of rhodopsin proteins, one belonging primarily to the bacterial and lower eukaryotes (Type 1) and the other being found so far in higher eukaryotes (Type 2). Type 1 rhodopsins are more relevant for this discussion, and function mostly as light-driven ion transporters or phototaxis receptors (including bacteriorhodopsin, halorhodopsin and sensory rhodopsins I and II). \textit{Anabaena} sensory rhodopsin belongs to this Type 1 family of rhodopsins.
Anabaena sensory rhodopsin (ASR) is a 261-residue, 26 kDa, seven-alpha-helical transmembrane protein, and the first sensory rhodopsin to be observed in the eubacterial domain. In these experiments, a C-terminally 6xHis-tagged form, truncated at position 229 was used for better expression. The amino acid sequence is shown in Figure 1-2.

ASR differs from microbial rhodopsins in the archaeal domain in a number of ways. First, on its cytoplasmic side, ASR has a hydrophilic cavity that contains a number of water molecules creating a network of hydrogen bonds from near the Schiff Base-forming lysine (Lys210) to the cytoplasmic surface, making it structurally different from archaeal rhodopsins. Also, ASR functions as a photosensory receptor, but unlike archaeal sensory rhodopsins, transmits signals through a soluble cytoplasmic protein, Anabaena sensory rhodopsin transducer (ASRT).

ASR provides an excellent candidate for studying water-protein interactions, since its high-resolution structure was obtained by X-ray diffraction, and solid-state NMR high-
resolution structure (which differs from the X-ray structure, revealing that ASR forms a trimer), and assignments for most of its residues are available. The X-ray structure reveals a number of bound water molecules, which are covalently bonded to intra-helical residues. Recently, MAS SSNMR experiments were performed which detected individual amino acids undergoing H/D exchange, both in the dark and under illumination, providing insight into the protein’s structural changes upon light activation. This H/D exchange data can also provide a reference, and guide the discussion of which areas of the protein might be in contact with solvent water.

1.4 Experimental Strategies

In this thesis, the T₂-filter method is employed in order to prepare a “non-equilibrium state” of magnetization, where only water and highly mobile species (those species with a long T₂, described above) are polarized. A similar approach, exploiting a non-equilibrium state of magnetization to study exchange processes, was introduced by Harbison et al., and the following experimental strategy is based upon this framework of selective excitation. These mobile species can interact with the non-polarized protein surface, and move toward an equilibrium state by diffusing its polarization through the various mechanisms described earlier (chemical exchange, homonuclear dipolar interaction, NOE transfer etc.). A proton-proton mixing period is used in these experiments to allow for this transfer to occur. At the same time, as the surface of the protein is polarized, the magnetization can diffuse further into the buried core of the protein through proton-proton spin diffusion. These main components, depicted in Figure 1-3, form the basis for all experiments used in this work, and will be referred to as “water-edited” experiments throughout. This approach has also been used previously by Etzkorn et al. in Ref [51] to distinguish signals of mobile, static, and water-exposed protein segments in
order to investigate the structure and topology of a seven-helical membrane protein, sensory rhodopsin II from *Natronomonas pharaonis* (NpSRII).\(^{51}\)

![Diagram of Water-edited Experiments](image)

**Figure 1-3:** “Water-edited” experiments in this work are made up of these three basic elements: preparation (a \(T_2\)-filter to select for mobile species), mixing (variable proton-proton mixing period for signal transfer), and detection (includes transfer of signal from proton to carbon or nitrogen nuclei followed by signal detection). In experiments with short mixing, we expect to see stronger signal from protein-solvent interface regions after the detection period.

At short mixing times, before the internal diffusion process has occurred, we expect to see primarily the exposed protein-solvent interface regions appear in a spectrum. If proton-proton spin diffusion happens much faster than polarization transfer across the protein-water interface, then two protons, one on the interface, the second coupled through space to the first, could not be distinguished, as they would obtain polarization on the same timescale. In a three proton case, with two protons coupled through space to the one on the interface, the internal diffusion process would be slowed, but protons would still be indistinguishable. With the addition of more internal protons, the number of relayed steps increases, and thus the diffusion process through the protein is significantly decreased. In principle, it is possible to differentiate between the exposed and internal residues based on the transfer rates, provided that the interaction across the interface occurs faster than the interaction between the protein protons. Our goal is to measure the buildup rates of residues in the protein to investigate which residues obtain polarization faster than others, and whether these correlate with regions which are exposed to solvent.
1.5 Summary of Goals

The most intuitive description of the protein-water interface in ASR includes exposed residues found in loop regions between helices, which can be detected in NMR experiments. In addition, there are internal immobilized water molecules found in the interior portion of the protein, which may be hydrogen-bonded to various polar residues in the protein core. These water molecules create another type of protein-water interface. Although these immobilized waters have a short $T_2^*$ (effective transverse relaxation time constant) and are expected to be suppressed by the $T_2$-filter, they could still rapidly exchange with bulk mobile water. Certain residues (e.g., Asp217, Ser86, Thr90, Lys210) are known to be interacting with internal water molecules through hydrogen bonds. In addition, residues like Lys, Arg, His, are expected to form hydrogen bonds, so they could be bound to internal water molecules as well. With this approach, our goal is to detect any interactions between the protein-water interface of these immobilized water molecules and internal “buried” residue groups.

In this thesis, we use a variety of spectroscopic techniques to investigate protein-water interactions in ASR. One-dimensional spectroscopy is used to observe which types of nuclei and residues obtain their signal faster in a water-edited experiment, and gain a sense of the timescales involved in the relevant interactions. These experiments are also used to disentangle transfer mechanisms which contribute to signal buildup (i.e., chemical exchange or dipolar interactions and spin diffusion). We attempt to determine which mobile protons are correlated to nuclei in the protein, and to observe potential interactions between lipids and protein. Two-dimensional spectroscopy is used to site-specifically observe residues that are located within the ‘first layer’ of the protein-water interface, and therefore obtain their polarization first. These residues are correlated with solvent accessible regions of the protein. The interactions between bound water
molecules and internal residues buried in the protein core are observed through detection of their sidechain signal buildup. Finally, we attempt to site-specifically quantify the signal buildup rates of each residue in ASR using these water-edited experiments.
Chapter 2
NMR Theory

2.1 Static NMR Background

NMR Spectroscopy provides a means of studying molecular structure, by using the quantum mechanical property inherent to all nuclear species, spin. This spin property can be manipulated and investigated in the presence of a magnetic field. Any system considered in NMR studies is governed by both internal interactions, and externally applied magnetic fields. The Zeeman interaction is one induced by an external magnetic field, while additional interactions include the chemical shift, dipolar, J-coupling, and quadrupolar interactions. To include these, we can describe the system by a Hamiltonian of the following general form, with each term described in more detail in the following sections:

\[
H = H_Z + H_{CS} + H_d + H_J + H_Q + H_{rf}
\]

(2.1)

For studies in SSNMR, most biological systems, and all those considered in this work, are isotopically-labelled with $^{13}\text{C}$ and $^{15}\text{N}$. These nuclear species, along with $^1\text{H}$, are spin ½ nuclei, so the quadrupolar interaction, which is only present in nuclei with spins greater than ½, need not be considered further. At common laboratory fields, all remaining internal interactions are at least three orders of magnitude smaller than the Zeeman interaction, which is on the order of hundreds of MHz, so only the parts which commute with the Zeeman field, known as secular components, are considered further. The following is a brief discussion and overview of these interactions, and other relevant aspects of SSNMR theory. For a more complete introduction, the reader should consult additional references, such as *Spin Dynamics* by Levitt\textsuperscript{52}, *Protein NMR Spectroscopy* by Cavanagh *et al*\textsuperscript{53}, and *Principles of Magnetic Resonance* by Slichter\textsuperscript{54}.
2.1.1 Zeeman Interaction

When a spin ½ nucleus is placed into a static magnetic field, the spin angular momentum vector tends to align itself with the magnetic field. When measured, two distinct energy levels are observed, corresponding to a positive and negative projection along the magnetic field, called “spin-up” and “spin-down” states, denoted as $|\alpha\rangle$ and $|\beta\rangle$ respectively. This splitting of the energy levels is called the Zeeman effect, and the spin states correspond to eigenfunctions of the Zeeman interaction Hamiltonian, which is expressed in frequency units:

$$H_Z = -\gamma B_o I_z$$  (2.2)

Here, we have chosen the frame in which the external magnetic field $B_o$ is aligned with the $z$-axis of the lab frame, $\vec{B} = (0,0,B_o)$ by convention, so $I_z$ is the spin $\frac{1}{2}$ operator, and $\gamma$ is the gyromagnetic ratio. The gyromagnetic ratio is a quantity which is unique to each type of nucleus, and is the ratio between the magnetic dipole moment and spin angular momentum of a nucleus. The eigenstates of the Hamiltonian defined by Eq. (2.2) have energies given by:

$$E_\alpha = -\frac{1}{2}\gamma B_o$$

$$E_\beta = \frac{1}{2}\gamma B_o$$  (2.3)

The energy difference between these two eigenstates of the Zeeman Hamiltonian is then given by the Larmor frequency:

$$\omega_o = -\gamma B_o$$  (2.4)

This quantity is different for each type of nucleus, since it depends on the gyromagnetic ratio. Transitions can be driven between these two energy states by applying a time-dependent magnetic field with a frequency which matches the energy difference between them, or their
resonant frequency. This is the basis for all NMR experiments. By convention, NMR spectrometers are referred to by the $^1$H Larmor frequency induced by its static magnetic field. For example, a spectrometer with a magnetic field of 18.8 T corresponds to a $^1$H Larmor frequency of 800 MHz (in angular frequency units).

When an ensemble of spins is placed into this static magnetic field, the Zeeman energy eigenstates are populated according to the Boltzmann distribution:

$$\frac{N_\alpha}{N_\beta} = \frac{\Delta E}{k_B T} \approx 1 + \frac{\gamma B_0}{k_B T}$$

This small population difference, where $k_B$ is the Boltzmann constant and T the temperature in Kelvin, creates a net magnetization aligned with the magnetic field. Thermal energy dominates the exponent at any temperatures higher than a fraction of a Kelvin. In this work, experiments are performed at temperatures ~278 K, so a “high temperature” approximation is used. With this approximation, and where N is the total number of spins:

$$\mathbf{M} = (0,0,M_o)$$

$$M_o = \frac{\gamma}{2} (N_\alpha - N_\beta) = \frac{N \gamma^2 B_0}{4 k_B T}$$

It can be seen that the magnetization created in a system by an external magnetic field depends on the total number of nuclei, the temperature, and the gyromagnetic ratio (which is nucleus-specific) of the system, as well as the strength of the field. The population difference in states is extremely small (on the order of $10^{-5}$ for $^1$H nuclei, which are the most abundant and have the highest gyromagnetic ratio among nuclear spins in biological samples), thus the net magnetization created is also a small quantity, but is still large enough to be distinctly probed for various nuclear species.
2.1.2 Chemical Shift Interaction

The chemical shift interaction refers to a shift in the Larmor frequency of a nucleus, caused by the electrons in its local environment. The external magnetic field perturbs electronic orbital motions, which then generate a secondary local magnetic field. The Larmor frequency of a nucleus depends on both the static external magnetic field, as well as this local magnetic field. As a result, depending on their local electron density distribution, otherwise identical nuclear species will resonate at slightly different frequencies, making them identifiable from one another. There are isotropic and anisotropic components of this interaction.

The “chemical shift” refers to the isotropic component, which describes the difference between the resonant frequency of a specific nucleus, and that of the same nucleus with no electronic shielding, and is reported in units of ppm (parts per million) when compared to a reference standard:

\[
\nu_{CS}[ppm] = \frac{\nu - \nu_{ref}}{\nu_{ref}} \times 10^6
\]

(2.8)

The reference standard used in this work is the \(^1\)H resonant shift of 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS).\(^{55}\) The \(^{13}\)C and \(^{15}\)N shifts are indirectly referenced to DSS, using tabulated gamma ratios.\(^{56,57}\)

The chemical shift interaction can then be described by the following general Hamiltonian:

\[
H_{CS} = (\omega_{ISO} + \omega_{CSA})I_x
\]

(2.9)

where \(\omega_{ISO}\) is the isotropic contribution which is not orientation dependent, while \(\omega_{CSA}\) represents the anisotropic contribution, which is dependent on the orientation of the molecule with respect to the external magnetic field. This orientation dependence, or anisotropy, is
averaged out by tumbling in solution state NMR, but in the case of solids, causes complicated lineshapes and a large amount of line broadening. In addition, in the case of large biomolecules, many regions of a sample have a similar electronic environment, leading to similar chemical shifts which are difficult to resolve. The chemical shift anisotropy is commonly removed in solids by Magic Angle Spinning (as discussed in Section 2.2).

### 2.1.3 Dipolar Interaction

The dipolar interaction refers to the through-space coupling between two spins, and can either be between two spins of the same nuclear species (homonuclear dipolar interaction), or two spins of different species (heteronuclear dipolar interaction). This coupling is represented in its general form by the following Hamiltonian (where \( j,k \) denote any generic spins in the system, and \( \gamma_i \) is the gyromagnetic ratio of the \( i^{th} \) spin):

\[
H_d = \sum_{j<k}^{N} \frac{\mu_0 \gamma_j \gamma_k}{4\pi r_{jk}^3} \left( \vec{I}_j \cdot \vec{I}_k - 3 \frac{\vec{I}_j \cdot \vec{r}_{jk}}{r_{jk}^2} \frac{\vec{I}_k \cdot \vec{r}_{jk}}{r_{jk}^2} \right)
\]

But by neglecting terms which do not commute with the Zeeman field, the expression for the truncated form of the heteronuclear dipolar interaction becomes:

\[
H_d = \sum_{j<k}^{N} \frac{\mu_0 \gamma_j \gamma_k}{4\pi r_{jk}^3} \left( 1 - 3 \cos^2 \theta_{jk} \right)(I_{jz}S_{kz})
\]

where \( \vec{r}_{jk} \) is the internuclear vector between the nuclei at \( j \) and \( k \), and \( \theta_{jk} \) is the angle between this internuclear vector and the external magnetic field. So, the dipolar interaction depends on both the orientation of the internuclear vector and the internuclear distance between the coupled nuclear spins. Because of this dependence, the dipolar interaction contains valuable structural
information about the system. However, this interaction is also orientation dependent, which again is averaged out through molecular tumbling in solution NMR studies. But in the case of solids where this averaging is absent, the dipolar interaction term contributes to significant line broadening and therefore extremely overlapped spectra.

2.1.4 J-Coupling

J-Coupling, also referred to as the J-interaction, or scalar coupling, denotes another type of coupling between two spins, but in this case is indirectly mediated by the bonding electrons. The homonuclear J-coupling term has the following form:

\[ H_J = 2\pi J \vec{I}_j \cdot \vec{I}_k \]  

(2.12)

This reduces to the following truncated form, after the secular approximation is made:

\[ H_J = 2\pi J l_j l_k \]  

(2.13)

for homonuclear and heteronuclear cases respectively. This interaction is very weak, and causes spectral line splitting. In the case of solids, each of the singlets can be broad, so this interaction contributes to the broadening of spectral lines. The J-coupling can be neglected from further discussion presented here.

2.1.5 Applied Radiofrequency Field

In order to excite transitions between the spin-up and spin-down energy states, a time-dependent magnetic field is applied with a frequency that is close to the energy difference between these states. As previously stated, the difference in Zeeman energy states is on the order of tens or hundreds of MHz for nuclei of interest in biological samples, which falls in the
radiofrequency range. These fields are applied in the form of pulses, and can be described in the following general form:

\[ H_{rf} = 2\omega_1 I_x \cos(\omega_{rf} t + \varphi) \]  

(2.14)

where \( \omega_1 \) is the strength of the applied pulse, and is related to the magnetic field as \( \omega_1 = -\gamma B_1 \), while \( \omega_{rf} \) is the frequency, and \( \varphi \) is the phase of this applied pulse. In this way, nuclear spins can be perturbed from their equilibrium states, and thus manipulated in the context of NMR experiments.

### 2.2 Magic Angle Spinning NMR

Unlike in solution state NMR, where the fast molecular tumbling averages out anisotropic interactions, resulting in high resolution spectra, the previously mentioned anisotropic interactions lead to significant line broadening and complicated line shapes in solid state NMR spectra. It is common to average out this anisotropy by physically spinning the sample about an axis tilted at 54.7° (the magic angle) with respect to the external magnetic field, as shown in Figure 2-1.\(^{58}\) This technique is commonly referred to as “magic angle spinning (MAS) NMR”.
Figure 2-1: Orientation of sample rotor, aligned at $\theta=54.7^\circ$ with respect to $B_0$, the external magnetic field. Spinning the sample around this axis with a frequency $\omega_r$ induces time dependence and averages out anisotropic terms from the dipolar and chemical shift interaction terms of the Hamiltonian.

The following is an overview of how MAS removes the effect of the dipolar interaction, and could be derived in a similar way for the chemical shift anisotropy.

An expression for the dipolar interaction can be written as follows:

$$ H = \vec{I} \vec{D} \vec{S} $$  \hspace{1cm} (2.15)

where $I$ and $S$ are two coupled spins, and $D$ is the axially symmetric dipolar coupling tensor, represented in Cartesian coordinates as follows:

$$ \mathcal{D} = \begin{bmatrix} -d & 0 & 0 \\ 0 & -d & 0 \\ 0 & 0 & 2d \end{bmatrix} $$  \hspace{1cm} (2.16)

where $d$ is the dipolar coupling constant between spins. This can also be expressed in general as a product of irreducible spherical, second rank, tensors:

$$ \mathcal{H} = \sum_{L=0}^{2} \sum_{m=-L}^{L} (-1)^m R_{Lm} T_{L,-m} $$  \hspace{1cm} (2.17)

where $R_{Lm}$ is the spatial part of the interaction, and $T_{L,-m}$ is the spin-operator part. Interactions are most conveniently expressed in their principal axis frame (PAS), shown in Figure 2-2, where the
interaction tensors are diagonal, and then transformed into the laboratory frame (LAB). It is ideal to transform the interaction into the laboratory frame, since it is in this frame that the secular approximation with the Zeeman field can be made. This transformation is accomplished by a series of rotations through various frames of reference, as shown in Figure 2-2:

Figure 2-2: Frames of reference used in NMR. Principle Axis System (PAS) is the frame in which the interaction tensor is diagonal. The dipolar interaction tensor is rotated from this frame to the rotor frame (ROTOR) through the Euler angles $\Omega_{PR}$, and finally from the rotor frame to the laboratory frame (LAB) which is defined by the external magnetic field, through the angles $\Omega_{RL}$. The interaction obtains its time-dependence from this final transformation.

These transformations are each accomplished using a different set of Euler angles $\Omega=(\alpha,\beta,\gamma)$ to rotate the tensors between frames, which are defined using the convention shown in Figure 2-3.
Figure 2-3: Definition of Euler angles $\Omega=(\alpha, \beta, \gamma)$ used to accomplish frame transformations. The original axis system is represented in blue with $(x,y,z)$, and the new rotated frame is represented in red with $(x',y',z')$.

The transformation from PAS $\rightarrow$ ROTOR uses Euler angles $\Omega^{PR} = (\alpha, \beta, \gamma)$ which are random variables in a powder sample. The second transformation from ROTOR $\rightarrow$ LAB uses time dependent Euler angles $\Omega^{RL}(t) = (\alpha^{RL}, \beta^{RL}, \gamma^{RL})$. The first term $\alpha^{RL}$ depends on the spinning frequency and on time and is defined as $\alpha^{RL} = a_0 - \omega_r t$, the term $\beta^{RL}$ is the angle between the rotor and the external magnetic field (will be defined later), and $\gamma = 0$ due to the symmetry of the frame.

Spherical tensors transform particularly easily between frames, because these transformations only involve tensors of equal rank, $L$, as follows:

$$R'_{Lm} = \sum_{m'=-L}^{L} R_{Lm} D^L_{mm'}(\alpha, \beta, \gamma) \quad (2.18)$$
where \( \mathcal{D}_{nm'm}^l(\Omega) = d_{nm'm}^l(\beta)e^{-im\alpha}e^{-im\gamma} \) (2.19)
is the Wigner rotation matrix, and is a function of the Euler angles. The dependence on \( \beta \) is described by the reduced Wigner rotation matrix elements \( d_{nm'm}^l(\beta) \) which can be found in Appendix B of *Multidimensional Solid-State NMR and Polymers* by Schmidt-Rohr and Spiess.\(^{59}\)

These equations can be used to express the tensors in the lab frame. First, since the Zeeman interaction dominates the interactions in the lab frame, any components of the dipolar expression which do not commute with the Zeeman term can be discarded. A complete list of the spin-part of the tensor product can be found in Ref [59], but after this approximation is made, the only terms remaining in the lab frame are listed below:

\[
T_{0,0} = \frac{1}{3}(\mathbf{I} \cdot \mathbf{S})
\]

\[
T_{2,0} = \frac{1}{\sqrt{6}}(3l_zs_z - \mathbf{I} \cdot \mathbf{S})
\]

so the dipolar interaction Hamiltonian can be described as:

\[
H = R_{0,0}T_{0,0} + R_{2,0}T_{2,0}
\]

(2.21)

The spatial part of the spherical tensor representing the dipolar interaction in the PAS is referred to as \( \rho_{Lm} \). Based on the form of the dipolar interaction given in Eq. 2.16, \( \rho_{2,0} \) remains as the only non-zero term. The Hamiltonian can be expressed using Eq. 2.17, which can then be expanded using the Euler angle convention:

\[
R_{2,0}^{LAB} = \sum_{m'} \rho_{2,0}^{(2)} d_{0m'}^{(2)}(\Omega) \mathcal{D}_{m'0}^{(2)}(\Omega_{RL})
\]

(2.22)

\[
R_{2,0}^{LAB} = \rho_{2,0}^{(2)} \sum_{m'} d_{0m'}^{(2)}(\beta)e^{-im\gamma}e^{-im\omega_t}t d_{m'0}^{(2)}(\beta_{RL})
\]

(2.23)

which we rewrite as a Fourier series because of the time dependence gained from spinning in the lab frame, at a frequency \( \omega_t \):
Putting this together with the spin-part of the tensor product from Eq 2.19, and factoring the zeroth Fourier component, the complete Hamiltonian can be written as the following:

\[ \mathcal{H} = \frac{R_{2,0}^{LAB}}{\sqrt{6}} \left( 3I_z S_z - \vec{I} \cdot \vec{S} \right) \]

\[ = \frac{1}{\sqrt{6}} \omega_d^{(0)} (3I_z S_z - \vec{I} \cdot \vec{S}) + \frac{1}{\sqrt{6}} \sum_{n \neq 0} \omega_d^{(n)} e^{i n \omega_d t} (3I_z S_z - \vec{I} \cdot \vec{S}) \]  

(2.26)

The non-zero \((n \neq 0)\) Fourier components are time-dependent, and contribute to spinning sidebands, as seen in Figure 2-4 below. The remaining discussion will focus on the first term in Eq. 2.26, the zeroth Fourier component, which is time-independent and therefore cannot be averaged out through spining. Since the goal is to remove the dipolar interaction, this term can be expanded using Eq 2.17, and using Ref [59]:

\[ \omega_d^{(0)} = \rho_{2,0} d_{00}^{(2)} (\beta^{PR}) d_{00}^{(2)} (\beta^{RL}) \]

\[ \omega_d^{(0)} = \rho_{2,0} \frac{1}{2} (3 \cos^2 \beta^{PR} - 1) \frac{1}{2} (3 \cos^2 \beta^{RL} - 1) \]

(2.27)

where \(\beta^{RL}\) again describes the angle between the \(z\)-axis of the rotor and the lab frame. Clearly, this entire term will go to zero if \(\beta^{RL} = \cos^{-1} \frac{1}{\sqrt{3}} = 54.7^\circ\). This angle is referred to as the “magic angle”. So, by physically aligning the rotation axis at the magic angle with respect to the external field \(B_o\) (which defines the lab frame), this time-independent term of the dipolar interaction Hamiltonian is removed.\(^{60}\)
Figure 2-4: The averaging effect of MAS on the lineshape of a static powder sample which is broad due to the effect of dipolar coupling (0 kHz). As the MAS frequency is increased, the peak splits into a single centreband with several spinning sidebands as the time-independent term of this interaction is. Spectra were produced using the simulation package SPINEVOLUTION\textsuperscript{61} at a $^1$H Larmor frequency of 600 MHz. The dipolar coupling constant was -2.8 kHz, corresponding to a $^{13}$C-$^{13}$C distance of 1.4 Å. Simulation produced by A. Gravelle.\textsuperscript{62}

2.3 NMR Experiments

2.3.1 Signal Detection

As a simple explanation, NMR spectrometers detect nuclear spins in a sample by detecting their precession at their Larmor frequency. Inside the spectrometer, the sample is inside a coil with its spins aligned along the z-axis (direction of the external magnetic field $B_0$) in equilibrium. The coil is aligned perpendicular to this axis (or at the magic angle in the case of MAS NMR). Radiofrequency pulses are applied with $B_1$ being along the coil axis, so nuclear spins can be tilted off the z-axis in such a way to create polarization in the transverse plane. Subsequently, when the pulse is removed, the spins will begin to precess in the transverse plane.
with a frequency equal to that of their intrinsic Larmor frequency. This phenomenon is known as “Larmor precession”. Because the nuclear magnetic moments rotate, they generate a time-dependent magnetic flux, which induces a small, but detectable, electric current in the coil.

It is the transverse magnetization components that are measured as a function of time; this signal is referred to as the free induction decay (FID). This oscillating signal in the time dimension can be Fourier Transformed and results in a peak at the resonant, or Larmor, frequency in the frequency domain. Since each spin species has a slightly unique Larmor frequency as a result of the chemical shift, this leads to resolvable peaks dispersed over a range of frequencies.

2.3.2 Cross-Polarization and Multidimensional Spectroscopy

Despite the dispersion of peaks due to unique chemical shifts, as well as line narrowing from MAS experiments, spectra are often extremely crowded and overlapped with many degenerate resonances, especially in the case of solids. A commonly used method to further disperse degeneracies in spectra is to perform a multidimensional experiment. Typically, these types of experiments are accomplished by correlating the chemical shifts of two or more spins, as shown in Figure 2-5.
Figure 2-5: Multidimensional spectroscopy: demonstration of a 1D carbon spectrum of UCN-ASR, with aliphatic carbons then being correlated with nitrogen chemical shifts in a 2D NCA spectrum.

A common way to excite and transfer polarization between two nuclear species, now known as cross-polarization (CP), was first introduced by Hartmann and Hahn. The CP technique encompasses, and is an integral part of, a large body of experiments and applications. This was applied by Pines, Gibby and Waugh to enhance the signal obtained from detecting rare spins like carbon or nitrogen, by first transferring polarization from abundant spins, such as protons. Here, the most important uses of cross-polarization are two-fold: first, for signal enhancement in nuclei with a low-gyromagnetic ratio, and as a mixing period in multidimensional experiments for spectral dispersion. Signal transfer is accomplished under MAS conditions by applying rf pulses to the nuclei, denoted I and S, simultaneously so that the Hartmann Hahn condition ($\omega_I \pm \omega_S = n\omega_r$) is satisfied. This technique is specific enough to be used for frequency-selective polarization transfer, for example, to perform transfer from nitrogen atoms to only alpha-carbons (CA), or only carbonyl atoms (CO).
Two-dimensional spectroscopy can be illustrated using a chemical shift correlation experiment, in which nitrogen chemical shifts are correlated with those of alpha-carbon nuclei (2D NCA).

Figure 2-6: A typical 2D NCA pulse sequence. Polarization is transferred via CP from proton spins to nitrogen spins, chemical shifts are recorded, and then polarization is again transferred to carbon spins where it is directly detected. The result is a 2D NCA spectrum, correlating nitrogen and carbon spins, shown below.

Figure 2-6 shows the 2D NCA pulse sequence used to accomplish this multidimensional experiment. First, nitrogen spins are polarized using proton-nitrogen cross-polarization to create transverse magnetization. These spins are then allowed to evolve freely under the chemical shift interaction (the only relevant interaction present in the system during evolution) for a time period $t_1$ in which the chemical shifts of the spins are recorded in the phase of the magnetization. In the transverse plane, nitrogen spins are mixed with CA spins using nitrogen-carbon CP, to transfer the polarization from one species to the other. Once the CA spins are polarized in the transverse plane as well, they undergo free precession, referred to as $t_2$ evolution, where the FID for the CA nuclei is directly observed. The period $t_1$ is then incrementated, which corresponds to a phase increment in the nitrogen magnetization before transfer to CA, and the experiment is repeated.
After many $t_1$ increments, the data recorded in both time dimensions is Fourier transformed into a 2D frequency spectrum. **Figure 2-7** shows a sample 2D NCA spectrum of GB3 protein, in which peaks are extremely well resolved. Peaks in a 2D spectrum such as this, referred to as crosspeaks, mainly reflect a correlation between directly bonded atoms (nitrogen and $\alpha$-carbons in the case of a 2D NCA experiment) since they are in close spatial proximity and the dipolar interaction between them is strong. Their positions report on the chemical shifts of two nuclei. With each added dimension, spectral resolution is improved as the peaks are dispersed in multiple dimensions.

**Figure 2-7**: 2D NCA sample spectrum of GB3 protein. Peaks correspond mainly to correlations between nitrogen and $\alpha$-carbons (CA) that are directly bonded. Because of the dispersion in two dimensions, peaks are well resolved.
2.4 Relaxation Rates

2.4.1 $T_1$ – Longitudinal Relaxation

A system which is initially at equilibrium in the presence of a static external magnetic field, (according to the Boltzmann distribution across the Zeeman energy levels), can be excited by an applied magnetic field. When this perturbation is removed, the system relaxes back to its initial state of Boltzmann distribution across the Zeeman transition, according to its longitudinal relaxation time constant, or $T_1$. In the case of solids, and specifically membrane proteins, this quantity is generally on the timescale of seconds, and is the biggest contributor to experiment length, since the system must be allowed to return to its initial state of equilibrium before the experiment can be repeated.

2.4.2 $T_2$ and $T_2^*$ – Transverse Relaxation

When a radiofrequency pulse is applied to a system in such a way that the polarization of the spins is rotated by $\frac{\pi}{2}$ radians, the net spin polarization becomes aligned along either the x- or y-axis (an axis perpendicular to the $B_0$ field). When the radiofrequency pulse is turned off, the bulk magnetization vector precesses around the static field at its Larmor frequency.

Due to homogeneous line broadening effects, such as microscopic fluctuations in the magnetic field and molecular motions, the net transverse magnetization will decay over time, according to the transverse relaxation time constant, or $T_2$, of the sample. In solids, coupling between spins (e.g. homonuclear dipolar interaction) provides additional homogeneous broadening. The resulting effective transverse relaxation time constant is referred to as $T_2^*$. In addition to this, there may be inhomogeneous line broadening due to inhomogeneity in the external magnetic field, $B_0$, or local structural heterogeneity of the sample. This causes each spin
in the sample to precess with a slightly different frequency, so the precession frequency of the entire system gradually dephases, resulting in an enhanced decay rate of the transverse signal. Homogeneous and inhomogeneous line broadening effects can be separated by performing a simple echo experiment (first introduced by Hahn$^{68}$) to refocus only the inhomogeneous components. In the following work, the effective or ‘apparent’ transverse relaxation time constant, referred to as $T_2^*$, of various proton species is often considered. Since the inhomogeneous line broadening mechanisms are refocused with the application of an echo (as performed in our experiments), these effects do not need to be considered here.

2.4.3 $T_{1p}$ – Transverse Relaxation in the Rotating Frame

There is one additional time constant, $T_{1p}$, which refers to the characteristic time for relaxation of the net transverse magnetization vector, viewed from the frame of an applied radiofrequency pulse. In NMR experiments, it is often necessary to keep the spin polarization of a specific nucleus aligned with a certain axis, for instance, in the transverse plane for an extended time. This is accomplished by applying an rf pulse in the direction of the desired magnetization for a long duration, referred to as a “spin-lock” pulse, since it suppresses free evolution, locking the magnetization along a particular axis. Within the frame of this time-dependent applied field, known as the “rotating frame”, the magnitude of the transverse magnetization undergoes relaxation, and appears to decay, according to the same mechanisms described above. This quantity is particularly important in cross-polarization experiments, because if $T_{1p}$ is too short, magnetization will disappear before it is further utilized – either detected or transferred to other nuclei. In this work, $T_{1p}$ is analyzed in the context of Lee-Goldburg decoupling, discussed in Section 3.2.4, which is the form of the spin-lock pulse applied to the system.
Chapter 3
Experimental Methods

3.1 Sample Preparation

3.1.1 Materials

Common chemicals used for preparation of proteins were purchased from Fisher-Scientific (Unionville, ON) or from Sigma-Aldrich (Oakville, ON). Cambridge Isotope Laboratories (Andover, MA) provided isotopically labelled compounds, such as $^{13}\text{C}_6$-labeled glucose and $^{15}\text{NH}_4\text{Cl}$. Lipids were purchased from Avanti Polar Lipids (Alabaster, AL).

3.1.2 Preparation of GB3

The expression, purification, and crystallization of a uniformly $^{13}\text{C}$, $^{15}\text{N}$-labeled sample (UCN) of GB3 was performed by Andrew Gravelle at the University of Guelph. Using a UCN sample allows for almost complete resolution of all residues.

3.1.3 Preparation of ASR

The expression, purification, and reconstitution of a uniformly $^{13}\text{C}$- and $^{15}\text{N}$-labeled sample of ASR (UCN-ASR) was performed by Dr. Izuru Kawamura at the University of Guelph, as described previously. ASR was reconstituted in DMPC/DMPA (Avanti Polar Lipids, Alabaster, AL) liposomes (9:1 w/w) at a protein:l lipid ratio of 2:1 (w/w). After lipid-reconstitution, approximately 11 mg of the sample was center-packed into a 3.2mm MAS thin wall rotor.
3.2 SSNMR Experimental Setup

Experiments on GB3 were performed on a Bruker Avance III 600 MHz spectrometer, using a 3.2mm triple channel $^1\text{H}$-$^{13}\text{C}$-$^{15}\text{N}$ E-free probe. Experiments on UCN-ASR were performed on a Bruker Avance III 800 MHz spectrometer, using a 3.2mm triple channel $^1\text{H}$-$^{13}\text{C}$-$^{15}\text{N}$ E-free probe. The MAS rate was set and controlled using a Bruker MAS control unit. The sample temperature was set to 271 K for all experiments, which results in an effective temperature of the sample environment of 278 K, after heating of the sample due to MAS is accounted for.

3.2.1 $T_2$ – Filter

A simple $T_2$ filter was often inserted on the proton channel at the beginning of pulse programs, prior to the initial CP step. This filter (as described by Harbison et al.\textsuperscript{41}) is based on a Hahn echo sequence\textsuperscript{68} and consists of a $\pi$-pulse centered between two equivalent delays, in this case referred to as $t_{\text{echo}}$. This $t_{\text{echo}}$-$\pi$-$t_{\text{echo}}$ period, when $t_{\text{echo}}$ is on the order of milliseconds, selects for any proton magnetization which has a long transverse relaxation time. During this delay, the proton magnetization dephases according to its transverse relaxation time ($T_2$). Since this is very short in solids, most protons found within the protein will decay more rapidly than the timescale of the echo delay, and therefore, the signal from these species is destroyed prior to transfer to other detectable nuclear species, such as carbon or nitrogen. The filter selects for protons which have a long $T_2^*$ relaxation time, including water protons, lipid protons, and any other mobile protons. This is what accomplishes the “water-edited” experimental preparation discussed previously.
3.2.2 N-CA Chemical Shift Correlation Experiments

Site-specific measurements of the buildup rate of signal in the protein sample is accomplished by using a 2D N-CA correlation experiment, with a T2-filter inserted at the beginning of the sequence, shown in Figure 3-1. In this way, only magnetization from protons with a long relaxation time, including water protons, will be retained, as any initial protein signal is destroyed as a result of dephasing. A mixing delay is inserted prior to the $^1$H/$^{15}$N CP step, with a $\frac{\pi}{2}$ pulse which flips proton magnetization to the z-axis, where it is stored during the mixing delay and then returned to the transverse plane with another $\frac{\pi}{2}$ pulse. This step allows for magnetization to be transferred (through either spin diffusion or chemical exchange mechanisms, discussed above) from mobile water protons to amide or other protein protons. Thus, after $^1$H/$^{15}$N and $^{15}$N/$^{13}$C CP transfer, any signal detected in the N-CA plane must originate from mobile protons, and the strength of this signal will depend on the length of mixing time.

Experiments were performed at 14.3 kHz spinning frequency. Typical $\pi/2$ pulses were 2.5μs for $^1$H, 5μs for $^{13}$C, and 7μs for $^{15}$N. For $^1$H/$^{15}$N cross-polarization (CP), contact time was 2ms, with a constant radio-frequency (RF) field of 37.5 kHz applied on nitrogen, while the proton lock field was linearly ramped around the n=1 Hartman-Hahn condition, using a 5% ramp (optimized experimentally). $^{15}$N/$^{13}$Cα band selective CP transfer was accomplished with 5ms contact time, with a constant RF field of $2.5 \cdot v_t$ (where $v_t = \omega_t/2\pi$, spinning frequency) applied on nitrogen, and with a field ramped linearly (10% ramp) around $1.5 \cdot v_t$ applied on carbon, corresponding again to the n=1 Hartman-Hahn condition. Optimized CW decoupling of 100kHz was applied during this CP step, while optimized SPINAL64 of ~83 kHz strength decoupling was applied during indirect and direct chemical shift evolutions. The length of the T2-filter delay
was chosen to be $t_{\text{echo}}=2.0975\text{ms}$, for rotor-synchronization, and after verifying the complete removal of all protein signal. The proton mixing delay, $t_{\text{mix}}$, was varied from 0 ms to 100 ms, at which point, a steady state was achieved for signal buildup.

Figure 3-1: 2D N-CA correlation pulse sequence, with a $T_2$-filter ($t_{\text{echo}}, \pi - t_{\text{echo}}$) and a $^1\text{H} - ^1\text{H}$ mixing delay ($t_{\text{mix}}$) inserted prior to the initial H-N CP step, used to measure signal build up rates. As in Figure 1-3, the preparation step destroys any protein proton signal and selects only mobile protons; in the mixing step polarization is transferred to protein protons through the various mechanisms previously described in Section 1.1, and in the detection step this polarization is further transferred to N and then CA for site-specific detection. Narrow and wide solid rectangles represent $\frac{\pi}{2}$ and $\pi$ pulses, respectively. The pulse sequence was phase cycled as follows: $\phi_1 = y, -y; \phi_2 = x; \phi_3 = y; \phi_4 = x, x, y, y, -x, -x, -y, -y; \phi_{\text{REC}} = x, x, y, y, -x, -x, -y, -y$.

3.2.3 C-C Chemical Shift Correlation Experiments

A series of 2D carbon-carbon (C-C) chemical shift correlation experiments with a $T_2$-filter inserted at the beginning of the pulse sequence was used for a site-specific investigation of the interaction of amino acid sidechains with protons, using the pulse sequence given in Figure 3-2. As before, only magnetization originating from protons with a long $T_2$ relaxation rate is retained following the $T_2$-filter. It is transferred to protein protons during the mixing period and then correlated with carbon spins. After $^1\text{H} / ^13\text{C}$ CP transfer, and $t_1$ chemical shift evolution of carbons, a mixing time is inserted (denoted $t_{\text{DARR}}$) which allows magnetization to be transferred
between carbons using the dipolar-assisted rotational resonance (DARR) method\textsuperscript{70}, before direct \(t_2\) detection. Thus, in this experiment, signal will be seen wherever a water proton, which survives the \(T_2\) filter, is in close proximity to carbon nuclei.

![Figure 3-2: Pulse sequence for 2D C-C correlation experiments. The length of \(T_1\) filter, \(t_{\text{echo}}\) was set to 2ms, while proton-proton mixing time was varied for each experiment. Length of DARR carbon-carbon mixing, \(t_{\text{DARR}}\) was set to 30 ms. The echo just prior to carbon detection is to improve the baseline. Narrow and wide solid rectangles represent \(\frac{\pi}{2}\) and \(\pi\) pulses, respectively. The pulse sequence was phase cycled according to the following: \(\varphi_1 = y, -y\); \(\varphi_2 = x\); \(\varphi_3 = y\); \(\varphi_4 = x, x, x, y, y, y, y\); \(\varphi_5 = y, y, -y, -y, -x, -x, x, x\); \(\varphi_6 = y, y, y, -x, -x, -x, -x, -x\); \(\varphi_{\text{REC}} = x, x, x, x, -x, -y, -y, -y, -y\).

Experimental pulse lengths were the same as described in Section 3.2.2, with a \(^1\text{H}/\text{C}\) CP contact time of 2ms, and a 10\% linear ramp on \(^1\text{H}\). The length of DARR mixing, \(t_{\text{DARR}}\) was set to 30 ms, while the length of the echo period on carbon, prior to detection, was 1.40 ms (equivalent to two rotor periods). Spectra were obtained for 2 different \(^1\text{H}-\text{H}\) mixing times, 10 ms and 50 ms.

3.2.4 Spin Lock at the Magic Angle

In both \(T_2\)-filtered NCA and C-C correlation experiments discussed in the previous sections, the proton polarization transfer across the protein-water interface can be either dipolar driven through-space, or exchange based. To differentiate between these two mechanisms, we
used proton mixing while keeping the proton magnetization locked at the magic angle (Lee-Goldburg decoupling).

It was shown by Lee and Goldburg that certain terms from the dipolar interaction Hamiltonian can be averaged out, through the application of off-resonance pulses. If the applied RF pulse offset (denoted $\Delta$) is chosen in such a way that

$$\sqrt{2} = \frac{\omega_1}{\Delta}$$  \hspace{1cm} (3.1)

then the effective field created by this pulse is tilted at an angle of $54.7^\circ$ with respect to the external magnetic field, $B_0$, as shown in Figure 3-3.

![Figure 3-3: Orientation of Lee-Goldburg irradiation (referred to as a spin-lock at the magic angle), applied off-resonance, with $\omega_1$ denoting the RF frequency, and $\Delta$ the RF pulse offset value. This off-resonance pulse creates an effective applied magnetic field, $B_{\text{eff}}$, oriented at the magic angle, $\theta_m=54.7^\circ$, around which axis the spin polarization precesses.](image)

Spins are irradiated with this rf field, referred to as a spin-lock, for a given time, which results in an averaging of the dipolar interaction Hamiltonian, in the tilted rotating frame.

Lee-Goldburg (LG) decoupling removes the homonuclear dipolar interaction in the first order, and significantly slows through-space proton-proton spin diffusion. Thus, any transfer across the protein-water interface would be accomplished by chemical exchange. A one-
A 38-dimensional experimental pulse sequence is shown in **Figure 3-4**. It contains a T$_2$-filter inserted at the beginning of the pulse sequence as the preparation step. Following this filter step, proton spins were locked along an axis oriented at the magic angle, which eliminated homonuclear dipolar transfer between protons and suppressed spin diffusion (as discussed above).

![Figure 3-4](image)

**Figure 3-4**: 1D $^{15}$N- or $^{13}$C- detected pulse program showing T$_2$-filter and then constant off-resonance $^1$H (LG Decoupling) irradiation. Flip angle $\theta_m$ corresponds to the magic angle, 54.7°. Narrow and wide solid rectangles represent $\frac{\pi}{2}$ and $\pi$ pulses, respectively. Phase cycling was as follows: $\varphi_1 = y$; $\varphi_2 = x$; $\varphi_3 = -y$, $y$; $\varphi_4 = x$, $x$, $y$, $-x$, $-y$, $-y$; $\varphi_{REC} = x$, $-x$, $y$, $-y$, $-x$, $x$, $-y$, $y$.

Experimental pulse lengths and power levels were the same as in Section 3.2.2, with a few exceptions. For $^1$H/$^{13}$C CP, contact time was 2ms, with a constant RF field of 50 kHz applied on carbon, with a 5% linear ramp around the n=1 Hartmann-Hahn condition. A block of two pulses, a magic angle prepulse corresponding to a tilt angle of 54.7° (denoted $\theta_m$ in **Figure 3-4**) immediately followed by a $\pi/2$ pulse applied to the proton channel align the magnetization with the magic angle axis, and perform spin temperature alteration. During the mixing time, a LG off-resonance irradiation was applied at an RF intensity of 80 kHz, with an off-resonance value of 56.64 kHz. The length of LG decoupling was chosen to be $t_{lock}=10$ms. Proton magnetization was then returned to the transverse plane by a $\pi/2-\theta_m$ pulse before performing CP to either $^{13}$C or $^{15}$N for direct detection.
3.2.5 H-(N)-CA Correlation Experiments

2D proton-carbon chemical shift correlation experiments, with a T$_2$-filter inserted at the beginning of the sequence, were used to investigate which proton species were contributing to the protein signal after transfer, using the pulse sequence shown in Figure 3-5a). As usual in these water-edited experiments, only proton signals with a long transverse relaxation rate are retained, and in this type of experiment, these protons are then detected. After a variable $^1$H-$^1$H mixing time, proton polarization is then transferred to carbon nuclei by a $^1$H/$^{13}$C CP step, before direct $t_2$ detection.
Figure 3-5: Pulse sequences for a) 2D H-C and b) 2D H-(N)-CA experiments. Proton detection occurs immediately following the T2-filter, prior to 1H-1H mixing. Narrow and wide solid rectangles represent \( \frac{\pi}{2} \) and \( \pi \) pulses, respectively. Phase cycling was performed as follows: a) \( \varphi_1 = y, y, -y, -y; \varphi_2 = x; \varphi_3 = y, -y; \varphi_4 = y; \varphi_5 = x, x, x, x, y, y, y, y; \varphi_6 = y, y, y, y, -x, -x, -x, -x; \varphi_{REC} = x, -x, -x, -x, -y, -y, -y, -y. \) b) \( \varphi_1 = y, y, -y, -y; \varphi_2 = x; \varphi_3 = y, -y; \varphi_4 = y; \varphi_5 = x, x, x, x, y, y, y, y; \varphi_{REC} = x, -x, -x, -x, -y, -y, -y, -y. \)

Figure 3-5b) shows a similar carbon-detected pulse sequence, but with signal being transferred to nitrogen prior to carbon nuclei. This is accomplished as usual, with additional \(^1H/^{15}N\) and then \(^{15}N/^{13}C\) CP steps inserted prior to direct detection of carbon nuclei, as before.
Experimental pulse lengths were the same as described in Section 3.2.2, with a $^1\text{H}/^{13}\text{C}$, $^1\text{H}/^{15}\text{N}$, and $^{15}\text{N}/^{13}\text{C}$ CP contact times of 2ms, 2ms and 5 ms, respectively, with a 10% linear ramp on $^1\text{H}$, and a 5% ramp on $^{13}\text{C}$. Spectra were obtained for short $^1\text{H}-^1\text{H}$ mixing times of 10 ms only.

With these experiments, signal will be seen for protons which have survived the $T_2$-filter, and at carbon regions which are in close proximity to these protons. The additional nitrogen transfer is performed to discriminate between lipid and protein resonances.
Chapter 4
Results and Discussion

4.1 Investigation of the Protein-Water Interface using 1D Spectroscopy

We begin by describing various proton species in the ASR sample and their dynamic properties. A typical ASR proton spectrum under MAS conditions is shown in Figure 4-2. Water protons are the most abundant, and therefore have the largest signal at ~5 ppm in a spectrum. This includes both mobile water protons (bulk water), and may also include immobilized water protons (bound water), likely buried under the strong bulk water line. In addition, there are protons found within the lipid membranes, which produce distinct peaks in the spectrum.\textsuperscript{72,73} Figure 4-1 shows the structure of DMPC lipids used in reconstitution of the ASR samples, with headgroup protons labeled.

![Figure 4-1: Structure of DMPC lipid reproduced from Prosser et al.,\textsuperscript{74} with $\alpha,\beta,\gamma$ protons on the headgroup labeled according to the nomenclature of Brown & Seelig.\textsuperscript{75}]

Finally, there are a large number of protons that are part of the protein itself. These include amide protons (bonded to the backbone nitrogens), alpha-protons (bonded to alpha-carbons) and others bonded to the sidechains. Depending on the types of bonds involved, some of these protons may be mobile (for example: those on the sidechains of some polar residues), and can
further undergo chemical exchange with water protons, while others are tightly covalently bound and are much less mobile as a result. In immobilized molecules such as ASR, protein protons are generally unresolvable because of the strong homonuclear dipolar couplings between them, so they appear as a low, broad region in a typical 1D $^1$H spectrum.

Typical bulk water appears in solid state spectra around 4.7 ppm, but this value is shifted according to the sample conditions, like temperature and pH, of the sample. Also, the difference in magnetic susceptibility of the rotor material and water as well as between the lipid membrane and water leads to a local magnetic field, which can further affect the line shape and shift the water peak from its expected position. For the UCN-ASR sample used in this work, the water peak appears at 5.03 ppm in each spectrum.

![Figure 4-2: 1D $^1$H spectrum of UCN-ASR. Different types of proton lines are visible, with the strongest centered at 5.03 ppm belonging to bulk water. Lines at 3.7 and 3.2 ppm belong to H$\beta$ and H$\gamma$ of DMPC lipid as seen in Figure 4-1, while the broad region centered ~1.3 ppm is made up of sidechain protein proton resonances, and is extremely broad due to strong homonuclear coupling interactions. Spectrum obtained at 14.3 kHz spinning.](image-url)
These vastly different dynamic properties of the bulk water (and lipids) and the protein protons result in different relaxation parameters for their spins. Specifically, the apparent transverse relaxation, $T_2^*$, is much longer for highly dynamic water and lipid species. This property can be exploited spectroscopically to create a non-equilibrium situation in which only mobile proton species retain their nuclear spin polarization, while the immobile protein proton spins dephase (by contrast, at equilibrium, both types of spins are polarized). Such a state is created by employing a $T_2$ filter, which is often inserted on the proton channel at the beginning of pulse sequences. During this step, all proton magnetization dephases according to its transverse relaxation time ($T_2^*$). As seen in the water-edited spectrum in Figure 4-3, the filter selects for protons which have a long $T_2^*$ relaxation time, including mobile water protons, lipid protons, and any other mobile protons.

**Figure 4-3**: 1D $^1$H spectrum at 14.3 kHz spinning, after the application of a $T_2$-filter, shown in black. The red dotted line is the original proton signal from Figure 4-1. All of the protein proton signal has effectively been removed in the water-edited spectrum. Small peaks remain around 2 and 1 ppm, which could be due to CH$_2$ or H3 groups in the lipid. Much of the bulk water signal remains since it is highly mobile, as well as the mobile lipid resonances. Because such a large fraction of proton magnetization dephases and results in a loss of signal, the intensity of this spectrum has been multiplied by a factor of 10 compared to the red lines.
It can be seen that the DMPC Hβ peak is more attenuated than the Hγ peak. As shown in Figure 4-1, the β protons belong to a CH₂ group in the DMPC headgroup, so they are less mobile than the γ protons which all belong to CH₃ groups coordinated to the nitrogen on the headgroup. By comparing the intensities of the water peak and the Hγ lipid peaks before and after the application of the T₂-filter, the lipid proton T₂ relaxation time constant is estimated to be ~7.7 ms, which is much longer than the filter, allowing these signals to survive.

4.1.1 Nitrogen Spectra

¹⁵N and ¹³C 1D spectroscopy provides a means to obtain general (sometimes even site-specific) information on the timescale and the mechanism of polarization transfer across the interface. Figure 4-4 shows a typical ¹⁵N spectrum of ASR.

![Figure 4-4: 1D ¹⁵N spectrum of UCN-ASR with no T₂-filter to show all residues under normal conditions. Spectrum was taken with 256 scans. Insert shows the unique resonances of His8, His69, and the Schiff Base (SB) forming Lys 210.](image-url)
The majority of signals are from the protein backbone and show in the region of ~110 - 140 ppm. Sharp peaks around ~30 ppm correspond to side chain lysine NZ, while the peaks around ~75 ppm and ~90 ppm are from arginine side chain NH and NE resonances respectively. In addition, there are three well-resolved resonances in the region of 160-180 ppm, which correspond to histidine sidechain nitrogens, and to the Schiff base (SB) forming Lys210 NZ.

An introduction of a T2-filter prior to $^1$H/$^{15}$N CP excitation almost completely suppresses the spectrum, (with the exception of mobile Lys NZ signals around 30 ppm, see below), in agreement with the disappearance of the protein signal in the proton spectra shown in Figure 4-3. As discussed in Section 1.4, an introduction of the proton-proton mixing period allows for the transfer of polarization from water to the protein, either through dipolar mechanisms or by chemical exchange. The buildup of various parts of the $^{15}$N spectrum is shown in Figure 4-5 as a function of mixing time. It appears from this series of spectra that the transfer is nearly complete after about 50 ms for all species. However, some signals build up much faster than others.
Figure 4-5: 1D water edited nitrogen spectra, taken at various $^1$H-$^1$H mixing times, showing the signal buildup of the backbone, Schiff base, and sidechain regions. Lysine sidechain peak intensity is visible in the first spectrum, and remains fairly constant with increasing mixing time. Polar arginine sidechains have fast buildup, while the Schiff Base (SB) and nearby histidine sidechains can be seen to buildup slowly (~35 ms).

In particular, the signals from Arg sidechain nitrogen atoms are visible immediately in the 3 ms mixing spectrum, and build up quicker than other resonances. Arginines are polar residues, found both in the solvent exposed loops (such as R27, R30, R124, R151, and R228), and in the core, where they can form salt bridges with other polar residues, or can be found in contact with water through hydrogen bonds.

As can be seen in the region from 200-150 ppm, assigned histidine Nδ atoms (His8 and H69) and the NZ atom of Schiff Base Lys210 appear in a spectrum with longer mixing of 35 ms. The amide proton on the Schiff Base is known to be exchangeable, and also bound to a water molecule, and thus contributes to signal buildup in the above nitrogen spectrum. As for histidines, they have positively charged sidechains with mobile protons. The resonance values of the nitrogen shifts of His8 and His69 indicate these residues are partially deprotonated in ASR.
The sidechain Lys NZ atom signals resonating around 35 ppm also appear in the first spectrum taken with the shortest mixing, and as the mixing time increases, the intensity of this peak remains the same. These sidechains appear to be very mobile, as attested by small linewidths of their NZ resonances and relatively long transverse relaxation, which allows them to survive the $T_2$ filter, making the sidechain nitrogen peak appear at relatively full intensity in the first spectrum.

While it is quite obvious from the 1D water-edited $^{15}\text{N}$ detected experiments that certain side chain regions build up their signal at a faster rate than the backbone, the mechanism by which this signal is transferred from water to the protein is unclear. There are two major mechanisms that can result in transfer on a millisecond timescale: proton-proton spin diffusion, and chemical exchange between exchangeable protein-protons and water-protons. To disentangle these two mechanisms, the water edited $^{15}\text{N}$-detected experiment was modified to include Lee-Goldburg (LG) homonuclear decoupling during proton-proton mixing. As explained in Section 3.2.4, this decoupling removes interactions between protons, so any signal that appears in the spectrum will be a result of chemical exchange between water protons and protein protons. Thus, by introducing homonuclear decoupling, we can separate the effects of chemical exchange from dipole-dipole interactions in ASR.

One possible complication may arise from the fact that during the applied decoupling spin-lock pulse, proton magnetization will decay according to its transverse relaxation time constant, $T_{1\rho}$. Experimentally estimated values of $T_{1\rho}$ for ASR were in the range of 12-18 ms for $^1\text{H}_2\text{O}$, $^1\text{H}_\text{N}$, and $^1\text{H}_\alpha$. Thus, we chose mixing times under LG conditions to be 10 ms, to ensure that, despite relaxation losses, a large part of the signal is retained (approximately 44%, as estimated by the shortest $T_{1\rho}$ value for $^1\text{H}_\alpha$).
Figure 4-6 shows the $^{15}$N-detected ASR spectrum, recorded with LG mixing, where a number of sharp peaks appear, most notably at a resonance value of ~35 ppm, which corresponds to the chemical shift of the sidechain NZ atoms of lysine residues. Peaks can also be seen at ~75 ppm and ~90 ppm, which both correspond to Arginine sidechain nitrogen resonances, NH and NE nuclei respectively. The broad region ~170 ppm is consistent with the chemical shift region of two previously assigned histidine sidechain nitrogens; Nδ1 of His8 and His69, and of the Schiff Base Lys210, which undergoes exchange in a few milliseconds in homologous proteins. Since LG decoupling experiments eliminate the homonuclear dipolar interaction, only residues which undergo chemical exchange with water will contribute to the signal. So, it appears that protons close to sidechain nitrogen atoms of arginine, lysine, and histidine have undergone chemical exchange with water. These are all residues with positively charged sidechains, indicating that their protons will readily undergo chemical exchange on a fast timescale, or survive the T<sub>2</sub>-filter, due to the mobility of these sidechain protons.

Figure 4-6: UCN-ASR 1D $^{15}$N spectrum taken under Lee-Goldburg decoupling conditions, with a spin-lock of 10 ms and 24576 scans. Spectrum was apodized with 40 Hz of line broadening. The sharp peak ~35 ppm corresponds to Lys NZ, the peaks ~75ppm and ~90ppm are from Arg NH and NE resonances respectively, and the region ~170 ppm could be from His NE2, or from SB which is exchangeable on the timescale of a few milliseconds in homogolous proteins. Apart from these sidechain nitrogens, the weak, broad region around 120 ppm is the buildup of exchangeable amide nitrogen atoms, which is almost completely removed with LG decoupling. Signal intensity is increased by a factor of 64 compared to Figure 4-4.
In addition to these sidechain resonances, a broad band can be seen around 120 ppm, which is mostly comprised of contributions from backbone amide nitrogen resonances. This region is strongly attenuated compared to those of the sidechains, indicating that their protons are buried or protected by either tight interhelical packing, interactions with lipids, or by strong intrahelical hydrogen bonds, making chemical exchange difficult, at least on the timescale of the LG mixing. Hydrogen bonds along the backbone (between backbone nitrogens and carbonyls) are involved in forming the secondary structure of the protein, so it requires higher energy for these amide protons to undergo chemical exchange. Notably, the backbone peaks are also strongly attenuated compared to the intensity in the full excitation nitrogen spectrum, taken without LG decoupling (Figure 4-4). Although amide protons are known to be exchangeable, this indicates that the kinetics of their exchange occur on a longer timescale than the timescale of LG mixing, 10 ms. Overall, these backbone nitrogen resonances are almost completely eliminated in the LG spectrum, since the dipolar transfer mechanism is removed.

### 4.1.2 Carbon Spectra

Additional information on the timescale and the dominant mechanisms of exchange can be obtained using 1D $^{13}$C spectroscopy. We are specifically interested to observe exchangeable side chains, many of which are known to interact with internal waters, or be directly exposed to solvent. A water-edited 1D carbon spectrum taken with a 50 ms $^1$H-$^1$H water-protein mixing is shown in Figure 4-7.
Figure 4-7: 1D water-edited Carbon spectrum of ASR taken with 50ms mixing for signal transfer.

The majority of resonances are seen in this spectrum, with an overall shape resembling that of a spectrum without a $T_2$-filter (shown in Figure 4-9a). Again, because of the lack of site-specific resolution in a one-dimensional spectrum, individual residues cannot be identified, but we can distinguish between different types of carbon nuclei, and determine which types of residues are contributing to the signal as it builds.
By comparing various regions of the carbon spectrum at increasing $^1$H-$^1$H mixing times, as shown in Figure 4-8, we can see certain types of residues obtaining their signal at different rates. At 0 ms mixing, most of the resonances are removed by the T2 filter, with only a few distinguishable peaks in the aliphatic region remaining. These are found at chemical shift positions of 16.2 ppm, ~24.5 ppm, 42 ppm, and 43 ppm. Because these signals appear in the spectrum before mixing has occurred, they must correspond to nuclei in close proximity to protons which are mobile enough to survive the T2-filter. Since lipid protons on the headgroup are highly mobile, it is possible these peaks arise from natural abundance lipid carbons. The small linewidth of the peak at 16.2 ppm (~0.3 ppm), indicates this peak originates from a lipid carbon. Indeed, the shift is consistent with the CH$_3$ group carbon nuclei in DMPC$^{73}$ (taking into account the different referencing standards used in assignments$^{55}$). The other peaks are broader (>0.5 ppm) and their shifts are consistent with those of polar groups with mobile protons, Lys CG and CE, His CB, Arg CD, and Thr CG, so they could originate from these mobile protons in...
the protein. Lysine sidechain protons are clearly sufficiently mobile to survive the T2-filter, as seen in the 1D nitrogen spectra, so it is expected that they also contribute to carbon signal.

Overall, the timescale of the signal buildup profiles of both nitrogen and carbon nuclei are similar, and this timescale is on the order of 50 ms for most nuclei.

A 1D water edited carbon-detected spectrum recorded with an LG spin lock is shown in Figure 4-9(b). A number of peaks appear, indicating the presence of residues with exchangeable, or highly mobile, protons. The overall intensity of the aliphatic region is significantly reduced, apart from these few individual peaks remaining in the spectrum due to chemical exchange. There are a few particularly sharp peaks in the aliphatic region, specifically ~16.5 ppm, ~24.5 ppm, ~42 and 43 ppm, and ~56 ppm. The first three of these sharp peaks are at the same positions as those seen previously in the carbon buildup spectra (Figure 4-8), appearing in the 0ms before $^1$H-$^1$H mixing has occurred. As discussed, these peaks must come from residues with mobile sidechains (Lys, Arg, Thr) since their protons appear to survive the T2-filter. The region from 60 – 75 ppm, primarily Thr CB and CA, and Ser CA resonances, is attenuated compared to the full carbon spectrum (taken under normal conditions) with the removal of the dipolar interaction, indicating that these nuclei primarily obtain their signal by through-space transfer, but the remaining peaks indicate some species are chemically exchangeable as well.

The strong peak ~158-160 ppm is consistent with CZ resonances from the sidechains and aromatic rings of arginines and tyrosine. Although tyrosine sidechains are hydrophobic, they contain an OH group which could be found in contact with water, and the peak in the corresponding spectrum indicates that this group undergoes chemical exchange.
Figure 4-9: a) Reference 1D CP $^{13}$C spectrum of UCN-ASR taken without a $T_2$ filter, where all resonances appear. Spectrum taken with 32 scans. b) Water-edited 1D $^{13}$C spectrum taken with a Lee-Goldburg spin lock mixing period of 10 ms and 8192 scans.

Similar to the nitrogen spectrum, the backbone regions are strongly attenuated in the carbon-detected spectrum taken under LG decoupling, when compared to the full carbon spectrum. Carbonyl (CO) resonances around 180 ppm are much weaker under LG conditions than in the CPMAS experiment, and are also much weaker than the visible sidechain peaks in the LG spectrum.
In summary, the majority of signal has been eliminated by applying LG decoupling, with the loss of signal being far greater than those due to $T_{1\rho}$ effects. This indicates that the dipolar interaction is the largest contributor to the mechanism of signal buildup in the protein. Thus, for signal to be obtained in most areas of the protein after a $T_2$-filter, a mobile water proton must be in close spatial proximity to the residue in order for this signal transfer to occur. Although some evidence for chemical exchange was seen for sidechain atoms, backbone protons were not seen to exchange on the timescale of 10 ms.

4.2 Water-Protein Correlation Experiments

If signal buildup occurs primarily through dipolar transfer and spin diffusion from water in close proximity to regions of the protein, we want to use this to investigate which regions specifically come in contact with mobile species. We want to investigate which residues are indeed close to this water, allowing time for dipolar transfer between proton species to occur, and potentially detect whether lipids are involved in this polarization transfer process.

The $T_2$-filter selects for any species with a long transverse relaxation time constant, which still includes various types of protons (as seen in Figure 4-3). To determine which of these proton species provides the initial source of magnetization for transfer to the protein, we perform additional two-dimensional (2D) experiments where proton chemical shifts are recorded in the indirect dimension. In the following $^1\text{H}-^{13}\text{C}$ chemical shift correlation experiments, magnetization is transferred from various proton species to nearby carbon nuclei. Figure 4-10 shows a 2D water-edited spectrum, recorded with no $^1\text{H}-^1\text{H}$ mixing, so crosspeaks only appear where mobile protons are in close proximity with carbon species, allowing for signal transfer during the CP step.
**Figure 4-10:** 2D H-C chemical shift correlation spectrum, with a T$_2$-filter, but with no $^1$H-$^1$H mixing period, so crosspeaks only come from correlations between carbon atoms in close spatial proximity to mobile protons. Spectrum was recorded with 68 scans and apodized with an exponential window broadening of 100 Hz in both dimensions.

In the proton dimension, lines are fairly narrow with linewidths of ~0.2-0.4 ppm (or ~300 Hz), as the peaks all arise from mobile species with a long T$_2^*$. This allows us to distinguish between various types of protons, and determine to which types of carbon nuclei they are close. As expected, the water proton line at 5.03 ppm forms crosspeaks with several carbon species. In the carbonyl region, correlations are seen with carbons centered around 181 and 179 ppm which could originate from a combination of sidechain resonances of Asp, Asn, Gln and Glu, or protein backbone carbonyl, as these types of carbon nuclei could interact with water. In the aromatic ring region, there are two clear peaks at 160.4 ppm and 158.8 ppm which are consistent with CZ resonances of arginines and tyrosine, indicating their close proximity or exposure to water. It is unexpected that tyrosine sidechain carbons would be in contact with water, as these sidechains are hydrophobic, and are found in helical regions at the membrane surface of ASR. However, they do have an OH group which could be found close to water. At least one tyrosine, Y51
found in helix B, is known to be in contact with bound water. Correlations also appear between water and carbon resonances at 67.7, 63.2, 56, 43, 29.8, 24.4 ppm which are consistent with both sidechain and backbone aliphatic carbons. These aliphatic resonances display relatively broad carbon linewidths, ranging from about 0.5 ppm to 0.8 ppm, indicating that they come from protein carbons in close proximity to water. Since lipid $^{13}$C nuclei are at natural abundance, their peaks do not experience broadening due to scalar couplings, so they would appear narrower than protein carbon peaks.

Apart from the main water line, there are a number of other proton resonances that appear in the spectrum, in the range of 3.8 to 4.3 ppm, which are consistent with shifts of protons found on the headgroup, Hα, Hβ, of the DMPC lipid (as seen in Figure 4-1, and according to the nomenclature of Brown & Seelig$^{75}$). As expected, these proton peaks are correlated with nearby mobile lipid headgroup carbon nuclei (resonating in the range of 56 and 77 ppm). These peaks, particularly those resonating at 56 and 68.4 ppm, display narrow linewidths of about ~0.2 ppm in the carbon dimension, indicating these come from headgroup carbons (β or γ).

There are a few noteworthy sharp peaks which appear at the same location as the peaks in the 1D water-edited carbon spectrum recorded with 0ms mixing (~16.5, ~25, 42 and 43 ppm). These peaks which originate from mobile protons that survive the $T_2$-filter, can now be seen dispersed in the proton dimension. One in particular comes from the Hγ of DMPC lipid, and is correlated to a carbon resonance at ~42.5 ppm, which is not consistent with any carbon shifts of DMPC. However, there is no $^1$H-$^1$H mixing in this spectrum, so it is unexpected that Hγ would form correlations with protein carbons. The other crosspeaks in this area are consistent with shifts of protons from the acyl chains of DMPC, which therefore must be sufficiently mobile to survive the $T_2$-filter.
The result of a water-edited H-C correlation experiment recorded with short (10ms) $^1$H-$^1$H mixing is shown in Figure 4-11. In this spectrum, the lines which survive the filter and have the most signal after short mixing are primarily the strong water line, and the Hγ line from the lipid. There are two additional isolated peaks, likely from sidechain protein Hβ atoms at 2.24 and 2.06 ppm, which are both correlated to carbon nuclei resonances at 24.75 ppm, which are consistent once again with 1D carbon results, and belong to either lysine or threonine CG nuclei, since these are mobile enough to survive the filter.

Correlations corresponding to the water line are very intense, because of its high mobility, and therefore long $T_2^*$. Water forms correlations with many carbon species, as shown by the 1D carbon trace, which resembles a full carbon spectrum (as shown in Figure 4-8a). Correlations from the Hγ line forms significant crosspeaks at the carbonyl, and aliphatic regions, but the aromatic ring regions are too weak to appear in this spectrum.
**Figure 4-11:** 2D water-edited H-C chemical shift correlation spectrum with 10ms $^1$H-$^1$H mixing. Window function is SSB 2.5 in both dimensions, and 84 scans. Cut at 8 sigma. We can see correlations mostly retained from the water peak, and the $H_\gamma$ peak of the DMPC lipid. Inserts are 1D carbon slices taken at water proton line of 5.03 ppm (bottom slice), showing correlations across the whole carbon dimension, and at the $H_\gamma$ proton of DMPC lipid line at 3.33 ppm (top slice) showing much weaker signal, and showing transfer limited to the carbonyl and aliphatic regions of the carbon dimension.

This water-edited spectrum provides a way to detect which mobile proton species contribute to polarization transfer and signal buildup in ASR. Based on the correlations seen above at short mixing, it is clear that the majority of signal originates from mobile water protons, with some contribution from mobile lipid protons found on the headgroup of DMPC ($H_\gamma$).

To further determine where signal comes from, an additional experiment is performed, in which proton chemical shifts are evolved following the $T_2$-filter, then short $^1$H-$^1$H mixing (10 ms) occurs, and polarization is then transferred from these protons to nitrogen nuclei, where it is stored before being transferred in a final CP step to carbon nuclei, which are directly detected.
Figure 4-12: 2D water edited H-(N)-CA chemical shift correlation experiment, with short (10ms) proton-proton mixing. Window function is SSB 2 in both dimensions, and 150 scans. Correlations are retained on two proton lines: at the water line at 5.03 ppm, and at the $H_\gamma$ of DMPC lipid line at 3.33 ppm.

In our lipid-reconstituted samples, some DMPC molecules may be in close spatial proximity to ASR. DMPC lipid headgroups contain nitrogen, which is coordinated by four carbons (three CH$_3$ groups, referred to as $\gamma$, and one CH$_2$ methylene group, referred to as $\beta$). Since the lipid is not isotopically-labeled, detectable species are at natural isotopic abundance, which is 0.36% for $^{15}$N and 1.1% for $^{13}$C.$^{78}$ In this experiment, any correlations seen in the spectrum must involve nitrogen nuclei. The observed signal may come from either lipid-protein correlations or lipid-lipid correlations, although these intralipid correlations will be attenuated by a factor of almost 30000 due to the isotopic abundance of spins. While the large number of lipids in the system (~20 lipids per ASR protein) partially counteracts this loss of signal, it is unlikely that these interactions could be detected. It is also possible to have endogeneous lipids,
originating from the expression of ASR in *Escherichia coli*, bound to the protein, which are UCN labeled. These lipids, while capable of producing strong signals due to their isotopic labeling, are much less numerous than DMPC (or others from reconstitution), so are unlikely to be seen. Thus, it is unlikely that peaks which appear in the spectrum originate from intralipid correlations, but rather indicate lipid-protein correlations.

**Figure 4-12** shows the result of this experiment, with correlations seen in the aliphatic carbon region, at two main proton lines: those of water at 5.03 ppm, and of the Hγ nuclei of the DMPC lipid, as was seen from the 2D spectrum in **Figure 4-11**. During ¹H-¹H mixing, polarization is transferred from Hγ to nearby protons. Those that are close to nitrogen nuclei will transfer their polarization, where it is stored before being transferred again to nearby carbons. Because the frequency of nitrogen in the experiments is ~120 ppm, and nitrogen nuclei in lipid choline groups resonate at about 43 ppm(*), this additional transfer step excludes transfer through the choline headgroup. In addition, protons in the DMPC headgroup (Hα, Hβ or Hγ), are not directly coupled to nitrogen, resulting in even more unlikely transfer through the lipid nitrogen. Correlations are seen in the spectrum between 55-67 ppm in the carbon dimension, which is consistent with many aliphatic protein carbon residues, and indeed some peaks are broad (~0.4 ppm) indicating lipid-protein correlations. Finally, we see more correlations than just the β (CH₂ group) and γ (CH₃ groups) carbon resonances at ~65.5 ppm and ~54 ppm respectively, further indicating that these are indeed due to lipid-protein correlations.

An unambiguous assignment of the specific carbon nuclei which are in close contact with these mobile lipid protons is not possible with the given data, and would require dispersion of spectral peaks in a three-dimensional experiment, but would provide valuable insight into the structural arrangement of lipid-embedded ASR.
4.3 Methodology of Site-Specific Detection

1D spectroscopy provides valuable information on the timescale and mechanism of the water-protein polarization transfer process. It appears that for the backbone amide protons, the process is dominated by through space transfers, at least on a millisecond timescale. For further site-specific investigations of water-protein interactions we resort to two-dimensional (2D) spectroscopy.

Figure 4.13 shows a 2D NCA chemical shift correlation spectrum collected from a UCN ASR sample. In this spectrum, taken without a $T_2$-filter as a reference for peak intensity, 87 residues are resolved of the 206 assigned in a uniformly-labelled sample of ASR. This allows for a site-specific investigation of water-protein interactions.
Figure 4-13: 2D NCA chemical shift correlation spectrum of UCN-ASR taken at 14.3 kHz. This is a reference spectrum, taken without a T$_2$-filter, so all peaks appear. 87 of the 206 assigned residues are resolved in this type of spectrum, forming the basis for the site-specific study of protein-water interactions.

N-CA correlation spectroscopy can be used to detect interactions of amide protons (H$_N$) with protons from the bulk water. This process was previously investigated in ASR using the H/D exchange technique. Briefly, in an H/D experiment, lipid reconstituted ASR is placed in a D$_2$O-based buffer for an extended period of time (~24 h), and its NCA (or more complicated 3D NCACX) spectrum is recorded. While in the buffer, the exchangeable protein protons interacting with solvent may undergo exchange with deuterons. The exchange rate depends on the water accessibility of the protein fragments, and hydrogen bond strength, etc. If an amide proton undergoes H/D exchange, the signal of a corresponding residue disappears from a 2D NCA spectrum. Thus, H/D exchange experiments report site-specifically on the solvent accessibility of the protein fragments. Two drawbacks of such an experiment is that H/D exchange data do
not provide a complete picture, since many amide proton groups, although solvent accessible, do not undergo chemical exchange because they are locked in strong intrahelical hydrogen bonds along the backbone. They also do not report on the kinetics of the amide proton-solvent exchange. In contrast, this new method allows us to detect and identify interactions between water protons and amide protons more directly, as the only requirement is spatial proximity between these species. Thus, with these N-CA experiments, we hope to detect H$_2$O-H$_N$ interactions in addition to those previously seen in ASR.

N-CA correlation experiments only provide information about nuclei along the backbone of the protein, and about how these nuclei interact with water protons. In order to gain additional information about other nuclei in the protein, we perform C-C chemical shift correlation experiments. With this type of spectrum, correlations between sidechain carbon atoms can be detected, and will reveal if bulk mobile or internal bound water is interacting with the sidechains. This could potentially answer if there are buried water molecules, because certain types of residues, such as the positively charged Lys, Arg, His, have mobile sidechains, with protons that are more likely to undergo chemical exchange than backbone amide protons. Also, the groups with polar sidechains, Ser, Thr, Asn, Gln, as well as residues with OH groups (Asp, Glu, Tyr, Trp), may form hydrogen bonds with bound water in close proximity.
As shown in Figure 4-14, many crosspeak regions of the C-C spectra can be resolved site-specifically. From the X-ray structure, a number of residues are known to interact with bound water molecules through hydrogen bonding, specifically Glu36, Arg72, Asp75, Thr90, and Lys210.49 Crosspeaks from these groups of residues can be resolved. This information can be used to detect interactions between water and internal residues.
4.4 Site-Specific Detection Results

4.4.1 Investigation of Backbone Polarization Transfer

Figure 4-15 shows a 2D water-edited NCA spectrum of UCN-ASR at 14.3 kHz spinning frequency, taken with 10 ms of $^1$H-$^1$H mixing. Many peaks are resolved and visible at this short mixing length, indicating that polarization is transferred quickly from mobile waters to amide protons of the detected residues. While some residues visible in the spectrum are known to undergo H/D exchange, and therefore are known to be solvent-exposed (primarily those located in the BC loop), there are a number of additional visible residues which were not seen to exchange in the H/D data (V12, A13, A53, M54, A55, R72, M77, I102, G116, G135, G191, and T196). Most of these residues are located within the helical regions of the extracellular half of ASR, and could be locked in strong hydrogen bonds, preventing them from undergoing exchange. But, their appearance in the NCA short mixing spectrum indicates their close proximity to mobile water protons. The majority of these residues are found directly adjacent to exchangeable regions of ASR, indicating the presence of mobile waters nearby, and providing insight into additional H$_2$O-H$_N$ interactions occurring in ASR. Consistent with the results from 1D experiments, many polar residues are visible including strong lysine and threonine peaks, as well as serine, asparagine and glutamine peaks. In addition to these polar residues, glycine and alanine peaks appear strongly in this spectrum, particularly residues A63, A64, G59 and G65, which are all found in the solvent-exposed B-C loop region. Finally, we see some peaks appear, corresponding to residues whose sidechains are known to interact with water molecules, including R72, D75, T90, Q93, and the Schiff base-forming K210.
Figure 4-15: 2D water edited NCA chemical shift correlation spectrum, with 10 ms $^1$H-$^1$H mixing, and 96 scans. At this short mixing, we see many resolved peaks, including R72, D75, T90, Q93, K210 which are known to interact with bound water, as well as other polar or positively charged residues. Data were processed using Lorentzian-to-Gaussian apodization functions and zero filled to $4096 \times 2048$ prior to Fourier transform. 12 Hz of Lorentzian line narrowing and 30 Hz of Gaussian line broadening were applied in the indirect dimension, while 30 Hz of Lorentzian line narrowing and 60 Hz of Gaussian line broadening were applied in the direct dimension. The first contour is cut at 4σ, with each additional level multiplied by 1.1.

By varying the mixing time used for proton-proton transfer, we can observe how the signal builds up in the protein, and gain a sense of which residues build up faster or slower, depending on how they interact with mobile water protons. Figure 4-16 shows a 2D NCA spectrum taken with 50 ms $^1$H-$^1$H mixing. At long mixing, we see that almost all residues have obtained their signal and are visible in the spectrum. Peaks which appeared in the short mixing spectrum are now much stronger.
Figure 4-16: 2D water-edited NCA chemical shift correlation spectrum taken with 50 ms $^1$H-$^1$H mixing, and 96 scans. Increased signal intensity can be seen in this longer mixing spectrum, and all residues have obtained their signal by this time. Data processing was performed in the same way as for Figure 4-14. The first contour is cut at 4σ, with each additional level multiplied by 1.1.

Because of the uniform labeling scheme, many residues are significantly overlapped in the N-CA plane, so an unambiguous analysis of their increasing signal intensity as a function of mixing time cannot be obtained from this experiment. However, we can use these experiments to visualize how signal propagates through the backbone of the protein as mixing time is increased in a water-edited experiment. This can be visualized as a two-step process, in which spin polarization is first transferred from mobile water to the nearby protein proton spins both by through-space and chemical exchange mechanisms (although this occurs primarily by through-space dipolar interaction, as seen previously), and then in the second step is diffused into the core of the protein because of strong intraprotein dipolar couplings. Given enough mixing time, we expect the entire protein to be re-polarized from water, which is obvious from the 1D $^{13}$C
spectrum at long mixing, shown in Figure 4-7, which closely resembles a regular CP spectrum, and shows signals from polar side chains, many of which interact with water, as well as from aliphatic carbons more likely to be found in the hydrophobic core.

As a demonstration of this signal propagation concept, a 56 residue protein, the β3 immunoglobulin binding domain of protein G (GB3), is used (Figure 4-17).

![Figure 4-17: Amino acid sequence and corresponding secondary structure of protein GB3.](image)

This protein, found on the cell surface of bacteria belonging to group G of streptococcal bacteria, was used as a model system since it has been extensively studied and characterized. Both its high-resolution 3D X-ray crystal structure\(^7^9\), as well as complete \(^13\)C and \(^15\)N solid-state NMR resonance assignments\(^8^0\) are available. GB3 is not a membrane protein, but rather a small globular protein, and therefore does not have the same type of protein-water interface as in the case of ASR. But, because of its small size, almost all its residues can be resolved in a 2D NCA spectrum, so the GB3 sample provides an ideal reference to obtain a complete picture of how polarization can propagate through a protein, before applying this to the larger membrane protein, ASR.

The same water-edited 2D NCA spectra were performed on GB3 with 12 kHz spinning frequency, at various \(^1\)H-\(^1\)H mixing times, in order to observe the buildup of polarization across the protein. The propagation of signal is visualized in Figure 4-18. Three residues in particular (T25, E27, and K28), located in the α-helical region nearest the β2-strand, appear in the first
spectrum, taken with only 1ms proton-proton mixing time. As mixing time is increased, the signal is diffused quickly from this region across the adjacent loop, to the rest of the helix, and by 10ms of mixing time, the entire protein is polarized.

![GB3 12 kHz](image)

_Figure 4-18:_ Structural model of GB3 derived from X-ray, showing which residues are polarized at various mixing times in the corresponding NCA spectra (taken at 12 kHz spinning). Blue regions correspond to residues which are visible in the spectrum, green regions do not appear in the spectrum, and the grey regions are residues which are unresolved due to spectral overlap. It is clear that residues in the lower portion of the helical region (T25, E27, K28) obtain signal quickly, and then this signal is diffused to nearby residues on the helix, and in the β-strands. In the case of GB3, all residues are polarized by 10 ms mixing time.

Spinning at higher frequencies slows the dipolar transfer mechanism, affecting the rate of signal propagation across the protein. _Figure 4-19_ shows a similar visualization of signal propagation across GB3 as proton-proton mixing is varied, but now from NCA spectra taken with a spinning frequency of 20.5 kHz. The same region of the helix obtains its signal quickly (~2ms), as seen at slower spinning, and as before, this signal diffuses outward to the adjacent loop region (~4ms), to the rest of the helix (~6ms), and eventually to the β-strands (~10 ms).
Figure 4-19: Similar to Figure 4-17 (now corresponding to NCA spectra taken at 20.5 kHz spinning). Again it is seen that residues in the lower portion of the helical region obtain signal quickly, and then signal is propagated to other regions as mixing time increases. In general, at faster spinning, residues take longer to obtain polarization since the dipolar transfer mechanism is removed.

The propagation of polarization occurs slower with a spinning frequency of 20.5 kHz than before at only 12 kHz spinning frequency. At faster spinning, the signal transfer mechanism, and therefore signal diffusion through the protein, is slowed by the effects of MAS, which averages out the dipolar interaction. Using GB3 as a demonstration of polarization propagation through a protein is important before applying this technique to ASR, where the process of spin diffusion gets complicated by many more residues.
Indeed, in the case of the large membrane protein ASR, we can visualize the flow of polarization across the protein-water interface and into the core of the protein in a similar way, as the mixing time is increased, as seen in Figure 4-20:

![Structural model of ASR derived from NMR, showing which residues are polarized at various mixing times in the corresponding NCA spectra. Blue regions correspond to residues which are visible in the spectrum, green regions do not appear in the spectrum, and the grey regions are residues which are unassigned, or unresolved due to spectral overlap.](image)

**Figure 4-20:** Structural model of ASR derived from NMR, showing which residues are polarized at various mixing times in the corresponding NCA spectra. Blue regions correspond to residues which are visible in the spectrum, green regions do not appear in the spectrum, and the grey regions are residues which are unassigned, or unresolved due to spectral overlap.

In the shortest mixing experiment, with only 5 ms proton-proton mixing, there are relatively few residues whose backbone nuclei are polarized. Most are found in the B-C loop region where we expect residues to obtain their signal faster than those in buried regions, since amide protons found on the backbone of solvent-exposed regions (B-C loop, for example) are known to undergo chemical exchange with water protons. However, a number of additional residues located relatively close to the surface of the protein, which were found to be exchangeable under illumination conditions in the H/D data (Y51, A71, Y73, W131, T170, Q195 and G220) are also polarized, despite being located in helical regions of the protein. Since they are found to exchange under illumination, this indicates their involvement in the protein’s
functioning, since upon light-activation it has been found that a cavity opens up in ASR. As mixing time increases, polarization propagates from these residues throughout the protein via spin diffusion or dipolar transfer. By 50 ms mixing, polarization has completely diffused throughout the protein and all the residues are visible in the corresponding spectrum (as shown previously).

Because this type of spectroscopy provides insight into backbone nitrogen atoms, we expect the residues which appear first to be consistent with previous H/D exchange data. In H/D exchange experiments, amide protons at exposed sites on the protein will undergo chemical exchange with deuterium in the surrounding solvent, which prevents polarization transfer to other nuclei in the residue. Peak intensities of solvent-exposed residues are therefore attenuated in the spectra. In the H/D exchange data shown in Figure 4-21 (obtained from Wang et al in Ref [46]), blue bars indicate the signal intensity of peaks when the protein was exposed to H$_2$O, while red bars indicate the signal intensity of each residue, after the protein was exposed to 100% D$_2$O buffer. So, residues which have decreased signal intensity after exposure to D$_2$O (red bars), compared to their signal intensity in H$_2$O (blue bars), have undergone chemical exchange and are therefore exposed to the solvent.
Figure 4-21: H/D exchange data, showing signal intensity (in units of RMS noise) both prior to deuteration (blue) and after exposure to 100% D$_2$O buffer solvent (red). Locations where signal disappears (i.e., no red bars present) have undergone H/D exchange. Courtesy of S. Wang, unpublished.

From this data, it is clear that residues along the entire B-C loop, have amide protons which undergo exchange with the solvent. As previously discussed, all resolved residues from the B-C loop appear in the shortest mixing 2D NCA experiment, consistent with their solvent accessibility.

4.4.2 Investigation of Sidechain Polarization Transfer

Amino acid sidechain nuclei are of more functional importance and interest than the backbone nuclei. So, to investigate the interactions between water and residue sidechains in the protein, water-edited experiments were performed using 2D C-C correlation spectroscopy for site-specific detection.
Protons on the sidechain of exposed residues will build up signal faster, since they can undergo chemical exchange or interact through the dipolar mechanism with nearby water protons. Also, residues on the protein-water interface which have sidechain carbons in close proximity to water molecules will appear faster, because signal is transferred via spin diffusion faster. Based on the x-ray structure for ASR, there are a few known water molecules bound within the protein, held in place through hydrogen bonds to nearby residues. We hope to detect these water molecules and the residues to which they are bound.

A water-edited spectrum taken with short $^1$H-$^1$H mixing (10 ms) is shown in Figure 4-22. In the aliphatic region, many crosspeaks are visible already at this short mixing time. All of the alanine CA/CB crosspeaks are visible, consistent with alanine residues appearing in NCA spectra at short mixing as well, despite being hydrophobic residues generally found buried within the helices of the protein. A few alanine residues are found in exposed loop regions of ASR, with three specifically located in the highly mobile B-C loop region: A63, A64, and A68. These residues produce strong, well-resolved crosspeaks, indicating fast transfer from mobile water protons. One alanine which is buried within a helical region is known to be in contact with a bound water molecule. This residue, A40, is not as strong in the short mixing spectrum as those found in the B-C loop. Although A40 is in contact with an internal water molecule, the signal likely takes longer to propagate to this residue than it does in the exposed loop regions.

As expected from previous observations of threonine sidechain carbons appearing at short mixing, most of the CB/CG2 crosspeaks from polar threonine residues appear here as well. Only one threonine is known to interact with a bound water molecule, T90, and it is visible in the short mixing spectrum, although this crosspeak is very small, indicating weak transfer of signal from water. Two residues, T100 and T154, are not visible in either the short or long mixing
spectra, but these residues also have very weak peaks in the reference spectrum. Similar information is seen in the Thr CA/CB region. Despite the apparent weakness of the T90 crosspeak, upon comparison to the intensity of the reference spectrum, this residue actually obtains a greater fraction of its total signal at short mixing compared to the other threonines, indicating fast polarization transfer to this residue.

In the case of serines, another amino acid with polar sidechains, many of these resonances are overlapped in the intense diagonal region of the spectrum. However, S47, S122, and S158 are particularly well-resolved, and visible at short mixing, indicating fast signal transfer to this polar residue type, even though many of them are buried within helices.

Based on previously discussed results, we also expect to see lysine crosspeaks in this spectrum at short mixing, as they have very mobile sidechains. There are very few lysine residues which are fully assigned. The CB/CG crosspeak of K60 is resolved and visible in this short mixing spectrum, consistent with its location in the solvent-exposed B-C loop. The same crosspeak of the Schiff-base forming K210 is only weakly visible in this spectrum, consistent with observations from the 1D experiments that this residue obtains its signal fairly slowly.
Figure 4.22: 2D C-C water-edited correlation spectrum of UCN-ASR taken with 10 ms $^1$H-$^1$H mixing. Many sidechain crosspeaks are resolved in these regions, providing site-specific information on which residues are interacting with water at this short timescale, obtaining their signal quickly through proton-proton transfer. Data were processed with Lorentzian-to-Gaussian apodization functions and zero filled to 16,384 (t1)×4096 (t2) prior to Fourier transform. 10 Hz of Lorentzian line narrowing and 80 Hz of Gaussian line broadening were applied in the indirect dimension, while 40 Hz of Lorentzian line narrowing and 80 Hz of Gaussian line broadening were applied in the direct dimension. The first contour is cut at 4σ, with each additional level multiplied by 1.1.
Crosspeaks from the remaining polar residues, glutamine and asparagine, are visible in the lower section of this short mixing spectrum due to their unique chemical shift values. CG/CD crosspeaks from Q66 and Q157 are both strongly visible, while Q195 is visible but is quite weak. This crosspeak for Q93 (known to interact with bound water) is unresolved, but the CB/CG does appear. These residues are all found in exposed loop regions, and due to their polar sidechains, are likely to be involved in hydrogen bonding, which could facilitate the transfer of signal to their protons. Many asparagine crosspeak signals are unresolved, but N28, found in the A-B loop region of ASR, is resolved and visible in this spectrum.

Residues which are already known to interact with internally bound waters can be resolved in these 2D spectra, and some appear at short mixing. The CB/CG crosspeak of the counterion, D75, appears at short mixing. Although this residue is buried in a helical region of the protein, and is not exchangeable, it is indirectly hydrogen bonded through water with the Schiff base and involved in the proton transport pathway, so it is important to the functioning of ASR. The same CB/CG crosspeak of Arg72 is also visible. This residue was not seen to undergo H/D exchange, but it is bound to an internal water molecule.

The CB/CG crosspeak of the glutamic acid residue E36, known to be in contact with a bound water molecule, is also visible in this spectrum, as are T90 and K210, discussed previously. There are a number of additional residues that are expected to be in contact with bound water molecules, but do not appear to have signal in this short mixing spectrum: Y51, W76, K96, D217. This could indicate that the signal is transferred on a slower timescale. Since bound water signal is likely destroyed by the T2-filter, these would have to undergo fast exchange with bulk water, and then propagate their polarization through to these residues, which would result in a slower buildup rate for these residues.
In the 2D NCA spectrum recorded at short mixing (5 ms), a number of residues were seen despite being located in helical regions of ASR (Y51, A71, Y73, W131, T170, Q195 and G220). Of these, the tyrosine and tryptophan crosspeaks are not resolved in the 2D C-C spectrum, and glycine lacks sidechain carbon atoms. Of the remaining residues, A71, T170, and Q195, all three have crosspeaks which are visible at short mixing in this C-C spectrum. Since these residues undergo H/D exchange, it is clear that they are solvent-exposed and interact with water molecules. Of the additional residues visible in the NCA spectrum that are not exchangeable (V12, A13, A53, M54, A55, R72, M77, I102, G116, G135, G191, and T196), some sidechain crosspeaks are unresolved or lacking. Of the remaining, sidechain crosspeaks of A53, A55, R72 and T196 appear in this short mixing C-C spectrum, supporting the indication that water molecules exist in close contact with these residues, which are unseen by H/D exchange experiments.

**Figure 4-23:** Structural model of ASR, derived from x-ray studies, showing which residues have sidechains which appear in the 2D C-C spectrum with short mixing. This indicates that they interact with mobile waters, and obtain polarization quickly. Red regions are residues which have visible sidechain crosspeaks, green are residues which do not appear in the spectrum, and blue spheres are bound water molecules. Smaller images have BCD (top) and EFG (bottom) helices removed to better view the core.
In summary, there are many residues with sidechains that obtain polarization on a fast timescale, on the order of 10 ms. Figure 4-23 shows these residues on a structural model of ASR, which includes known bound water molecules. As expected, most polar residues, and those with mobile sidechains appear. There is one stretch (residues 156-164) where all sidechains are polarized, located on the cytoplasmic half of helix F. This same stretch of residues on the cytoplasmic side of the helix was shown to be partially exchangeable, indicating its dynamical nature. It is proposed that this helix is tilted when ASR is illuminated with light, this conformational change is thus involved with the function of the protein. The water contact seen in our current data supports the hydrophilic nature of this proton conduction pathway. Based on the H-C correlation experiments, tyrosine residues appear to be interacting with water molecules, but their crosspeaks are not visible, even in the reference spectrum, so it is not clear whether their sidechains are indeed in contact with water. Finally, there are many residues which are in close proximity to internally bound water molecules, as is seen in the cartoon model of ASR, which obtain their signal at short mixing. But there are also many residues surrounding these water molecules which are not yet polarized, which could indicate that some internal waters obtain their polarization on a longer timescale than this short mixing experiment.
Chapter 5
Final Summary and Future Directions

By using solid-state NMR spectroscopy we have investigated the protein-water interface of *Anabaena* sensory rhodopsin through the site-specific detection of water molecules on the surface of this membrane protein. One of the subsequent goals was to determine the nature of the water-protein polarization transfer mechanism, and disentangle the contributions of chemical exchange or through-space interactions, toward this transfer. Finally, with these investigations of the water-protein interface, an additional goal was to detect any lipid-protein interactions present in the system.

Water-edited 2D experiments allow us to site-specifically detect the residues which obtain polarization first, indicating their proximity to water molecules, and then to follow the flow of polarization across the protein. Using 2D NCA experiments, it was seen that residues found along the solvent-accessible B-C loop, and some additional residues in the helical regions, obtained their signal fastest, indicating that they form part of the protein-water interface. Although the helical residues are not located in the solvent-exposed loop regions, some of these residues undergo H/D exchange on a longer timescale, seen in previous work\textsuperscript{46}, which is consistent with the proposed interaction with water in the present work. However, some residues which are buried in the helices and were not previously seen to undergo exchange, have been detected in these NCA experiments, indicating additional H\textsubscript{2}O-H\textsubscript{N} interactions in ASR.

Water-edited 2D C-C correlation experiments have also allowed for the detection of interactions between sidechains and water molecules. Polar residues and those with mobile sidechains in general showed contact with water at short mixing. We have also found that
sidechains of residues known to be interacting with internally bound water molecules can be detected. In addition to this, we see new residues which indicate water-protein interactions, which are in good agreement with those seen in the NCA spectra. These interactions can be correlated with important structural features of the protein, giving further insight into the protein’s function.

Through the application of Lee-Goldburg homonuclear decoupling to remove the dipolar interaction between protons, it was found that the through-space interaction was the main contribution to polarization transfer, as opposed to chemical exchange. This is in contrast to previous work where it was found that chemical exchange was the primary source of water-protein polarization transfer, and further supports the assertion that water molecules are found in close spatial proximity to specific residues for a timescale sufficiently long for polarization transfer to occur across the protein-water interface.

Finally, in a series of 2D HC experiments, it appears that the main proton species which contribute to signal buildup in ASR are water and Hγ protons of the DMPC lipid headgroup. Our data also indicated correlations originating from lipid-protein interactions. This should be investigated further, as it would provide insight into the structural arrangement of ASR in its lipid-embedded environment.

While we have shown that detecting water-protein interactions on a site-specific basis is feasible, there remain unanswered questions regarding the details of these interactions. Thus, in order to further understand the kinetics of polarization transfer, future work should focus on the quantification of transfer mechanisms. It would be valuable to quantify the rates of both the initial water-protein interface transfer and the internal spin diffusion rate, in order to better
distinguish these processes from one another. Future work could also aim to resolve crystal water molecules, and thus be able to determine their position in ASR and how they are related to its function and interaction with the cytoplasmic soluble transducer. Further investigation of lipid-protein interactions through the use of three-dimensional (3D) experiments would also provide valuable insight into the structure of ASR. Understanding membrane protein function is of great importance and using these site-specific SSNMR techniques to study the protein-water interface, we are able to gain valuable structural and functional information. This methodology can be extended to other protein systems to further characterize their protein-water interface, which is of crucial importance for understanding protein function.


