Solid-State NMR Investigations of Transmembrane Proteins - New Approaches for Signal Enhancement and In Situ Studies of Anabaena Sensory Rhodopsin

by

Meaghan Elizabeth Ward

A Thesis
presented to
The University of Guelph

In partial fulfillment of requirements
for the degree of
Doctor of Philosophy
in
Physics

Guelph, Ontario, Canada
© Meaghan Elizabeth Ward, June 2016
ABSTRACT

SOLID-STATE NMR INVESTIGATIONS OF TRANSMEMBRANE PROTEINS – NEW APPROACHES FOR SIGNAL ENHANCEMENT AND IN SITU STUDIES OF ANABAENA SENSORY RHODOPSIN

Meaghan Elizabeth Ward
University of Guelph, 2016

Advisors:
Dr. V. Ladizhansky
Dr. L.S. Brown

Membrane proteins account for 30% of all proteins and perform many important roles in the cell, yet are difficult to study due to the necessity to maintain their lipidic environment during all stages of purification and characterization. Often during these processes the lipid environment is drastically altered, and high resolution studies of membrane protein structure commonly require the use of membrane-mimetic environments. Solid-state NMR (SSNMR) spectroscopy using magic angle spinning (MAS) is a rapidly developing and promising technique to study such proteins in their native, lipid-associated environment. However, these experiments are generally limited by sensitivity. This thesis centers upon the development and application of novel SSNMR approaches to study large, seven transmembrane (7TM) α-helical proteins – a class of proteins to which microbial rhodopsins, which we use as model systems for these experiments, belong. First, the sensitivity available for such samples under ultrafast MAS conditions (> 50 kHz) was investigated. As smaller sample volumes are necessary at these spinning frequencies, low power decoupling and paramagnetic enhancement of the signal relaxation rate were implemented to facilitate a condensed data collection scheme. Under these conditions, it was found that the paramagnetic relaxation enhancement was uniformly distributed throughout the proteins and that sensitivity comparable to that available in larger rotors was
obtainable with proton detection. Next, the implementation of proton detection to specifically detect the mobile regions of proteins was developed. Using these methods it was found that for the mobiles regions, ~10x increase in sensitivity was available and that both the loop regions and lipid and carbohydrates tightly bound to these proteins could be studied. Finally, methods with which to characterize membrane proteins in the native *E. coli* membrane environment were developed and implemented, using *Anabaena* sensory rhodopsin (ASR) as an example. Small, site-specific perturbations in the structure of ASR, which occur as the local membrane milieu changes, indicate that the protein can subtly adapt to its environment without large structural rearrangement. In summary, this work has advanced our ability to use MAS-SSNMR spectroscopy as a structural probe for large and oligomeric membrane proteins.
Acknowledgements

I cannot even begin to express the huge debts of gratitude that I owe to so many people who have helped and influenced me throughout this process. First and foremost, I must thank the boss-men, my advisors Dr. Vladimir Ladizhansky (Big Daddy) and Dr. Leonid Brown (The Jellyfish), for all their eternal patience, help, support, and infallible good attitudes over the last few years. Without their constant encouragement and enthusiasm there is no way I could have completed nearly as much as I did.

In addition, I would also like to thank the members of my advisory committee, Dr. John Dutcher and Dr. George Harauz, and my examining committee, Dr. Steffen Graether, Dr. Simon Sharpe and Dr. Michelle Oliver, for their time, knowledge, and comments on this work. I would also like to particularly thank Dr. George Harauz for all of the moral support and advice.

Many of the projects presented in this work are collaborative efforts. I owe many thanks to Dr. Ivan Hung and Dr. Peter Gor’kov of the National High Magnetic Field Laboratory, Dr. Hongjun Liang, for his expertise in CD spectroscopy, Dr. Bryan Bryska, and Doug Grahame, for their assistance with CD spectroscopy, Dr. Dyanne Brewer, Dr. Armen Charchoglyan for their mass spectrometry expertise and Valerie Robertson, Dr. Andy Lo, and Peter Scheffer of the NMR centre for their technical assistance and willingness to help find a solution to the many problems I managed to get myself into.

Throughout my years at Guelph, I have benefited greatly from the assistance of many of the staff members of the Department of Physics and B.I.G. A thousand thank you’s are owed to Bill Teesdale for keeping my monstrosity of a computer working, as well as to Steve Wilson and the other employees of the shop who could often find a way to fix whatever it was I had broken.
Furthermore, I would like to thank Reggie Vaillilee, David Atkinson, Cynthia Cheeseman and Dr. Hermann Eberl for all of their help throughout the years.

I have been very lucky during my degree to have been surrounded by amazing officemates and peers. Many thanks are owed to Dr. Shenlin Wang, David Bolton, Daryl Good, Emily Ritz, Matthew O’Halloran, David Park, Andrew Harris, Rachel Munro, Dr. Miranda Schmidt, Dr. Mumdooh Ahmed, Ivana Komljenovic, Dr. Vladimir Bamm for all of their help and advice on a wide range of topics, and for putting up with me. I sincerely appreciate all of your time, patience, and office snacks.

Finally, I would never have made it through this with the few shreds of sanity I still possess without the constant love and support of my family and friends. In particular, I remain in complete amazement that Chris Skoog, the wonderful human being that he is, not only stayed with me through all the ridiculousness that I’ve thrown at him, but was also fool enough to marry me.
# Table of Contents

Chapter 1: Introduction to Rhodopsins and Magic Angle Spinning Solid-State Nuclear Magnetic Resonance Spectroscopy ................................................................. 1
  1.1 Cell Membranes and Membrane Proteins .............................................. 2
  1.2 Solid-State NMR (SSNMR) Studies of Membrane Proteins ..................... 4
  1.3 Rhodopsins ......................................................................................... 8
    1.3.1 Anabaena sensory rhodopsin .......................................................... 10
    1.3.2 Proteorhodopsin .......................................................................... 12
  1.4 Statement of Thesis ............................................................................ 13

Chapter 2: Review of Experimental Methods ............................................... 15
  2.1 Introduction to Solid-State NMR .......................................................... 16
    Zeeman Interaction .................................................................................. 16
    Chemical Shift Interaction ...................................................................... 18
    Dipolar Interaction .................................................................................. 19
    J-Interaction .......................................................................................... 21
    Interaction of Spins with Radio-Frequency Field .................................... 21
    2.1.1 Magic Angle Spinning ................................................................. 23
    2.1.2 Multidimensional NMR and Recoupling Techniques ....................... 25
    Dipolar Based Recoupling ....................................................................... 26
    J-Based Recoupling ............................................................................... 29
    2.1.3 Pulse Sequences ........................................................................... 30
    2.1.4 Fast MAS and PACC ..................................................................... 36
    2.1.5 Relaxation ..................................................................................... 37
    T<sub>1</sub> Relaxation ................................................................................ 37
    T<sub>2</sub> Relaxation ................................................................................... 38
  2.2 Fourier Transform Infrared Spectroscopy ................................................. 39
  2.3 Circular Dichroism Spectroscopy .......................................................... 40
  2.4 Small Angle X-Ray Scattering .............................................................. 40

Chapter 3: Materials and Methods .............................................................. 42
  3.1 Materials ............................................................................................ 43
  3.2 Expression and Purification of Membrane Proteins ................................ 43
    3.2.1 Anabaena Sensory Rhodopsin ......................................................... 44
    3.2.2 Proteorhodopsin .......................................................................... 44
    3.2.3 Reconstitution into Synthetic Lipids ................................................. 45
Expression of Anabaena Sensory Rhodopsin with reduced isotopic labelling of background proteins .................................................................................................................. 101
Isolation of Anabaena Sensory Rhodopsin in the E. coli membrane............................ 103
6.2.2 Magic Angle Spinning Solid-State NMR .................................................................. 105
6.2.3 Fourier Transform Infrared Spectroscopy ................................................................. 107
6.2.4 Circular Dichroism Spectroscopy ............................................................................ 107
6.2.5 Small Angle X-Ray Scattering ................................................................................ 107
6.3 Results and Discussion ............................................................................................... 108
6.3.1 Sample Optimization and the State of ASR in E. coli Membranes......................... 108
6.3.2 2D and 3D SSNMR spectra of ASR in cellular membranes ....................................... 117
6.3.3 Structural evaluation of ASR in the E. coli inner membrane .................................... 120
6.4 Summary and Conclusions ......................................................................................... 126
Chapter 7 Conclusion and Future Directions .................................................................... 131
7.1 Concluding Remarks .................................................................................................. 132
7.2 Continued Investigations of Membrane Proteins in Native Environments ................. 133
7.3 Detailed investigation on the effect of lipid acyl chain length on membrane protein structure ...................................................................................................................... 134
List of Tables

Table 4-1 Sample preparation conditions and bulk $^1H T_1$ Values of N148C ASR and PR samples under 50 kHz MAS ......................................................................................................................................................... 56

Table 4-2 Sensitivity comparison between the 3.2-mm E-free probe and the 1.3-mm probe under optimized PACC conditions ............................................................................................................................................. 63

Table 5-1 Spin systems identified in the 2D $^{13}C$-$^{13}C$ INEPT-TOBSY and 3D (H)CHH experiments on UCN ASR ..................................................................................................................................................... 83

Table 5-2 Amino acid systems identified in the 3D (H)CHH experiments on UCN MBP ......... 94

Table 6-1 Summary of chemical shifts in EM-ASR (BMRB entry 25290) and PL-ASR ....... 128
List of Figures

Figure 1-1 Cartoon representation of the signal transduction cascade of Anabaena Sensory Rhodopsin (ASR) .................................................................................................................................................................................................................. 11
Figure 2-1 Representation of the interactions occurring in a spin ½ system ........................................... 18
Figure 2-2 A) Representation of MAS in SSNMR. B) the effect of MAS on the anisotropic interactions. ....................................................................................................................................................................................................... 25
Figure 2-3 Basic pulse sequences .............................................................................................................. 31
Figure 2-4 Pulse sequences for dipolar based experiments ........................................................................ 32
Figure 2-5 A) NCA spectra of ASR reconstituted in proteoliposomes. Representative two-dimensional planes of B) the 3D CANCO and C) NCACX experiments collected on ASR .................. 33
Figure 2-6 Pulse sequences for J-based experiments ................................................................................. 35
Figure 2-7 Pulse sequences used for A) T1 and B) T2 relaxation measurements .................................... 39
Figure 4-1 A) Schematic representation of lipid reconstituted ASR B) Absorption spectra of liposomes containing ASR reconstituted in the presence of 5 mM pyranine ............................................ 58
Figure 4-2 2D NCA spectra of A) diamagnetic ASR, B) ASRCu C), ASR40Cu, and D) ASR80Cu. 1D traces extracted from 2D spectra of E) ASR, F) ASRCu, G) ASR40Cu, and H) ASR80Cu .................................................................................. 60
Figure 4-3 Site-specific comparison of the relative S/N of ASRCu (blue), ASR40Cu (red), and ASR80Cu (green) for residues which are resolved in all four 2D NCA spectra ........................................................................... 63
Figure 4-4 1H-15N heteronuclear correlation spectra of U-[2H-15N-13C]-PR. Diamagnetic PR at a spinning frequency of 20.5 kHz A), 50 kHz B), and C) 50 kHz with 80 mM CuEDTA added to the solvent ........................................................................................................... 65
Figure 5-1 A) Schematic representation of 18.5-kDa MBP. B) Schematic representation of lipid reconstituted ASR .............................................................................................................................................................. 72
Figure 5-2 1D 13C A) INEPT- and B) CPMAS-based spectra of UCN MBP. C) 1D 13C INEPT- and D) 1D 13C CPMAS-based spectra of UCN ASR ................................................................................................. 77
Figure 5-3 A) 2D 13C-13C spectra of UCN ASR obtained with INEPT excitation and TOBSY mixing (black) and CPMAS excitation and DARR mixing (red). B) Enlarged regions of the same spectra ................................................................................................................................................................................................. 81
Figure 5-4 1D 1H spectra of A) UCN MBP and B) UCN ASR in 100% H2O buffer (black) and D2O buffer (red) ........................................................................................................................................................................................................ 84
Figure 5-5 Comparison of the aliphatic region of the 1H-13C INEPT HSQC spectra of MBP and ASR. A) 13C-detected and B) 1H-detected HSQC spectrum of MBP. D) 13C-detected and E) 1H-detected HSQC spectrum of ASR. C) and F) 1D traces extracted from the 2D 1H-detected (black) and 13C-detected spectra (grey) of MBP and ASR, respectively ......................................................................................................................... 87
Figure 5-6 A) 2D spectrum of an (H)C(H)H experiment on MBP (black) overlaid with the 2D 1H-13C INEPT HSQC (blue) ........................................................................................................................................................................ 92
Figure 5-7 A) 2D spectrum of an (H)C(H)H experiment on ASR (black) overlaid with the 2D 1H-13C INEPT HSQC (blue) .................................................................................................................................................................................................. 94
Figure 6-1 Overview of isotopic labelling and sample preparation strategies ........................................... 102
Figure 6-2 Summary of isotopic labelling, sample preparation strategies, and sample compositions ................................................................................................................................................................................................ 105
Figure 6-3 A) Sypro Ruby Protein and B) Invison His6-tag stained SDS/PAGE analysis of E. coli membranes from BL21-Codon-plus-RIL cells containing the plasmid for ASR which have been induced (ASR) and not induced (NIC) ...................................................................................................................................... 109
Figure 6-4 A) FTIR spectra of NA EM-ASR throughout the inner membrane isolation process. B) One-dimensional $^{15}$N spectra of i) UCN PL-ASR, ii) UN EM-ASR, and iii) rbUN EM-ASR. C) One-dimensional $^{15}$N spectra of UCN PL-ASR and rbUN EM-ASR.

Figure 6-5 A) FTIR spectra of rbUCN ASR before and after the IM isolation process.

Figure 6-6 A) Visible range CD and B) absorption spectra of PL-ASR and EM-ASR. Small angle X-ray scattering spectra of C) PL-ASR and D) EM-ASR.

Figure 6-7 A) Small angle x-ray scattering of green proteorhodopsin (PR) in the E. coli membrane. B) Absorption spectra of EM-NIC.

Figure 6-8 NCA SSNMR spectra of A) PL-ASR, B) rbUCN EM-ASR, and C) rbUCN EM-NIC. Overlaid representative 2D planes of D-G) 3D NCACB and H) CANCO experiments performed on PL-ASR (red), EM-ASR (black), and EM-NIC (blue).

Figure 6-9 Additional overlaid representative 2D planes of 3D CANCO (A, B, H) and NCACB (C-G, I) experiments performed on PL-ASR (red), EM-ASR (black), and EM-NIC (blue).

Figure 6-10 Site-specific chemical shift perturbations and signal-to-noise ratio (SNR) comparisons between EM-ASR and PL-ASR.

Figure 6-11 A) Structural model of a monomer of trimeric ASR. B) Close-up of the retinal pocket. C) Side view and D) view from the extracellular side of the interaction interface.

Figure 7-1 Structural models of the lipids used in this study and a 3D model of PR.
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2/3D</td>
<td>one-/two-/three-dimensional</td>
</tr>
<tr>
<td>ASR</td>
<td><em>Anabaena</em> sensory rhodopsin</td>
</tr>
<tr>
<td>ASRT</td>
<td><em>Anabaena</em> sensory rhodopsin transducer</td>
</tr>
<tr>
<td>BR</td>
<td>bacteriorhodopsin</td>
</tr>
<tr>
<td>CARA</td>
<td>computer assisted resonance assignment</td>
</tr>
<tr>
<td>CD</td>
<td>circular dichroism</td>
</tr>
<tr>
<td>CP</td>
<td>cross polarization</td>
</tr>
<tr>
<td>CPMAS</td>
<td>magic angle spinning cross-polarization</td>
</tr>
<tr>
<td>CW</td>
<td>continuous wave</td>
</tr>
<tr>
<td>DARR</td>
<td>dipolar assisted rotational resonance</td>
</tr>
<tr>
<td>DDM</td>
<td>n-dodecyl β-d-maltoside</td>
</tr>
<tr>
<td>DMPA</td>
<td>1,2-dimyristoyl-sn-glycero-3-phosphate</td>
</tr>
<tr>
<td>DMPC</td>
<td>1,2-dimyristoyl-sn-glycero-3-phosphocholine</td>
</tr>
<tr>
<td>DMPG</td>
<td>1,2-dimyristoyl-sn-glycero-3-[phospho-rac-(1-glycerol)]</td>
</tr>
<tr>
<td>DNP</td>
<td>Dynamic Nuclear Polarization</td>
</tr>
<tr>
<td>DREAM</td>
<td>Dipolar Recoupling Enhanced by Amplitude Modulation</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamineteraacetic acid</td>
</tr>
<tr>
<td>EM</td>
<td><em>E. coli</em> membrane</td>
</tr>
<tr>
<td>FTIR</td>
<td>Fourier Transform Infrared Spectroscopy</td>
</tr>
<tr>
<td>GB1</td>
<td>first immunoglobulin binding domain of protein G</td>
</tr>
<tr>
<td>GPCR</td>
<td>G-protein coupled receptor</td>
</tr>
<tr>
<td>GPR</td>
<td>green proteorhodopsin</td>
</tr>
<tr>
<td>HORROR</td>
<td>homonuclear rotary resonance recoupling</td>
</tr>
<tr>
<td>HR</td>
<td>high resolution</td>
</tr>
<tr>
<td>HSQC</td>
<td>heteronuclear single-quantum coherence</td>
</tr>
<tr>
<td>IM</td>
<td>inner membrane</td>
</tr>
<tr>
<td>INEPT</td>
<td>Insensitive Nuclei Enhanced by Polarization Transfer</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>LB</td>
<td>Lysogeny Broth</td>
</tr>
<tr>
<td>LG</td>
<td>Lee-Goldburg</td>
</tr>
<tr>
<td>LUV</td>
<td>large unilamellar vesicles</td>
</tr>
<tr>
<td>MAS</td>
<td>magic angle spinning</td>
</tr>
<tr>
<td>MBP</td>
<td>myelin basic protein</td>
</tr>
<tr>
<td>MMTS</td>
<td>methyl methanethiosulfonate</td>
</tr>
<tr>
<td>MTS-EDTA</td>
<td>[S-methanethiosulfonylcysteaminy1 ethylenediamine-(N,N',N',N')-Tetraacetic Acid</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>NA</td>
<td>natural abundance</td>
</tr>
<tr>
<td>NIC</td>
<td>non-induced cells</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>NOE</td>
<td>nuclear Overhauser effect</td>
</tr>
<tr>
<td>OM</td>
<td>outer membrane</td>
</tr>
<tr>
<td>OmpA/F</td>
<td>outer membrane protein A/F</td>
</tr>
<tr>
<td>OS</td>
<td>oriented sample</td>
</tr>
<tr>
<td>PACC</td>
<td>paramagnetically-assisted, condensed data collection</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>PI</td>
<td>pre-induction</td>
</tr>
<tr>
<td>PL</td>
<td>proteoliposome</td>
</tr>
<tr>
<td>PM</td>
<td>purple membrane</td>
</tr>
<tr>
<td>ppm</td>
<td>parts per million</td>
</tr>
<tr>
<td>PR</td>
<td>proteorhodopsin</td>
</tr>
<tr>
<td>PRE</td>
<td>paramagnetic relaxation enhancement</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>rotational resonance</td>
</tr>
<tr>
<td>R&lt;sup&gt;3&lt;/sup&gt;</td>
<td>heteronuclear rotary resonance recoupling</td>
</tr>
<tr>
<td>RAP</td>
<td>reduced adjoining protonation</td>
</tr>
<tr>
<td>rb</td>
<td>reduced background</td>
</tr>
<tr>
<td>RF</td>
<td>radio frequency</td>
</tr>
<tr>
<td>RFDR</td>
<td>radio-frequency driven recoupling</td>
</tr>
<tr>
<td>SAXS</td>
<td>small angle X-ray scattering</td>
</tr>
<tr>
<td>SB</td>
<td>Schiff base</td>
</tr>
<tr>
<td>SDS/PAGE</td>
<td>sodium dodecyl sulfate / polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SPECIFIC CP</td>
<td>spectrally induced filtering in combination with CP</td>
</tr>
<tr>
<td>SRII</td>
<td>sensory rhodopsin II</td>
</tr>
<tr>
<td>SSNMR</td>
<td>solid-state nuclear magnetic resonance</td>
</tr>
<tr>
<td>TEA</td>
<td>triethylamine</td>
</tr>
<tr>
<td>TM</td>
<td>transmembrane</td>
</tr>
<tr>
<td>TOBSY</td>
<td>Total Through-Bond Correlation Spectroscopy</td>
</tr>
<tr>
<td>TPPM</td>
<td>two pulse phase-modulated</td>
</tr>
<tr>
<td>UCN</td>
<td>Uniformly &lt;sup&gt;13&lt;/sup&gt;C- and &lt;sup&gt;15&lt;/sup&gt;N-labeled</td>
</tr>
<tr>
<td>UCND</td>
<td>Uniformly &lt;sup&gt;13&lt;/sup&gt;C, &lt;sup&gt;15&lt;/sup&gt;N, and &lt;sup&gt;2&lt;/sup&gt;H-labeled</td>
</tr>
<tr>
<td>UN</td>
<td>Uniformly &lt;sup&gt;15&lt;/sup&gt;N-labeled</td>
</tr>
<tr>
<td>WS</td>
<td>water suppression</td>
</tr>
<tr>
<td>WT</td>
<td>wild-type</td>
</tr>
</tbody>
</table>
Chapter 1:
Introduction to Rhodopsins and Magic Angle Spinning Solid-State Nuclear Magnetic Resonance Spectroscopy

Parts of this chapter have been published in the review article:

1.1 Cell Membranes and Membrane Proteins

The native cell membrane is a complex environment which is composed of a wide range of lipids and membrane-spanning or -associated proteins. While the primary function of cell membranes is the physical separation of cellular components from the exterior environment and from each other, the membrane proteins embedded in these bilayers serve to transport signals and materials across these barriers in a controlled manner. Many Gram-negative prokaryotic cells, including *E. coli*, contain two surrounding membranes, denoted as the inner and outer membranes, whereas eukaryotic cells contain only one surrounding plasma membrane, as well as many internal membranes which separate cell compartments and organelles from one another and create micro-environments within cells.

The protein and lipid compositions of cellular membranes can vary significantly amongst species and even within the cell. For example, whereas the inner membrane of Gram-negative bacteria is largely composed of phospholipids and contains mainly α-helical proteins, the outer membrane is rich in lipopolysaccharides, which are not typically found in the inner membrane, and contains mainly β-barrel intrinsic membrane proteins. Membrane composition can also be altered to perform a specific function. For example, in mitochondria and chloroplasts the inner membranes are folded to increase surface area and have a high density of energy-producing proteins. Similarly, the purple membrane (PM) of *Halobacterium halobium* is composed of a two-dimensional crystal lattice of bacteriorhodopsin with minimal lipid incorporation.

Biological membranes must be composed of lipids which support optimal functioning of the embedded membrane proteins. The majority of lipids which surround intrinsic membrane proteins serve to protect the hydrophobic core of the proteins from polar aqueous medium, therefore acting as a general solvent for the protein. Whereas these annular lipids do not form
strong, long-lasting interactions with the protein and often exchange rapidly with the surrounding environment, many proteins have been observed to interact strongly and with high specificity with a small number of lipid molecules. These non-annular lipid molecules can be essential for proper protein function, are often located between transmembrane α-helices, and have been observed to act as protein cofactors [1]. Non-annular lipids have also been observed to mediate protein-protein interactions and stabilize microenvironments within the membrane [2–4].

The lipid membrane environment influences both the conformation and function of membrane proteins [5, 6], but the specific mechanisms of this modulation are not well understood. Many factors have been hypothesized to be the basis for lipid regulation of membrane protein function, including bilayer fluidity or stiffness [7, 8], bilayer-protein hydrophobic mismatch [9], intrinsic lipid curvature [10] or curvature frustration [11], bilayer deformation energy [12], acyl chain packing or lipid packing stress [13, 14], bilayer free volume [15], or the lateral pressure profile [16]. Many of these properties are closely related and often are not independent of each other. In addition, the presence of proteins in the lipid bilayer can have an effect on the bilayer properties. For example, it is not obvious in cases of mismatch between the lengths of the hydrophobic domains of the proteins and lipids involved, whether the lipids stretch or compress in order to match the hydrophobic region of the protein, or whether the helices of the protein tilt or flex to match the protein to the hydrophobic thickness of the lipids [9, 17].

In addition to interactions with lipids, membrane protein structure is also affected by interactions with other cellular components, such as other proteins or small molecules [6, 11, 18, 19]. Many membrane proteins form oligomers which are necessary for proper function, or exist
as part of a protein complex. The protein-protein interactions formed under these circumstances can have a significant effect on the protein structure.

1.2 Solid-State NMR (SSNMR) Studies of Membrane Proteins

The study of the structure and dynamics of membrane proteins in a cellular environment in which they are subjected to a wide range of interactions with the natively present cellular components, e.g., other proteins, nucleic acids, lipids, cofactors, etc., is a major aspect of structural biology. Removal of these interactions can lead to perturbations of the native protein structure and possible reduction or loss of protein function, therefore limiting our understanding of the biological roles of these molecules. As cellular environments are intrinsically complex, structural heterogeneity is not uncommon and, with a few exceptions [20, 21], such systems are not readily amenable to traditional high-resolution structural biology methods such as X-ray crystallography, cryo-electron microscopy, etc. Thus, the majority of structural studies undertaken to date have involved purification of the protein of interest through removal from its native membrane, isolation, and subsequent crystallization, or solubilization in suitable detergents. Through this general strategy, the structures of many membrane proteins have been elucidated in detergent micelles [22–26] and detergent-based crystals by solution NMR and X-ray crystallography [27, 28], respectively. Though these proteins have been removed from the native membrane environment, fragments of lipid molecules, most likely representing non-annular lipids which are tightly bound and interact specifically with the protein, are occasionally observed in crystal structures structures[29–36].

Solid-state NMR (SSNMR) spectroscopy, the theory of which is discussed in detail in Chapter 2, is a rapidly-developing technique which provides unique opportunities for the study
of membrane proteins. Whereas solution NMR and X-ray crystallography are limited by the necessity of fast tumbling rate (meaning limited overall size), or by the capacity to form large three-dimensional (3D) crystals, respectively, SSNMR allows for the study of membrane proteins of any size in a lipid milieu which more closely mimics the native cellular membrane environment. Because molecular tumbling is suppressed in such samples, their NMR spectra are broadened by strong anisotropic interactions. Many of the first SSNMR structures were obtained using oriented sample (OS) [37, 38] approaches, which focused on small α-helical proteins. More commonly, magic angle spinning (MAS) [39, 40] in combination with high power decoupling [41] is utilized to average out the anisotropic interactions and to re-establish high resolution.

To date, the majority of applications of MAS SSNMR to large α-helical membrane proteins have relied on long-established methods, employing carbon detection, robust polarization transfer schemes, moderate spinning frequencies (10-20 kHz), and high radio-frequency (RF) power decoupling in order to achieve sufficiently narrow linewidths. These studies have significant limitations for membrane structural biologists for several reasons. Typically, milligram quantities of the membrane protein which has been isotopically labelled, purified from its cellular environment, and reconstituted into a membrane-mimetic environment are required. Such samples are not always easily obtainable due to complications in protein expression and purification, particularly of mammalian membrane proteins. Furthermore, these procedures could have undesired effects on the structures of the proteins studied. The necessity of the lipid environment leads to an overall reduction in the amount of protein which can be present in the sample, and thus a reduction in experimental sensitivity. The typically high molecular mass of membrane proteins also reduces sensitivity, and leads to poor spectral
resolution due to the high repetitiveness of hydrophobic residues, the dominance of α-helical secondary structure, and high spectral congestion. Therefore, the use of numerous complementary, yet less-sensitive, experiments of high dimensionality is necessary. Furthermore, the peripheral solvent-exposed regions of membrane proteins are often mobile and may be disordered, creating additional sensitivity challenges for their detection. Finally, the experimental conditions described above necessitate the use of long recycle delays between successive scans (1.7 s or longer) to avoid probe and sample damage caused by the application of high power RF. Were these delays not necessary, experiments could be repeated much more rapidly, leading to an increase in experimental sensitivity.

Many of these limitations have been addressed through the implementation of ultra-fast MAS frequencies and proton-detected experiments. The implementation of proton detection provides an increase in sensitivity, additional chemical shift data, and more sensitive probes of conformational and environmental changes [42, 43]. However, applications of proton detection are generally limited by the broad lines caused by the strong proton-proton interactions, which are not sufficiently averaged under standard SSNMR conditions. Recent advances in high magnetic field and fast MAS probe technologies have extended the range of spinning frequencies up to ~110 kHz [44–48], and proton linewidths of ~0.2 ppm (160-180 Hz at 800 MHz field strength) could be achieved on fully-protonated samples at MAS rates of 40-60 kHz [49–53]. Proton linewidths could be significantly improved by combining high spinning frequencies with perdeuteration of the sample and the re-introduction of protons at exchangeable sites through back-exchange with protonated buffer [48, 54–61].

Ultra-fast MAS frequencies not only facilitate proton-detected experiments [49, 50, 62], but also enable the use of low-power decoupling [63–65], and suppress coherent contributions to
relaxation rates, thus providing direct access to dynamic information [66, 67]. An important challenge in the use of the special rotors designed for ultra-fast MAS is the drastically reduced sensitivity caused by the severely decreased sample volume (i.e., the volume of a 1.3-mm rotor is \(~15\times\) less than that of a 3.2-mm rotor). It is, therefore, important to investigate the extent to which the benefits offered by ultrafast MAS, such as the increased coil efficiency, accelerated acquisition schemes, and the potential for proton detection, can compensate for this drastic reduction in the amount of sample. As these special rotors have a higher sensitivity per unit sample, ultrafast MAS probes are an attractive option for the study of proteins that can only be produced in small quantities, particularly eukaryotic membrane proteins [68–70]. However, these methods are still somewhat limited because only exchangeable amide and sidechain protons are detectable. Thus, SSNMR experiments on fully-deuterated samples are restricted to those which specifically detect these atoms, and sidechain proton assignments are difficult to obtain. In \(\alpha\)-helical bundles, the number of detectable residues is further reduced because the hydrophobic core of these proteins is often completely shielded from the solvent, and is therefore not amenable to backbone exchange [58, 59, 71–77]. To combat these deficiencies, additional protons can be sparsely incorporated into the sample at low levels through the reduced adjoining protonation (RAP) labeling scheme [57], or by the selective reintroduction of protons into methyl-bearing sidechains [78–81], without major sacrifices to the linewidth.

Under standard MAS SSNMR conditions (~10-20 kHz MAS, fully protonated samples), proton detection has typically been restricted to the indirect detection of mobile resonances. While membrane embedded proteins are composed mainly of rigid secondary structural elements which are linked by short, flexible, and often less structured, loops and turns, many membrane-associated proteins have been observed to contain longer, more mobile loop regions. Within the
less structured regions, local, sub-microsecond molecular motions of sufficiently large amplitude can lead to an averaging of the dipolar interaction, therefore facilitating through-bond polarization transfers, which are based on the comparatively weak J-interactions. Excellent resolution in the indirect proton dimension has been demonstrated on the mobile regions of fully-protonated integral and peripheral membrane proteins [82–85], histones [86], and amyloid fibrils [87, 88].

Despite the above challenges, advances in SSNMR methodologies under standard conditions have already resulted in several lipid-reconstituted membrane protein structures [89–98] and have paved the way for detailed investigations of structural perturbations caused by non-physiological environments. Although an artificial lipid bilayer usually represents a good experimentally tractable model environment for structural and functional studies of membrane proteins, the natural cellular membrane is invariably more complex and includes a diverse array of lipids as well as peripheral and transmembrane proteins. Recently, several studies have demonstrated the feasibility of studying over-expressed peptides and proteins in the native E. coli membrane using SSNMR, particularly when enhanced by Dynamic Nuclear Polarization (DNP) [99–104]. To date, these studies have focused on small, highly abundant α-helical proteins [99, 101], or the general structural characterization of outer-membrane proteins [100] and membrane-associated complexes [105] with minimal site-specific characterization of the structure.

1.3 Rhodopsins

Rhodopsins are a class of integral membrane proteins which bind a retinal chromophore and form light-absorbing complexes [106]. These proteins share a common overall structural architecture of a seven transmembrane (7TM) α-helical bundle which surrounds a retinal
molecule that is covalently bound to a Schiff-base forming lysine in the seventh helix, helix G. Rhodopsins have been observed to perform a wide variety of functions, including the creation of ion gradients by acting as a proton pump in many single celled organisms and fungi, and as photoreceptors in many organisms such as bacteria, fungi, algae, higher-order eukaryotes etc. [106].

Microbial rhodopsins have traditionally acted as a testing ground in membrane protein research due to the ease of production and isolation of these proteins when compared to the structurally similar, yet more medically relevant, G-protein coupled receptors (GPCRs). The first structure of a membrane protein determined by electron microscopy in 1975 was that of an Archaeal proton pump, bacteriorhodopsin (BR) [107], and this structure was extensively used as a template to model GPCRs, until the first crystallographic study of visual rhodopsin was published in 2000 [108]. Microbial rhodopsins were also studied in some of the first applications of solution NMR to membrane proteins [109, 110]. Likewise, many SSNMR methodologies have been developed using BR, from early applications of rotational resonance ($R^2$) [111] and dipolar tensor correlation methods [112], to the determination of retinal conformation in selectively isotopically-labeled samples, to the demonstration of high proton resolution in polytopic $\alpha$-helical membrane proteins [58], to pioneering studies of the BR function using DNP [113]. Several other rhodopsins, including proteorhodopsin (PR) [114–116] sensory rhodopsin II (SRII) [82], Leptosphaeria rhodopsin [117], and Anabaena Sensory Rhodopsin (ASR) [93] have also been structurally characterized using SSNMR.

In our lab, expression and purification protocols for the production of uniformly $^{13}$C- and $^{15}$N-labeled ASR and PR samples for SSNMR have been well established (see references throughout the following section). As discussed below, the large amount of SSNMR
spectroscopic data available for ASR, which includes a high-resolution structure, make it an ideal system for site-specific investigations of signal enhancement, as well as of the effect of the membrane environment on the protein structure. Though the development of a protocol for the production of a deuterated sample of ASR are underway (Bolton et. al., unpublished) this work is still ongoing. Therefore, for applications of proton detection to the rigid regions of our membrane protein systems we investigate PR, which has previously been successfully produced in a perdeuterated form [59].

### 1.3.1 Anabaena sensory rhodopsin

*Anabaena* Sensory Rhodopsin (ASR) ([Figure 1-1](#)) is a cyanobacterial light-sensitive receptor which is likely responsible for chromatic adaptation in its host cell [29, 118]. ASR is unique among the known rhodopsins in that it is believed to be the only microbial rhodopsin to interact with a cytoplasmic soluble transducer (ASRT)[118–120]. ASR can be overexpressed in *E.coli*, and reconstituted in lipids at a high protein-to-lipid ratio in a fully functional state. It forms stable trimers in both detergents and lipids [121] and forms a two-dimensional (2D) lattice in DMPC:DMPA lipids [122]. These properties are important for establishing high spectral resolution and sensitivity in SSNMR, which facilitates the structural and dynamical characterizations of this protein.
Figure 1-1 Cartoon representation of the signal transduction cascade of *Anabaena* Sensory Rhodopsin (ASR). ASR, shown as a seven-helical bundle embedded in the bilayer, with retinal (red) covalently linked to seventh helix G, interacts with a soluble tetrameric transducer (ASRT) and is believed to be responsible for the regulation of the genes of several light-harvesting proteins.

Solid-state NMR resonance assignments for approximately 90% of ASR’s residues have been obtained by our group [123, 124], and a high-resolution trimer structure has been calculated from inter-nuclear distances, torsional restraints, and paramagnetic relaxation enhancements (PRE) [93, 121]. This structure reveals significant deviations from the structure previously solved using X-ray crystallography [29], including differences in the arrangement of some of the peripheral loops, as well as looser helical packing on the cytoplasmic side [93] of ASR in proteoliposomes. Secondly, whereas the SSNMR structure reveals a single all-trans conformation of the retinal chromophore and of the retinal binding pocket in the dark state, yet double, structurally very similar, conformations for a number of residues located on the cytoplasmic side [124], the all-trans and 13-cis conformations of retinal coexist in crystals [29], which results in double conformations for some of the residues in the vicinity of retinal (e.g., K210, S86), but does not cause perturbations in the rest of the protein. Furthermore, the crystal structure reveals dimerization whereas the SSNMR structure has a trimeric organization.
The availability of a high resolution SSNMR structure of ASR allows for the site-specific determination of many properties. For example, using a series of 2D NC experiments, dipolar order parameters could be measured for many residues [125]. These measurements have suggested the presence of collective motions in some of the exposed, ordered loops as well as in the transmembrane \( \alpha \)-helices [125]. These observation are consistent with the higher solvent accessibility of the cytoplasmic side of the protein, both in the dark state [123] and under illumination [126]. The increased conformational adaptability on the cytoplasmic side may be related to the function of ASR, in particular to its ability to interact with its transducer.

1.3.2 Proteorhodopsin

Proteorhodopsins are found in many species of marine bacteria in the photic zones of the oceans [127, 128]. In particular, green proteorhodopsin (henceforth referred to simply as PR) acts as a light driven proton pump and has been the focus of several studies [127, 129, 130]. Though the predicted differences in growth rates or cell yields between dark and light cultures were not initially observed [131], PR has been shown to play an important role in a complex cellular response that maintains cell functions during periods of carbon starvation [132].

Currently, little structural information is available for PR. Low-resolution atomic force and cryo-electron microscopy, as well as DNP-NMR data show that PR can assemble into pentameric or hexameric complexes [133–136]. Though PR produces two-dimensional (2D) crystals, it does not form well-diffracting 3D crystals [137], which precludes x-ray crystallography studies, however the structure of several close homologs have been elucidated in this way [29, 32, 138, 139]. Though solution NMR HSQC spectra of detergent-solubilized PR initially contained an insufficient number of resolvable peaks [137], a structure has since been
obtained using this method [116]. However, this structure contains a number of significant differences from the other existing models of retinylidene proteins, and significant structural differences exist between the multiple conformers obtained.

Solid-state NMR studies on PR reconstituted into lipid vesicles with the aim of solving a high-resolution structure for PR are currently underway in our laboratory [114, 140]. To date, 168 of 238 amino acids have been assigned, with the majority of the residues being located in the transmembrane regions. Difficulties in the acquisition of assignments of the loop regions of the protein may be due to increased motions in these areas [84, 114, 140]. In addition to carbon- and nitrogen-detected experiments, a fully-deuterated sample of PR has also been studied using proton-detected experiments [59]. Through these experiments, the solvent-accessible regions of the protein could be identified, and the presence of a hydrophilic cavity in the transmembrane region of the protein was indicated.

1.4 Statement of Thesis

In this doctoral thesis, novel methods for signal enhancement and in situ studies of membrane proteins using solid-state NMR are developed and applied to ASR. In Chapters 2 and 3, we review the principles of NMR spectroscopy and the experimental procedures. In Chapter 4, we investigate the application of ultra-fast MAS and paramagnetically-assisted, condensed data collection (PACC) to ASR and deuterated PR [141]. As studies of membrane proteins in more complex environments will lead to a lower concentration of the protein of interest in many samples, the potential increase in sensitivity available from these types of experiments is an attractive prospect. In Chapter 5, the application of proton detection to the mobile regions of proteins is investigated [142]. Although the use of indirect proton detection is well-established
for these regions, the direct detection of protons has not been. These experiments allow not only for the characterization of sufficiently mobile loop regions of membrane proteins, but also reveal tightly-associated carbohydrate and lipid molecules. Finally, in Chapter 6, methods for the study of membrane proteins in the *E. coli* membrane are developed [122]. Through these methods, the structure of ASR can be compared in the proteoliposome environment and the much more complex environment of the *E. coli* membrane. In Chapter 7, the ongoing research pertaining to the previous chapters is discussed. More precisely, the studies of the effect of the membrane environment on ASR are extended through specific investigations on the effect of lipid acyl chain length on membrane protein structure.
Chapter 2:
Review of Experimental Methods
2.1 Introduction to Solid-State NMR

Nuclear magnetic resonance (NMR) spectroscopy is based on the inherent quantum mechanical spin property of atoms. Specifically, in protein NMR, the nuclei of interest (\(^1\)H, \(^{13}\)C, \(^{15}\)N) all have a spin value of \(\frac{1}{2}\). As \(^{13}\)C and \(^{15}\)N do not occur at a high natural abundance, enhanced isotopic labelling of samples with these isotopes is necessary. A system composed of spin \(\frac{1}{2}\) atoms can be represented by the Hamiltonian

\[
H = H_Z + H_{CS} + H_D + H_J + H_{RF}
\]

(2.1)

where

\[
H_Z = \text{Zeeman Interaction} \quad (2.1a)
\]

\[
H_{CS} = \text{Chemical Shift Interaction} \quad (2.1b)
\]

\[
H_D = \text{Dipolar Interaction} \quad (2.1c)
\]

\[
H_J = J - \text{Interaction} \quad (2.1d)
\]

\[
H_{RF} = \text{Interation of nuclear spin with radio frequency field} \quad (2.1e)
\]

While \(H_{CS}, H_D, \text{ and } H_J\) are intrinsic interactions, and the effect of these interactions on the NMR spectra is summarized in Figure 2-1, \(H_Z\) and \(H_{RF}\) are the result of the applied magnetic field and an applied radio-frequency pulses, respectively. The quadrupolar interaction, which is only present in nuclei with spin values greater than \(1/2\), need not be considered here [143].

Zeeman Interaction

Under the conditions of a standard NMR experiment, the Zeeman interaction is several orders of magnitude larger than the intrinsic interactions. Therefore in standard NMR theory, components of these interactions which do not commute with the Zeeman interaction are generally neglected, while the eigenstate energies are corrected by secular terms which do
commute. The Zeeman interaction describes the interaction between spins and an applied, static magnetic field and can be represented by the Hamiltonian

\[ H_Z = \omega_0 I_Z \]  

(2.2)

where

\[ \omega_0 = -\gamma \hbar B_0, \]  

(2.3)

and \( \gamma \) is the gyromagnetic ratio of the nucleus, \( B_0 \) is the strength of the magnetic field, which is chosen to be along the \( z \)-axis, \( \vec{B} = (0,0,B_0) \), \( I_z \) is a spin \( \frac{1}{2} \) operator, \( \hbar \) is the reduced Planck's constant, and \( \omega_0 \) is the Larmor frequency of the nucleus. By convention, energy and field strength are measured in frequency units and \( \hbar \) is set to equal 1.

The Larmor frequency is equal to the energy difference between the two eigenstates of the Hamiltonian, \(|\alpha\rangle, |\beta\rangle\), as seen in Figure 2-1, which correspond to the \( \pm \frac{1}{2} \) projections of the \( z \)-component of the spin. These states have energies, in frequency units, of

\[ E_\alpha = -\frac{1}{2} \gamma B_0, \quad E_\beta = \frac{1}{2} \gamma B_0 \]  

(2.4)

and their respective populations \((n_\alpha, n_\beta)\) can be defined by the Boltzmann distribution,

\[ n_\alpha = e^{\frac{-E_\alpha}{k_B T}}, \quad n_\beta = e^{\frac{-E_\beta}{k_B T}}. \]  

(2.5)

A net magnetization \((M_0)\) along the \( z \)-direction is formed due to inequalities in the above populations

\[ M_0 = \frac{\gamma}{2} (n_\alpha - n_\beta) \]  

(2.6)

Using the Taylor Series expansion and considering that at relevant temperatures the energy difference is small compared to \( k_B T \), the magnetization can be approximated as

\[ M_0 \approx \frac{N \gamma^2 B_0}{4 k_B T} \]  

(2.7)
The resulting magnetization, and therefore the strength of any experimental signal, is dependent on the gyromagnetic ratio ($\gamma$), the strength of the magnetic field ($B_0$), the number of atoms, $N$, and the temperature, $T$.

**Chemical Shift Interaction**

The external magnetic field has an effect on the orbital motions of electrons which generate small, local magnetic fields. This is known as the chemical shift interaction and it results in a shift in the Larmor frequency of the nuclei. This interaction has an anisotropic, or orientation dependent ($\omega_{CSA}$), as well as an isotropic component ($\omega_{ISO}$) and therefore can be represented by the Hamiltonian:

$$H_{CS} = (\omega_{ISO} + \omega_{CSA})I_Z$$

(2.8)

![Figure 2-1](image)

**Figure 2-1** Representation of the interactions occurring in a spin $\frac{1}{2}$ system. When spins are subjected to a magnetic field, $B_0$, the unequal splitting of spins into energy levels leads to a net magnetization, and therefore a detectable NMR signal. These energy levels are altered by the chemical shift interaction, which perturbs energy levels and results in a dispersion of the nuclear species. Finally, the dipolar interaction leads to broadening of the energy levels.
The anisotropic component of the chemical shift is averaged by molecular tumbling in solution NMR, or by MAS in SSNMR, as will be discussed further in Section 2.2. The isotropic component of this interaction results in otherwise identical nuclear species becoming distinguishable from one another, as they will resonate at slightly different frequencies (Figure 2-1). The magnitude of this shift is dependent on the electronegativity of neighboring atoms, the average of molecular motions, the molecular structure, and the isotopic labelling of neighboring nuclear sites. The chemical shift is most often reported in units of parts per million (ppm), which can be calculated using

\[ \delta = \left( \frac{\omega_{ISO} - \omega_0}{\omega_0} \right) \times 10^6. \]  

(2.9)

Whereas solution NMR spectra are generally referenced with respect to an internal standard, commonly using a deuterium based field lock, SSNMR spectra are typically referenced to external standards [144]. In this work all chemical shifts are referenced to DSS (2,2-dimethyl-2-sila-pentane-5-sulphonic acid) using the \(^{13}\)C adamantane downfield peak resonating at 40.48 ppm.

**Dipolar Interaction**

The dipolar interaction is the through-space coupling between spins. As each spin itself is magnetic, it will generate a weak magnetic field which loops around the surrounding space. The through-space interaction between these dipoles can be represented by the Hamiltonian

\[ H_D = \sum_{j<k} d_{jk}(\vec{I}_j \cdot \vec{I}_k - 3 \frac{(\vec{I}_j \cdot \vec{r}_{jk})(\vec{I}_k \cdot \vec{r}_{jk})}{r_{jk}^2}) \]  

(2.10)

Where
\[ d_{jk} = \frac{\mu_0 \gamma_j \gamma_k}{4\pi r_{jk}^3} \]  

(2.11) and \( \gamma_{jk} \) is the gyromagnetic ratio of the \( j \) or \( k \) nucleus, and \( r_{jk} \) is the distance between the \( j \) and \( k \) nuclei. As the strength of the dipolar interaction is several orders of magnitude smaller than the Zeeman interaction, this expression can be simplified, by neglecting non-secular terms which do not commute with the Zeeman Hamiltonian, to

\[ H_D = \sum_{j<k} \omega_D^{jk} ((2l_j l_k) - \frac{1}{2}(l_j^+ l_k^- + l_j^- l_k^+)) \]  

(2.12)

where

\[ I^\pm = I_x \pm iI_y \]  

(2.12a)
in the homonuclear case, and to

\[ H_D = \sum \omega_D^{jk} (2l_j S_k) \]  

(2.13)
in the heteronuclear case, where in both cases

\[ \omega_D^{jk} = d_{jk} (1 - 3\cos^2\theta_{jk}) \]  

(2.14)

with \( \theta_{jk} \) being the angle between \( \vec{r}_{jk} \) and the static magnetic field, \( B_0 \) [143]. Often, \( \theta_{jk} \) is restrained by the orientation of the nuclei within a molecule. As the dipolar interaction is orientation dependent, it therefore leads to peak splitting and anisotropic broadening (Figure 2-1). These effects are eliminated by isotropic tumbling of the molecules in solution NMR or through the combination MAS and decoupling in SSNMR, as will be described further in Section 2.2.
**J-Interaction**

The J-interaction is mediated by the electrons which form chemical bonds. Though this interaction is small, it is often sufficient to modify the energy levels of the system and cause peak splitting. The J-coupling Hamiltonian can be approximated as

$$H_J = 2\pi J_{jk}(\vec{I}_j \cdot \vec{I}_k)$$  \hspace{1cm} (2.15)

in the homonuclear case and as

$$H_J = 2\pi J_{jk}(l_{jz}l_{kz})$$  \hspace{1cm} (2.16)

in the heteronuclear case, where $J_{jk}$ is the isotropic J-coupling in both instances. The heteronuclear approximation can most often be applied in the homonuclear case as well, since only carbon-carbon homonuclear pairs must be considered for this interaction in proteins, and the chemical shift difference between directly bonded carbons is often large (i.e. ~ 100 ppm for CA and CO). The J-interaction is much weaker than the dipolar interaction and therefore is often neglected in considerations of SSNMR. However, when the dipolar interaction is sufficiently averaged, though either MAS, molecular motions, or a combination of the two, the J-interaction will prevail and can cause peak splitting or be used for polarization transfer, as will be discussed further in **Section 2.3.2**.

**Interaction of Spins with Radio-Frequency Field**

Radiofrequency (RF) pulses are used to excite transitions between the $|\alpha\rangle$ and $|\beta\rangle$ states and therefore cause a perturbation of the system from its equilibrium state, or the transfer of polarization between nuclei. These pulses can be described by the Hamiltonian

$$H_{RF} = 2\omega_1 I_x \cos(\omega_{RF} t + \varphi)$$  \hspace{1cm} (2.17)
where $\omega_1$ is the strength of the field, $\omega_{RF}$ is the frequency at which the pulse is applied, and $\phi$ is the phase of the pulse. This field can be expressed as the sum of two fields which rotate in opposite directions about the static magnetic field. Of these two fields, the one which rotates in the same direction as the nuclear magnetic moment will interaction strongly with the magnetic moment, while the one which rotates in the opposite direction of the moment will have a minimal effect. Therefore the RF Hamiltonian can be simplified by expressing it in a frame which rotates at the frequency, $\omega_{RF}$ in the direction of the magnetic moment. In this frame of reference the pulse appears time independent and can be written as

$$H_{RF,rot} = \omega_1 I_x$$

(2.18)

When an RF pulse is applied on resonance with the Larmor frequency then the effective magnetic field in the rotating reference frame will be along the x- or y-axis. This results in a torque on the magnetic moment and subsequently, the magnetic moment rotating into the x-y plane and becoming distinguishable from the applied magnetic field, $B_0$.

Due to the dispersion of the chemical shifts in most samples, frequently $\omega_{rf}$ is close but not equal to $\omega_1$. Under these conditions the nuclear spin Hamiltonian is only partially determined by $H_{RF}$. When a short pulse is applied at a high frequency (a strong pulse) then $\omega_0 - \omega_{rf} \ll \omega_1$ and the Hamiltonian can be approximated to be dominated by the RF pulse and other terms can be neglected. On the other hand, when a low frequency pulse is applied (a weak pulse), then $\omega_0 - \omega_{rf} \gg \omega_1$, and the RF pulse will have a negligible contribution to the total Hamiltonian and can therefore be neglected. The property is useful when selective excitation of chemically distinct species is desired, as in N-CA or N-CO cross polarization, as is described further in Section 2.1.3.
2.1.1 Magic Angle Spinning

In solution NMR the anisotropic contributions to the dipolar and chemical shift interactions are averaged by rapid molecular tumbling and high-resolution spectra can easily be obtained. However, as molecular size increases, the rate of molecular tumbling is consequently reduced and resolution is decreased. Therefore, for larger protein complexes, such as protein-detergent micelles, whose molecular size is increased by the necessity of the hydrophobic detergent environment, solid samples are often studied and Magic Angle Spinning (MAS) is employed in order to average anisotropic contributions to the linewidth and reduce spectral overlap. In this technique, the sample rotor is tilted at an angle of ~ 54.7° with respect to the magnetic field (Figure 2-2) and the sample is physically spun about this axis. Under these conditions, the anisotropic interactions become periodically time independent and average to zero. For example, under MAS the dipolar interaction can be represented as

\[ \omega_D^{jk}(t) = d_{jk}(g_0 + g_1 \cos(\omega_r t + \varphi) + g_2 \cos(2\omega_r t + 2\varphi)) \]  

(2.19)

Where

\[ g_0 = \frac{(3\cos^2\theta_{jk} - 1)(3\cos^2\theta_m - 1)}{2} \]  

(2.19a)

\[ g_1 = \frac{3}{4}\sin2\theta_m\sin2\theta_{jk} \]  

(2.19b)

\[ g_2 = -\frac{3}{4}\sin^2\theta_m\sin^2\theta_{jk} \]  

(2.19c)

And \( \omega_r \) is the spinning frequency, \( \theta_m \) is the angle between the external magnetic field and the rotor axis, and \( \theta_{jk} \) is a polar angle of \( \vec{r}_{jk} \) in a rotating reference frame as defined by the sample rotor at a time \( t = 0 \). The time independent contribution to the dipolar interaction \( g_0 \) can be eliminated by setting \( \theta_m = \cos^{-1}(1/\sqrt{3}) = 54.7^\circ \). The remaining terms, \( g_1 \) and \( g_2 \), are modulated by time dependent terms and, therefore, with the implementation of spinning will be
averaged to zero over a rotor period. As is illustrated in Figure 2-2B, under MAS the broad powder spectra resulting from the dipolar coupling and anisotropic chemical shift are first split into a strong center-band at the isotropic chemical shift frequency as well as a number of spinning sideband peaks, occurring at multiples of the spinning frequency. As the spinning rate is increased to the magnitude of the couplings, the intensity of the side bands decreases and a single strong center-band is observed at the isotropic chemical shift frequency.

Magic-angle spinning can sufficiently suppress anisotropic interactions when the interaction Hamiltonian is self-commuting, or inhomogeneous in the sense of Maricq & Waugh [145]. When the Hamiltonian is not self-commuting, or homogenous, as in the case of multi-spin homonuclear dipolar interactions, higher-order terms will contribute to the average Hamiltonian. For homonuclear carbon and nitrogen interactions this effect is small and can generally be neglected, since the Larmor frequencies of neighboring carbon atoms are well dispersed, nitrogen has a low population, and both atoms have low gyromagnetic ratios. However, the homonuclear proton dipolar interactions are strong due to the high gyromagnetic ratio and the number of different interactions involved, and therefore the contribution from the higher-order terms cannot be ignored and MAS is generally not sufficient to average homonuclear $^1\text{H}-^1\text{H}$ interactions.
Figure 2-2 A) Representation of MAS in SSNMR. In these experiments the rotor is aligned at an angle of 54.7 ° with respect to the external magnetic field, $B_0$. Rotation of the sample around this axis induces a time averaging effect on the anisotropic dipolar and chemical shift interactions. In B) the effect of MAS on the anisotropic interactions is shown. As the MAS frequency increases, broad peaks split into a single centre-band surrounded by several spinning sidebands. Spectra were produced using SPINEVOLUTION [146].

2.1.2 Multidimensional NMR and Recoupling Techniques

Although peaks are often well dispersed and the anisotropic broadening is removed by MAS, one-dimensional spectra are most often heavily overlapped due to the large size of the membrane proteins. In order to identify and site-specifically assign resonances multidimensional spectroscopy is necessary. In these experiments, polarization is transferred between nuclei. As membrane proteins are largely rigid, polarization transfer is generally accomplished by recoupling the dipolar interaction, which is otherwise averaged by MAS, through the application of specifically tailored RF pulses. However, in more mobile regions of the protein the dipolar interaction is significantly averaged by motions and therefore cannot be efficiently recoupled. In these instances polarization transfer can be mediated though the J-interaction. In Figure 2-3 the basic polarization transfer blocks which are combined to build the more complex two- and three-dimensional experiments used in this work are presented. The basic principles of these methods are discussed below.


**Dipolar Based Recoupling**

Though the averaging of the dipolar interaction is crucial in order to resolve individual resonances, recoupling of this interaction is equally important for the implementation of multi-dimensional experiments in SSNMR. Although many heteronuclear recoupling methods have been developed (i.e., REDOR [147], TEDOR [148], etc.) in our experiments cross-polarization is used to re-couple the heteronuclear dipolar interaction. Many approaches to homonuclear recoupling have been developed as well. Here, the basic theory behind the recoupling techniques used throughout this work is discussed.

**Cross Polarization**

To transfer magnetization through cross polarization (CP) (Figure 2-3A) under MAS conditions the two heteronuclei involved are irradiated simultaneously such that the Hartmann-Hahn condition ($\omega_I \pm \omega_S = n\omega_r$) [149–151], is satisfied. Often, in order to match this condition for a range of nuclei resonating at slightly different frequencies, the irradiation is ramped around the matching condition. This condition holds only when the RF strengths used greatly exceed the offset of the isotropic chemical shifts ($\omega_{1I,S} \gg \Delta^{I,S}$). This experiment can be described using the CP Hamiltonian [152]

$$H(t) = \omega_{1I} I_x + \omega_{1S} S_x + \Delta^I I_z + \Delta^S S_z + 2\omega_d(t)(I_z S_z)$$  \hspace{1cm} (2.20)

where $\omega_{1I}$ and $\omega_{1S}$ are the RF intensities for spin I and S, and $\Delta^I$ and $\Delta^S$ are the resonance offsets. The sensitivity of this experiment to the offset of the isotropic chemical shifts allows it to be tailored for specialized purposes. When the RF strengths used in these experiments greatly exceed the offset of the isotropic chemical shifts ($\omega_{1I}^{I,S} \gg \Delta^{I,S}$), the experiment becomes insensitive to the offset of the isotropic chemical shifts, and all resonances are excited. The
signal enhancement of low gamma nuclei (\(^{13}\)C or \(^{15}\)N) through polarization transfer with the dipolar coupled proton baths is achieved in this way.

On the other hand, when the resonance offsets and the RF powers are of comparable intensities, the Hartmann-Hahn condition becomes

\[
\left( \omega_{1,\text{eff}} \pm \omega_{S,\text{eff}} = n\omega_r \right)
\]  

(2.21)

where the effective field experienced by a spin is expressed as

\[
\omega_{iS}^{\text{eff}} = \sqrt{(\Delta i S)^2 + (\omega_{1,i}^{lS})^2}
\]

(2.22)

The dependence of the matching condition on the isotropic chemical shifts under these conditions allows for frequency-selective transfers to be performed, and this type of experiment is commonly used to selectively cross polarize either CO or CA from N.

Irradiation applied directly between the two carbon frequency ranges (i.e. CA \(\approx 45-60\) ppm and CO \(\approx 170-180\) ppm, therefore irradiation would occur at \(~100\) ppm) at a high power will simultaneously recouple both interactions. However, one of these interactions can be eliminated by centering the irradiation at the frequency of one type of atom and using a lower RF power. For example, if the carbon carrier frequency is set at CO, CA will have an off-resonance component of \(~120\) ppm, or 24 kHz at 800 MHz proton frequency. Under 14.3 kHz MAS spinning, the Hartmann-Hahn condition can be matched for CO and N by choosing \(\frac{\omega_{1,N}}{2\pi} = 35.7\) kHz, and \(\frac{\omega_{1,C}}{2\pi} = 50\) kHz, whereas under these conditions the Hartmann-Hahn condition for CA and N will be mismatched by \(~5.2\) kHz, which is much larger than the dipolar coupling strength between these nuclei (\(~1\) kHz). In this way, the effective Hamiltonian is established for the N-CO pair, while the N-CA interaction will be averaged by MAS.
**Dipolar Assisted Rotational Resonance**

Dipolar assisted rotational resonance (DARR) [153, 154] is a recoupling technique that exploits the heteronuclear $^1$H-$^{13}$C and homonuclear $^1$H-$^1$H dipolar interactions to enhance $^{13}$C-$^{13}$C polarization transfer. Irradiation of the proton nuclei at the spinning frequency ($\omega_r$), as seen in **Figure 2-3C**, leads to the recoupling of the $^1$H-$^{13}$C and $^1$H-$^1$H dipolar interactions through heteronuclear rotary resonance recoupling ($R^3$) [155] and homonuclear rotary resonance recoupling (HORROR) [156], respectively. This results in broad-banded polarization transfer, which is often desirable, but can be less sensitive than specific polarization transfers. Additionally, this transfer has been observed to be significantly less efficient at fast MAS.

**Dipolar Recoupling Enhanced by Amplitude Modulation**

More selective $^{13}$C-$^{13}$C polarization transfer can be achieved through Dipolar Recoupling Enhanced by Amplitude Modulation (DREAM) [157] recoupling. In this experiment the RF pulse undergoes a tangential sweep around the HORROR [158] recoupling condition, as seen in **Figure 2-3C**. Recoupling through the HORROR condition occurs when $\omega_r = 2\omega_1$, and under this irradiation polarization is transferred with a high efficiency due to the elimination of the dependency on the $\gamma$ powder angle.

In order to achieve band-selective carbon-carbon polarization transfer between nuclei with significantly different chemical shifts (i.e., CA and CO or CA and CB), the matching condition becomes

$$\omega_{eff,i} + \omega_{eff,j} = \sqrt{\Delta^i^2 + \omega_1^2} + \sqrt{\Delta^j^2 + \omega_1^2} = \omega_r.$$  

During DREAM excitation the amplitude of the RF irradiation is not fixed, as in HORROR but is slowly swept through the HORROR condition. Therefore, separate spin pairs are matched at
different times during this sweep. Through the use of this adiabatic pulse more complete polarization transfer is obtainable.

**J-Based Recoupling**

Typically, membrane proteins consist largely of rigid secondary structural elements which are linked by more flexible, and often less structured, loops and turns. Within these less structured regions, local, sub-microsecond molecular motions of sufficiently large amplitude can lead to an averaging of the dipolar interactions. This can result in a reduction of the linewidths, but also often leads to reduced sensitivity for these residues in the dipolar-based two- and three-dimensional chemical shift correlation experiments which are used to study the well-structured regions of proteins. However, the averaging of the strong dipolar interactions facilitates through-bond polarization transfers, which are based on the comparatively weak J-interaction.

**Insensitive Nuclei Enhanced by Polarization Transfer (INEPT)**

Insensitive Nuclei Enhanced by Polarization Transfer (INEPT) experiments are based on J-couplings and favour the relatively mobile regions of proteins [82, 83, 85, 159]. The first 90° proton pulse (Figure 2-3B) creates transverse magnetization and in the subsequent delay, both the 1H-13C spin coupling and the proton chemical shift evolve. The chemical shift evolution is then refocused by the echo pulse applied to proton, while the echo pulse applied to carbon allows the second delay period to be modulated by only the H-C coupling. By setting the length of this echo period to $\tau_2 = 1/2J(C,H)$ pure antiphase coherence of the proton nuclei with respect to carbon can be obtained. This antiphase coherence ($l_xS_z$) is converted to carbon antiphase coherence ($l_zS_y$) through the application of simultaneous 90° pulses on carbon and proton. For
the direct detection of this magnetization in the HC experiment, a final refocusing period, with a delay equal to $\tau_2 = 1/4J(C,H)$ must be applied to create in-phase carbon magnetization ($S_y$).

**Total Through-Bond Correlation Spectroscopy (TOBSY)**

Total through-bond correlation spectroscopy (TOBSY) can be utilized to establish correlations between bonded carbon atoms (Figure 2-3C) [160, 161] through the J-interaction by the application of specific RF pulses which suppress the chemical shift as well as the homo- and hetero-nuclear dipolar interactions. TOBSY mixing was performed with the $P9_6^1$ mixing sequence. The basic unit of this sequence is timed to span six rotor cycles, and is broken into 9 C-elements. C-elements are designed to result in no net evolution of the nuclear spin states, when evolution under only the RF field is taken into account, while ensuring stability towards isotropic offset and RF-field inhomogeneity [162, 163]. To better suppress the dipolar interactions during this mixing period, Lee-Goldburg (LG) homonuclear decoupling is implemented [164].

**2.1.3 Pulse Sequences**

The pulse sequences used to collect all two- and three-dimensional experiments presented are shown in Figure 2-4 and Figure 2-6.

**2D Dipolar-Based Experiments**

The pulse sequences for the two-dimensional CC, HN, and NCA experiments collected throughout this work are presented in Figure 2-4A-C. These spectra are generally well resolved, (for an example see Figure 2-5A) with linewidths of $\sim$0.5 ppm in the direct and indirect $^{13}$C, $\sim$1 ppm in the indirect (truncated) $^{15}$N dimensions when collected on fully protonated samples of
ASR and PR, and proton linewidths of ~0.2 ppm when spectra are collected on perdeuterated samples of PR back-exchanged with buffer containing 40% H$_2$O. Though these experiments all result in a number of well-resolved, isolated peaks, significant overlap is still prevalent in the spectra due to the large size of the proteins studied and the high occurrence of hydrophobic residues (Figure 2-5A). These experiments are therefore useful for initial characterization of sample quality under differing experimental conditions, whereas if a site-specific analysis of changes in intensity or chemical shifts across different samples is required, three-dimensional spectroscopy is often necessary in order to resolve the maximum number of resonances.

**Figure 2-3** Basic pulse sequences used for heteronuclear recoupling using A) dipolar based cross polarization and B) J-based INEPT recoupling. C) Homonuclear recoupling was achieved through either dipolar-based methods such as DARR and DREAM or the J-based TOBSY method. TOBSY is a rotor-synchronized spin-locking pulse sequence with nine-fold symmetry. The basic unit of the TOBSY sequence spans six rotor cycles, and is broken into 9 C-elements, which are designed to result in no net evolution of the nuclear spin states.
Figure 2-4 Pulse sequences for dipolar based experiments A) 2D C-C CPMAS-DARR B) 2D \(^1\text{H}\)-detected H-N C) 2D NCA and D) 3D CANCO and E) 3D NCACB chemical shift correlation experiments. Filled and hollow bars represent \(\pi/2\) and \(\pi\) pulses respectively. In all experiments TPPI [165] phase-sensitive detection is used in the indirect dimensions. In panel A) the following phase cycling was used: \(\phi_1 = x, x, x, y, y, -x, -x, -x, -y, -y, -y; \phi_2 = y, y, -y, -y, -x, -x, -y, -y, y, x, x, x, -x, -x, -y, -y, -y, -y, -x, -x, -x; \phi_3 = y, y, y, y, x, x, x, -y, -y, -y, -y, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x, -x,
Figure 2-5 A) NCA spectra of ASR reconstituted in proteoliposomes. Representative two-dimensional planes of B) the 3D CANCO and C) NCACX experiments collected on ASR. In C) the NCACX experiment was collected with the pulse sequence shown in Figure 2-4E with 50 ms of DARR mixing in the place of DREAM mixing. In all spectra peaks are labeled based on the N[i]–Ca[i] assignments.

**CANCO and NCACB Experiments**

Three-dimensional (3D) CANCO and NCACB experiments were collected using the pulse sequences shown in Figure 2-4D and Figure 2-4E respectively. The 3D CANCO experiments were performed to obtain backbone assignments, whereas 3D NCACB experiments provide information on side chain atoms and help validate the identification of shifted residues from the CANCO data in cases of overlapping peaks or where there are multiple assignments.
possible. Through the use of three-dimensional spectroscopy, significantly more residues can be resolved than in the two-dimensional spectra, as can be seen in Figure 2-5B, C.

In the CANCO experiment, $^{15}\text{N}/^{13}\text{CA}$ and $^{15}\text{N}/^{13}\text{CO}$ band-selective transfers [166] were implemented in order to specifically obtain CA[i]-N[i]-C[i-1] correlations while in the NCACB experiment, a band selective $^{15}\text{N}/^{13}\text{CA}$ polarization transfer is implemented and followed by DREAM [157] recoupling with tangential sweep around the HORROR [156] condition which was optimized to facilitate CA to CB polarization transfer. Peaks in the resulting spectra therefore represent N[i]-CA[i]-CB[i] spin systems. The CA[i]-N[i]-C[i-1] and N[i]-CA[i]-CB[i] systems can be combined by comparing the overlapping N[i] and CA[i] resonances, allowing for confident reassignment in the case of shifted peaks in these spectra.

**Two-Dimensional HC and HCH experiments**

Two-dimensional $^{13}\text{C}$- and $^1\text{H}$-detected $^1\text{H}-^{13}\text{C}$ INEPT HSQC spectra were recorded using the standard pulse sequences shown in Figure 2-6B and Figure 2-6C, respectively [167–169]. Although an in-phase coherence is generally required for indirect detection in SSNMR, due to the need for continuous decoupling, we find that the residues excited by INEPT are sufficiently mobile to yield narrow carbon linewidths during indirect detection, even in the presence of only a single refocusing $\pi$-pulse applied to protons. Therefore, to avoid the signal loss in our $^1\text{H}$-detected experiments, which would occur during the anti-phase to in-phase $^{13}\text{C}$ coherence conversion period, and during the conversion back to proton in-phase coherence through INEPT, we indirectly record carbon chemical shifts thorough the evolution of carbon antiphase magnetization.
Figure 2-6 Pulse sequences for J-based experiments. A) $^{13}$C-$^{13}$C INEPT-TOBSY B) $^{13}$C detected and C) $^1$H-detected $^1$H-$^{13}$C INEPT HSQC experiments, and D) 3D (H)CHH INEPT based experiments. Filled and hollow bars represent π and π/2 pulses, respectively. In panel A) TOBSY mixing was performed with the $P_9^6$ mixing sequence at a MAS frequency of 14.3 kHz, and Lee-Goldburg decoupling [164] was applied during this period. In panel A) the following phase cycling was used: $\phi_1 = x, x, y, y, y, y, -x, -x, -x, -x, -x, -y, -y, -y, -y; \phi_2 = x, x, x, x, y, y, -y, -y; \phi_3 = x, x, x, x, y, y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y,$ $\phi_4 = x, x, x, x, y, y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y,$ $\phi_{rec} = x, x, x, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y,$ and phase-sensitive detection was obtained in the indirect $t_1$ dimension by incrementing $\phi_3$ by 90°. In panel B), the following phase cycling was used: $\phi_1 = x, x, x, x, x, x, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y,$ $\phi_3 = x, x, x, x, x, x, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y, -y,$ $\phi_{rec} = x, x, x, x, x, x, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y, y,$ and phase-sensitive detection was obtained in the indirect $t_1$ dimension by incrementing $\phi_1$ by 90°. In panel C), phase-sensitive detection was obtained in the indirect $t_1$ dimension by incrementing $\phi_1$ and $\phi_2$ by 90°. In D), phase-sensitive detection was obtained in the indirect $t_1$ and $t_2$ dimensions by incrementing $\phi_1$, $\phi_2$ and $\phi_3, \phi_4, \phi_{rec}$, respectively, by 90°. In all experiments, $\tau_1$ was optimized to 877 μs and $\tau_2$ was optimized to 1.4 ms, and TPPI [165] was used for phase-sensitive detection.

(H)CHH Experiment

The two-dimensional $^1$H-detected INEPT-HSQC experiment can be easily extended through the incorporation of a proton-proton mixing period. Mixing based on the nuclear Overhauser effect (NOE) [170–172] was found to result in effective polarization transfer, with the additional benefits of ease of set-up and lack of high-power irradiation, therefore allowing for faster experimental recycling. An additional echo period is added at the end of the pulse sequence to suppress broad proton signals which arise from immobile residues as a result of the final mixing step. Three-dimensional (H)CHH chemical shift correlation experiments were recorded using the pulse sequence shown in Figure 2-6D.
2.1.4 Fast MAS and PACC

Our standard SSNMR experiments are performed on 3.2-mm probes with MAS of 14.3 kHz. Under these conditions large amounts of isotopically-labeled protein are necessary to fill the sample rotor and maximize the experimental sensitivity, and high-power decoupling is necessary in order to fully decouple heteronuclear interactions and obtain narrow linewidths. Often, the large amounts of protein required to fill a 3.2-mm rotor cannot be easily produced, and in these cases it is beneficial to perform experiments on smaller diameter rotors, which have the additional benefit of being designed to spin at much higher frequencies (>40 kHz). In these systems, the sensitivity lost due to decreased sample volume can be compensated for in several ways, such as the increased coil efficiency, accelerated acquisition schemes, and the potential for proton detection.

Due to the more complete averaging of the dipolar couplings, fast spinning frequencies enable the use of low-power decoupling [63–65] and facilitate proton detected experiments through increases in resolution [49, 50, 62]. The use of low-power decoupling removes the need for long recycle delays during which the probe recovers from the application of high-power RF. Therefore, experiments can be recycled more quickly and the recycle delay becomes governed by the $T_1$ relaxation (described further in Section 2.1.6) of the nuclei. Recycle delays can be further reduced through the incorporation of paramagnetic dopants to the sample which reduce $T_1$ values. Such paramagnetic dopants cause rapid spin-lattice relaxation of nearby protons while rapid inter-proton spin diffusion redistributes the paramagnetic relaxation enhancement effect throughout the sample. Recently, Ishii and co-workers used fast MAS (40 kHz), low power decoupling, and short recycle delays enabled by exogenously added Cu(II)-EDTA as a paramagnetic relaxation agent to develop the paramagnetic relaxation-assisted condensed data.
collection (PACC) scheme [173]. They have demonstrated that the PACC scheme enables SSNMR studies of small, nanomolar amounts of $^{13}$C,$^{15}$N-labeled biomolecules. However, it is important to quantify the extent to which the PACC scheme can compensate for the drastic reduction in the amount of sample, and therefore determine the utility of these rotors for studies of membrane proteins using $^{13}$C- and $^1$H-detected experiments.

2.1.5 Relaxation

Relaxation describes the processes through which the NMR signal excited through RF irradiation returns to its equilibrium state. The bulk relaxation is characterized by two time constants, $T_1$, the longitudinal relaxation time, and $T_2$, the transverse relaxation time. The longitudinal relaxation time describes the time for the net magnetization to return to its equilibrium position along the z-axis following the application of a perturbing RF pulse, while the transverse relaxation time is characterized by the decay time of the magnetization in the transverse plane.

$T_1$ Relaxation

The longitudinal relaxation parameter, $T_1$, can be measured using a saturation recovery experiment, as show in Figure 2-7A. The purpose of the initial 90° pulse and delay, $d$, is to completely saturate the $^1$H magnetization, resulting in the magnetization being completely destroyed when $\tau$ is set to zero. As $\tau$ is incremented the polarization recovery will be governed by the equation

$$M_Z(t) = M_{eq}(1 - e^{\frac{t}{T_1}})$$  \hspace{1cm} (2.23)
where \( M_{eq} \) is the equilibrium magnetization and \( M_Z \) is the component of the magnetization which is parallel with the external magnetic field, \( B_0 \).

**\( T_2 \) Relaxation**

When a net magnetization is created in the transverse plane is created through RF irradiation, magnetization will precess around the static magnetic field at the Larmor frequency. Homogenous line broadening effects, such as molecular motions or fluctuations in the magnetic field, will lead to a decay of the transverse magnetization. This relaxation is governed by the time constant, \( T_2 \). In SSNMR, the residual couplings between spins leads to additional relaxation, and therefore the relevant time constant to describe the transverse relaxation is the coherence lifetime, \( T_2' \). This parameter can be measured using a spin-echo experiment, as shown in Figure 2-7B. During the echo-period, inhomogeneous interactions are refocused by the 180° pulse, and thus do not contribute to the signal decay during this time. By performing a series of spin-echo experiments with increasing \( \tau \) [143], the observed signal decay in the resulting signal can be modeled by

\[
M_X(t) = M_{eq}e^{-\frac{2\tau}{T_2'}}
\]

(2.24)
Figure 2-7 Pulse sequences used for A) $T_1$ and B) $T_2$' relaxation measurements.

2.2 Fourier Transform Infrared Spectroscopy

Fourier transform infrared spectroscopy (FTIR) is a technique which is widely used in the structural characterization of proteins [174]. As separate functional groups absorb at specific frequencies, information on many components of samples can be obtained. For example, proteins generate several infrared-active vibrational modes, of which the amide I and amide II modes have been the most extensively analyzed. In natural abundance samples, the amide I band appears between 1600-1700 cm$^{-1}$ and is generated primarily by the C=O stretching vibration, with minor contributions from CN and CCN out-of-plane bending vibrations and the NH in-plane bending mode. This band is therefore sensitive to the secondary structure, with the $\alpha$-helical contributions appearing between 1647-1658 cm$^{-1}$ and the $\beta$-strand contributions appearing between 1632-1638 and 1675-1695 cm$^{-1}$, and to isotopic labelling with $^{13}$C. The amide II band appears between 1540-1570 cm$^{-1}$ and is generated primarily by NH in-plane bending, along with CN stretching and other vibrations of the amide group. This peak is therefore sensitive to H-D
exchange and isotopic labelling with $^{15}$N and $^{13}$C. In addition to protein peaks, many lipid peaks can be observed in FTIR spectra, allowing for the determination of the protein to lipid ratio in membrane protein samples.

2.3 Circular Dichroism Spectroscopy

Circular Dichroism (CD) measures the difference in absorption of left- and right-hand circularly polarized light. This technique is a well-established as a tool for analysis of the structure and conformational changes of proteins, and can be used to study proteins without the need for crystallization. Folded proteins absorb in the UV range (250-300 nm) primarily due to the presence of aromatic residues and disulfide bonds. In this region the spectra are sensitive to the tertiary structure whereas in the far-UV region (< 250 nm) the spectral shape is determined primarily by protein secondary structure.

In addition to the application to secondary and tertiary structural analysis, CD in the visible range can be used to determine the oligomeric organization of retinal-containing proteins (as well as other chromoproteins). The appearance of a bilobed shape in the visible CD spectra of oligomers formed by microbial rhodopsins is related to the excitonic coupling of retinal chromophores of the monomers, with the exact shape of the spectra depending on the relative orientation of retinals, as well as additional contributions from aromatic sidechains [175–177].

2.4 Small Angle X-Ray Scattering

The scattering of x-rays at small angles provides structural information on inhomogeneities of the electron density. Though the resolution of small angle x-ray scattering (SAXS) is not sufficient to determine the atomic-level structure of samples, low resolution
models of proteins can be obtained from the precise scattering patterns obtained from these experiments. In addition, data such as membrane protein pore size, surface to volume ratio, and lattice type and dimensions can be determined [178].

When molecules are densely packed, as in a membrane protein lattice, the inter-molecular distances will be on the same order of magnitude as the intra-molecular distances. Therefore, the interference pattern will contain contributions from both types of distances. The additional interference created by inter-molecular distances multiplies with the form factor of the single molecule and is known as the lattice factor. When samples contain a crystal lattice of significant domain size, a set of narrow and intensive peaks at well-defined angles known as Bragg peaks will appear in the SAXS spectra. It can be shown that the ratios of the peak positions on the q-scale have typical values, which reveal the crystal symmetry.
Chapter 3 : Materials and Methods
3.1 Materials

Common chemicals of reagent grade were purchased from either Fisher Scientific (Unionville, Ontario, Canada) or Sigma-Aldrich (Oakville, Ontario, Canada). Isotopically labeled compounds, such as $^{15}\text{NH}_4\text{Cl}$ and $^{13}\text{C}_6$-glucose were obtained from Sigma-Aldrich or Cambridge Isotope Laboratories (Andover, MA). The Ni$^{2+}$NTA (nitrilotriacetic acid) agarose resin was purchased from Qiagen (Mississauga, Ontario, Canada). Lipids were purchased from Avanti Polar Lipids (Alabaster, AL).

3.2 Expression and Purification of Membrane Proteins

Many samples were used throughout the chapters of this thesis. In Chapter 4 we investigate the incorporation of paramagnetic labels into our samples using N148C ASR and UCND PR. In Chapter 5 we perform proton detected experiments on the mobile regions of membrane proteins. These methods are tested on Myelin Basic Protein (MBP), which is generally very mobile, before being extended to ASR, which is mainly rigid but contains mobile loop regions. In Chapter 6 we perform 3D experiments on UCN ASR as well as a sample of ASR prepared in the E. coli membrane (EM-ASR). In Chapter 7 we begin to specifically investigate the effect of varying the lipid environment on the structure and dynamics of PR. Here, I present the sample preparation protocols which were mainly established prior to my work, whereas those which were developed over the course of this work are included in their respective chapters.
3.2.1 Anabaena Sensory Rhodopsin

[U-\textsuperscript{13}C,\textsuperscript{15}N]–labeled (UCN) and natural abundance (NA), C-terminally truncated, Histagged, wild-type (ASR) and N148C mutant ASR (N148C ASR) were produced according to a protocol published for wild-type (WT) ASR [114, 123]. Proteins were expressed in BL21-Codonplus-RIL \textit{E. coli} grown on M9 minimal medium at 30 °C, using 4 g glucose and 1 g of NH\textsubscript{4}Cl per litre of culture as the sole carbon and nitrogen sources. For isotopically labelled samples \textsuperscript{13}C\textsubscript{6}-glucose and \textsuperscript{15}NH\textsubscript{4}Cl were used. When cultures reached a target cell density of \(A_{600}=0.4\) OD protein expression was induced by the addition of IPTG to a concentration of 1 mM. At this time retinal was also added to a final concentration of 7.5 μM. After ~21 hrs the cells were collected by centrifugation, pre-treated with lysozyme (12 mg/L of culture) and DNAase I (600 units per litre of culture) and then broken by sonication. The membrane fraction was then solubilized in 1 % DDM (\(n\)-dodecyl \(\beta\)-d-maltoside) at 4 °C, and purified following the batch procedure described in the Qiagen Ni\textsuperscript{2+}-NTA resin manual. Approximately 7 mg of N148C ASR or 5 mg of WT ASR could be purified from one litre of culture. The molar amount of ASR was determined by the absorbance of opsin-bound retinal, using the extinction coefficient of 48,000 M\textsuperscript{-1}cm\textsuperscript{-1} [179]. Purified proteins were buffer-exchanged using an Amicon Ultra-15 10K centrifugal filter (Millipore, Massachusetts, MA, USA) into pH 8.0 buffer (5 mM Tris, 10 mM NaCl, 0.05% DDM), and concentrated to ~1 mg/ml (36 μM).

3.2.2 Proteorhodopsin

NA and UCN PR samples were produced as described previously [114, 134, 140]. The procedure for producing and purifying these proteins is the same as that described above for ASR, except that PR solubilization was performed using 1% Triton X-100, which was
exchanged for DDM on the resin during purification. \([\text{U}^{-13}\text{C},^{15}\text{N},^2\text{H}]-\text{PR}\) (UCND PR) samples were produced using a derivative of the above protocol which was optimized for expression in 100% D\(_2\)O, and has been described previously [59]. For these samples, the cells were grown using \([\text{U}^{-2}\text{H},^{13}\text{C}]-\text{labeled glucose (4 g)}\) as the sole carbon source and with D\(_2\)O (99.9% purity) in the expression media. The cells were grown to a high cell density in 2 ml (24 hrs) and then 25 ml (15 hrs) volumes before being added to the 1 L culture to create a target cell density of \(A_{600}= 0.1\) OD. After induction with IPTG, retinal was added at intervals of \(~7\) hrs to a final concentration of 7.5 \(\mu\text{M}\). Approximately 15 mg of UCND PR was purified from one litre of culture. The molar amount of PR was determined by the absorbance of opsin-bound retinal, using the extinction coefficient of 44,000 M\(^{-1}\)cm\(^{-1}\) [129].

3.2.3 Reconstitution into Synthetic Lipids

To prepare liposomes, lipids were mixed in chloroform and dried under either nitrogen gas (for unsaturated lipids) or air (saturated lipids) before being further dried through lyophilization (unsaturated lipids) or desiccation (saturated lipids). The lipids were hydrated and collected in buffer and added to protein samples at the appropriate protein to lipid ratio along with a small amount of Triton-X. For standard samples DMPC and DMPA were mixed at a 9:1 ratio (w/w), at a protein:lipid ratio of 2:1 (w/w). The protein/lipid mixture is then incubated for at least 6 h at 4º C before Bio-Beads SM (Bio-Rad Laboratories) were added for detergent removal. The functionality and protein to lipid ratio of the reconstituted protein have been tested using visible and FTIR spectroscopy, as described previously [114, 180].
3.3 Magic Angle Spinning Solid-State NMR spectroscopy

3.3.1 Equipment

Three spectrometers were used over the course of this work. The majority of experiments were performed on a Bruker Biospin Avance III standard bore spectrometer operating at 800.230 MHz. On this spectrometer three separate probes were used. The work presented in Chapter 4 was performed on a 1.3-mm $^2$H-$^1$H-$^{13}$C-$^{15}$N probe used in triple resonance $^1$H/$^{13}$C/$^{15}$N mode. This probe can achieve spinning frequencies of $> 50$ kHz, which leads to better averaging of the anisotropic interactions which are removed by high power decoupling at lower spinning frequencies. The removal of the requirement for high-power decoupling allows for experiments to be recycled more quickly, and the investigation of the limitations of this approach on our samples is the focus of Chapter 4. The work in Chapter 5 performed on ASR was collected using a Bruker 3.2-mm TL2 $^1$H-$^{13}$C-$^{15}$N MAS probe. This probe was chosen as we have found that it provides better sensitivity than the E-FREE probe for proton-detected experiments. Finally, all experiments performed on UCN ASR and samples of uninduced cells (EM-NIC) and the one- and two-dimensional experiments performed on ASR samples prepared in the E. coli membrane (EM-ASR) in Chapter 6, as well as all experiments performed in Chapter 7 were collected using a 3.2-mm $^1$H-$^{13}$C-$^{15}$N E-FREE MAS probe.

Two other spectrometers were also used. All 3D experiments on EM-ASR were performed on a 900 MHz Bruker Avance III NMR spectrometer equipped with a low-E triple resonance $^1$H-$^{13}$C-$^{15}$N 3.2-mm magic angle spinning probe designed and built at the National High Magnetic Field Laboratory, Tallahassee, FL [181]. The high sensitivity and fast recycle delays afforded by the 900 MHz probehead were critical for the detection of small amounts of ASR in this sample. All NMR experiments on MBP were performed on a Bruker Biospin
Avance III spectrometer operating at 600.13 MHz using a Bruker 3.2-mm TL2 $^{1}\text{H}-^{13}\text{C}-^{15}\text{N}$ MAS probe, as the resolution and sensitivity available from the 800 MHz spectrometer was not necessary for these preliminary experiments.

Specific details of the experimental set-up and parameters for each spectrometer and probe are given in the respective chapters. Pulse sequences and their descriptions are presented in Section 2.1.3.

3.3.2 Data analysis

One-dimensional spectra and the coherence life time ($T_2'$) [182] analysis and estimations of bulk $T_1$ were performed using Topspin 3.1 (Bruker, Karlsruhe, Germany). Two- and three-dimensional spectra were processed with NMRPipe [183] and a cosine bell squared function was used to apodize the data in all dimensions. Noise analysis and peak picking were performed in the CARA [184] environment.
Chapter 4:
Paramagnetically-enhanced solid-state NMR spectroscopy of membrane proteins at fast magic angle spinning

This work has been published:
Magic angle spinning SSNMR experiments often suffer from low sensitivity due, in part, to the long recycle delays required for magnetization and probe recovery, as well as detection of low gamma nuclei. In ultrafast MAS experiments sensitivity can be enhanced through the use of low power decoupling sequences combined with paramagnetically-enhanced relaxation times to reduce recycle delays, as well as proton detected experiments. In this work we investigate the sensitivity of $^{13}$C and $^1$H detected experiments applied to 27 kDa membrane proteins reconstituted in lipids and packed in small 1.3-mm MAS NMR rotors. We demonstrate that spin diffusion is sufficient to uniformly distribute paramagnetic relaxation enhancement, provided by either covalently bound or dissolved CuEDTA, over 7TM α-helical membrane proteins. Using paramagnetic enhancement and low power decoupling in carbon detected experiments we can recycle experiments ~13 times faster than under standard conditions (3.2-mm probe). However, due to the small sample volume the overall sensitivity per unit time is still lower than that seen in the 3.2-mm probe. Proton detected experiments, however, showed increased efficiency and it was found that the 1.3-mm probe could achieve sensitivity comparable to that of the 3.2-mm per unit time. This is an attractive prospect for samples of limited quantity, as this allows for a reduction in the amount of protein that needs to be produced without the necessity for increased experimental time.

4.1 Introduction

Typically, SSNMR studies require milligram quantities of sample and employ moderate spinning frequencies (10-20 kHz) and high power RF decoupling. Such experimental conditions require long recycle delays between successive scans (1.7 s or longer) to avoid probe and sample damage caused by the application of high power RF.
Recent advances in probe design have extended the range of spinning frequencies up to \( \sim 110 \text{ kHz} \) \([46, 48]\). It has recently been demonstrated that recycle delay requirements can be reduced under fast MAS conditions, as the application of fast spinning frequencies (40 kHz or higher) averages anisotropic interactions more efficiently, eliminating the need for high power decoupling \([63–65]\). In addition, these MAS frequencies facilitate proton detected experiments \([49, 50, 62]\) and suppress coherent contributions to relaxation rates, thus providing direct access to dynamic information \([66, 67]\). However, an important challenge in the use of these rotors is the drastically reduced sensitivity caused by the decrease in sample volume. For example, when experiments using 1.3-mm and 3.2-mm rotors are compared the volume ratio contributes a factor of \( \sim 15 \) or more to the relative sensitivity. However, as these rotors have a higher sensitivity per unit sample, ultrafast MAS probes are an attractive option for the study of proteins that can only be produced in small quantities, particularly eukaryotic membrane proteins \([68–70]\). Therefore, it is important to investigate the extent to which the benefits offered by ultrafast MAS, such as the increased coil efficiency, accelerated acquisition schemes, and the potential for proton detection, can compensate for this drastic reduction in the sample volume.

When entirely low-power decoupling and polarization transfer schemes are combined with the introduction of paramagnetic dopant to reduce proton relaxation times, which govern the recycle delay when the necessity for probe recovery is removed, recycle delays can be reduced by an order of magnitude \([185]\). Such paramagnetic dopants cause rapid spin-lattice relaxation of nearby protons while rapid inter-proton spin diffusion redistributes the paramagnetic relaxation enhancement effect throughout the sample. Recently, Ishii and co-workers used fast MAS (40 kHz), low power decoupling, and short recycle delays enabled by exogenously added Cu(II)-EDTA as a paramagnetic relaxation agent to develop the paramagnetic relaxation-assisted
condensed data collection (PACC) scheme [173]. They have demonstrated that the PACC scheme enables SSNMR studies of small, nanomolar amounts of $^{13}$C,$^{15}$N-labeled biomolecules.

The PACC scheme has further been extended to systems containing covalently attached paramagnetic tags. The Jaroniec lab has demonstrated similar spin-lattice relaxation time reduction in microcrystalline GB1 (the first immunoglobulin binding domain of protein G), using EDTA tags loaded with Cu(II) (CuEDTA) covalently attached to cysteine residues introduced by site directed mutagenesis [186]. An additional benefit of the covalent introduction of Cu(II) is that it allows for quantification of PREs through the $^{15}$N spin-lattice relaxation rates, $T_1$, which can be used to extract distance restraints for protein structure determination [187]. The effects of paramagnetic relaxation enhancement, and the performance of the PACC scheme, were recently investigated in membrane proteins. The groups of Ramamoorthy and Rienstra have demonstrated that copper-chelated lipids and CuEDTA added to the buffer can be used to speed up SSNMR data acquisition in membrane proteins [188, 189]. The latter study, which was performed on the transmembrane protein DsbB, employed 1.6-mm rotors and spinning frequencies of 36 kHz. Although very short proton $T_1$ values of 50-80 ms were obtained, relatively long recycle delays of 0.3-0.5 s were still necessary to avoid sample damage and to minimize probe heating effects.

In this work we investigate the performance of the PACC scheme when applied to larger, 27 kDa membrane proteins reconstituted in lipids and packed in small 1.3-mm MAS NMR rotors. We examine the sensitivity of $^{13}$C-detected experiments using a lipid-reconstituted 7TM $\alpha$-helical protein, Anabaena Sensory Rhodopsin (ASR) [118]. ASR forms stable trimers in the lipid environment [93, 121] and gives well-resolved, high signal-to-noise ratio spectra with typical carbon and nitrogen line widths of 0.5 ppm [123]. We evaluate two different ways of
introducing paramagnetic dopants, through the addition of CuEDTA into the buffer and by covalently attaching CuEDTA to a cysteine introduced by mutagenesis.

In contrast to previous studies on GB1 [186], we find that in the larger, 27 kDa ASR the single covalently attached CuEDTA leads to only a small overall decrease in proton spin-lattice relaxation times. On the other hand, we found that CuEDTA added to the buffer results in significant, and fairly uniform, shortening of $^1$H $T_1$ relaxation in ASR samples, increasing in effectiveness as the CuEDTA concentration increases. Overall, although we observe a significant increase in the signal-to-noise ratio per unit sample, we conclude that even with extensive paramagnetic doping and fast experimental recycling, data collection for $^{13}$C-detected experiments with the 1.3-mm probe is ultimately much less efficient than with the 3.2-mm probe.

We further investigate the effects of fast MAS and paramagnetic doping on proton detection experiments on a fully deuterated 7TM protein, proteorhodopsin (PR), which is similar to ASR in terms of its stability and the obtained spectral quality [114, 140]. For proton-detected experiments the sensitivity per unit time of the 1.3-mm probe is found to be comparable to that of the 3.2-mm probe, which allows for experiments of similar sensitivity to be performed on a drastically reduced sample quantity, in agreement with previous observations on the $\alpha$-spectrin SH3 domain when the RAP (Reduced Adjoining Protonation) labeling scheme was utilized [190].

4.2 Methodology

4.2.1 Incorporation of Paramagnetic Labels

To investigate the paramagnetic relaxation enhancement (PRE) effect on our membrane proteins four samples of uniformly labeled N148C ASR were created: (1) a control sample
without copper (ASR), (2) a sample with CuEDTA covalently attached to 148C (ASRCu), and samples with (3) 40 mM (ASR40Cu), and (4) 80 mM CuEDTA (ASR80Cu) in the solvent (summarized in Table 4-1). The latter three samples are estimated to contain approximately 1, 4 and 8 paramagnetic centers per protein, respectively. The number of CuEDTA molecules per protein was estimated using the relative intensities of the water and lipid peaks in a 1D proton detected spectra. Given these relative intensities and the known protein to lipid ratio, the volume of water available per protein molecule can be calculated, and thus the number of CuEDTA molecules in that volume of water for a given concentration can be determined.

It has been previously shown that the native cysteine residues present in ASR (at positions 134, 137, and 203) are not accessible for reaction, as these residues are buried in the intramembrane region [121]. Thus, non-native, solvent-accessible cysteines introduced into the wild-type background of ASR can be selectively labeled through the addition of cysteine-specific paramagnetic tags. N148 was chosen as a mutation site as it is located on the E-F loop and therefore should be accessible for binding and lead to good dispersion of the paramagnetic labels throughout the trimer. Furthermore, expression of N148C ASR has been found to be more efficient than that of WT ASR and other cysteine mutants, such as S26C, which we have examined [121]. Direct comparisons of solid-state NMR spectra between WT ASR and the N148C mutant reveals that the mutant retains essentially the same fold as the wild-type, with only a few perturbations to line widths and chemical shifts being observed in residues located close to the mutation site.

To obtain covalently CuEDTA-labeled ASR (ASRCu in the following) solubilized N148C ASR was incubated for 1 hour at room temperature with a 20-fold molar excess of [S-Methanethiosulfonylcysteaminy]ethylenediamine-N,N,N’,N’-Tetraacetic Acid (MTS-EDTA,
purchased from Toronto Research Chemicals Inc., Ontario, Canada). Unreacted reagent was removed by buffer-exchange, and the completeness of spin labeling was monitored and confirmed by Matrix Assisted Laser Desorption Ionization Mass Spectroscopy (MALDI-MS).

ASR samples with CuEDTA incorporated in the buffer (ASR40Cu and ASR80Cu) were prepared by first mixing solubilized protein with a 20 fold molar excess of methyl methanethiosulfonate (MMTS, purchased from Toronto Research Chemicals Inc., Ontario, Canada) [191] to block the cysteine SH group. Unreacted reagent was then removed by buffer-exchange and the proteins were reconstituted into liposomes as is described in Section 3.2.3. After reconstitution liposomes were resuspended in a large volume of buffer containing the desired concentration of ethylenediaminetetraacetic acid (EDTA), tetrasodium salt, and copper sulfate and incubated for 1hr at 4º C before being ultracentrifuged for packing. It is important to note that the CuEDTA concentration was calculated excluding the volume of the protein and lipids in the sample, reporting on its amount in the solvent surrounding the proteoliposomes. For consistency, these ASR samples were all prepared using N148C ASR.

For proton detected experiments two samples of UCND PR were made: (1) a control sample without copper, which we denote as deuterated PR (DPR) and (2) a sample with 80 mM CuEDTA in the solvent (DPR80Cu). For the DPR and DPR80Cu samples CuEDTA incorporation was performed as described above for ASR. After reconstitution the UCND PR proteoliposomes were incubated in a 40:60 H2O:D2O buffer for ~24 h before the addition of CuEDTA and collection by ultracentrifugation.

4.2.2 Solid-State NMR Spectroscopy

For each sample approximately 0.75-1 mg of UCN N148C ASR or UCND PR was center-packed in a 1.3-mm rotor and sealed using silicone plugs. All NMR experiments were
performed on a Bruker Biospin Avance III standard bore spectrometer operating at 800.230 MHz using a 1.3-mm $^2$H-$^1$H-$^{13}$C-$^{15}$N probe used in triple resonance $^1$H/$^{13}$C/$^{15}$N mode at a spinning rate of 50 kHz and an effective temperature of ~15 °C, unless otherwise stated. As the heating induced by MAS in this probe is extensive, cooling was achieved using nitrogen gas cooled through a liquid nitrogen heat exchanger. The effective sample temperature was calibrated using the temperature dependence of the chemical shift of $^{79}$Br in KBr [192]. Although higher spinning frequencies were achievable, maintaining the sample temperature in the range optimal for sample stability required a much higher flow of cold nitrogen gas, resulting in overcooling of the shims in a standard bore magnet.

Typical $\pi/2$ pulses were 1.6 $\mu$s for $^1$H, 4 $\mu$s for $^{13}$C, and 5.6 $\mu$s for $^{15}$N. The $^1$H/$^{15}$N CP [151] contact time was 2 ms, with a constant RF field of 30 kHz on nitrogen, while the proton lock field was ramped linearly around the $n=1$ Hartmann-Hahn condition [149]. $^{15}$N/$^{13}$CA band-selective transfers [166] were implemented with a contact time of 5 ms with a constant lock field of 20 kHz strength applied on $^{15}$N, while the $^{13}$C field was ramped linearly (10% ramp) around 30 kHz. No proton decoupling was applied during $^{15}$N/$^{13}$CA CP. Low power TPPM48 [65] proton decoupling optimized around 12 kHz was used during $^{15}$N and $^{13}$C chemical shift evolutions. Proton detected $^1$H-$^{15}$N correlation experiments used WALTZ-16 [193] for $^{15}$N decoupling during direct acquisition.

The bulk proton $T_1$ of each sample was measured using the pulse sequence shown in **Figure 2-7A** (discussed in the Results section) and the recycle delay $D_1$ was adjusted to $1.3xT_1$ to maximize the signal to noise in a given time ($S/N$) as described by **Equation 4-1**:

$$
(S/N)_T \propto (1 - e^{-D_1/T_1}) \sqrt{1/D_1}.
$$

(4.1)
Here, we account for both the signal loss due to incomplete magnetization recovery, \((1 - e^{-D_1/T_1})\), and the signal enhancement due to the increased number of scans, \(\sqrt{1/D_1}\). Samples were stable over the course of the NMR experiments, and showed no signs of degradation even in measurements with recycle delays of 100 ms.

To investigate site-specific \(^1\text{H}\) relaxation enhancement and possible chemical shift and line width perturbations caused by the different paramagnetic labeling schemes, 2D NCA experiments were run on all four samples: ASR, ASRCu, ASR40Cu, and ASR80Cu. The pulse sequence for this experiment is given in Figure 2-4C. All spectra were collected over similar lengths of time (~ 13 hrs), and with recycle delays \(D_1=1.3T_1\) adjusted according to the measured bulk \(T_1\) values (Table 4-1). As \(T_1\) values are progressively shortened with increasing levels of paramagnetic doping, one expects that the sensitivity would increase according to the increasing number of scans.

To further investigate the effect of CuEDTA on our deuterated samples we collected 2D \(^1\text{H}\)-\(^{15}\text{N}\) heteronuclear correlation spectra on both PR samples using the pulse sequence show in Figure 2-4B.

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>ASR</th>
<th>ASRCu</th>
<th>ASR40Cu</th>
<th>ASR80Cu</th>
<th>DPR</th>
<th>DPR80Cu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Preparation Conditions</td>
<td>ASR with MTS-EDTA bound to C148</td>
<td>UCN N148C ASR with MTS-CuEDTA bound to C148</td>
<td>UCN N148C ASR with MMTS bound to C148 with 40 mM CuEDTA in the solvent</td>
<td>UCN N148C ASR with MMTS bound to C148 with 80 mM CuEDTA in the solvent</td>
<td>UCND PR</td>
<td>UCNDPR with 80 mM CuEDTA in the solvent</td>
</tr>
<tr>
<td>Bulk (^1\text{H} T_1) (ms)</td>
<td>340 ± 6</td>
<td>227 ± 10</td>
<td>113 ± 4</td>
<td>105 ± 5</td>
<td>460 ± 20</td>
<td>189 ± 16</td>
</tr>
</tbody>
</table>
4.3 Results and Discussion

4.3.1 Solvent Accessibility

Unlike soluble proteins, membrane embedded ASR and PR have their hydrophobic cores and much of their outer surfaces protected from solvent. Membrane protein accessibility to polar solutes may be further affected by their orientation in the lipid bilayer. To illustrate this, in Figure 4-1A we show a schematic representation of a proteoliposome with the extracellular surface of ASR facing the interior, which could be inaccessible to CuEDTA. Although the morphology of our sample is most likely very different because of the high protein to lipid ratio (approximately 1:20 molar ratio) and low water content, some protein molecules could still be inaccessible to CuEDTA. To study the accessibility of both sides of the proteins to solvent after reconstitution, an unlabeled sample of N148C ASR was reconstituted in the presence of the hydrophilic negatively charged dye 8-hydroxyl-1,3,6-pyrenetrisonate (pyranine). Pyranine is known to be membrane impermeable and is retained within well-formed liposomes for days, and has absorption bands centered at 370, 400, and 460 nm of the visible spectra [194]. The proteoliposomes were spun down and then resuspended in a small volume of fresh buffer. Visual inspection of absorption spectra, shown in Figure 4-1B, confirmed that the proteoliposomes do not retain pyranine upon buffer exchange, indicating that both sides of the protein are equally accessible to polar solutes.
Figure 4-1 A) Schematic representation of lipid reconstituted ASR at a high protein to lipid ratio and low water content, shown to illustrate that some surfaces of the protein may be inaccessible to CuEDTA. Only one of the two possible orientations of the protein in the lipids is shown. B) Absorption spectra of liposomes containing ASR reconstituted in the presence of 5 mM pyranine immediately following removal of biobeads (red) and following pelleting and resuspension in pyranine-free buffer (black) at pH =8.0 which shows that pyranine is not retained in the proteoliposomes.

4.3.2 Paramagnetic Relaxation Enhancement in N148C ASR $^{13}$C-detected experiments

To optimize sample preparation conditions and the parameters of NMR experiments, we first measured the bulk $^1$H $T_1$ of ASR, ASRCu, ASR40Cu, and ASR80Cu using an $^{15}$N-detected proton saturation recovery experiment, as shown in Figure 2-7A [195]. These samples are estimated to contain 0, 1, 4 and 8 CuEDTA molecules per protein molecule, as described in Section 4.2. Measured $^1$H $T_1$ values are summarized in Table 4-1 and discussed in the following.

We observe a relatively short $^1$H $T_1$ value of 340 ms for the diamagnetic ASR sample, which falls into the range of values previously reported for other diamagnetic proteins, including ubiquitin (230 ms) [173], GB1 (450 ms) [186], DsbB (490 ms) [189], and Aβ amyloid fibrils (560 ms) [173]. This variation in $T_1$ values is likely related to the different states of these proteins (microcrystalline, precipitated, lipid embedded), as well as the different spinning frequencies and temperatures at which experiments were performed.
In ASRCu, the addition of a single covalently attached tag leads to a moderate decrease in the bulk proton $T_1$ value from 340 ms to 227 ms. We note that a much larger bulk proton PRE effect was observed in microcrystalline GB1, with the bulk $T_1$ being reduced from 450 ms to 120 ms after the addition of a single covalently attached CuEDTA [186]. As the possibility of incomplete labeling has been ruled out through MALDI-MS, it is likely that the significantly reduced enhancement of the bulk $^1\text{H} \ T_1$ is due to the much larger size of ASR, as the surface to volume ratio is much smaller in ASR, and the interproton spin diffusion required to transfer the paramagnetic enhancement throughout the entire protein may be less efficient. Thus, it appears that a single covalently attached CuEDTA paramagnetic tag is not sufficient to achieve the desirable shortening of the proton spin-lattice relaxation times in our samples.

As we estimate that a single covalently attached CuEDTA is equivalent to ~10 mM of free CuEDTA, it is likely that an increase in CuEDTA concentration would further reduce the bulk $^1\text{H} \ T_1$. To verify this, we prepared ASR samples with 40 mM and 80 mM CuEDTA in the solvent. Not only are the concentrations of CuEDTA in ASR40Cu and ASR80Cu estimated to be approximately four and eight times higher than that in ASRCu, but the paramagnetic dopants will now have access to both sides of the protein. This allows loops on both sides of the protein to be in contact with the paramagnetic dopant, which likely doubles the effective surface area to volume ratio and reduces the distance which the proton spin polarization must travel. Accordingly, we see a large decrease in the bulk $^1\text{H} \ T_1$ to 113 ms in ASR40Cu. Further increase in the CuEDTA concentration to 80 mM results in only marginal decrease of the bulk proton $T_1$ to 105 ms, which indicates that the spin-lattice relaxation enhancement effect is approaching saturation in this range of concentrations.
In Figure 4-2 we show 2D NCA spectra measured on the four samples with different levels of paramagnetic doping. We expect average sensitivity increases of ~1.2, 1.55, and 1.6 in ASRCu, ASR40Cu, and ASR80Cu respectively when compared to ASR, based on the number of scans performed in each experiment. The resolution of all four spectra collected under low power decoupling conditions [65] is comparable to that obtainable in a 3.2-mm Bruker E-free probe [123], albeit the absolute sensitivity (i.e., sensitivity from a fully packed rotor) is lower (discussed in the following). The overall structure of the spectra remains the same with the addition of paramagnetic doping, with only a few insignificant perturbations to peak positions being observed. Slight inhomogeneous line broadening, mostly in the indirect $^{15}$N dimension, was observed in ASRCu and ASR40Cu. The broadening is uniform over the affected spectra and is not present in the ASR80Cu sample. As Cu(II) mostly enhances longitudinal relaxation rates
and has a small effect on $T_2$ [196, 197], the observed broadening is likely due to variations in the reconstitution conditions. We did not observe instances in which a peak appears in the spectrum for the diamagnetic sample but is not present in the paramagnetic samples, indicating that the addition of CuETDA, either covalently bound or in the buffer, does not perturb the structure of ASR or significantly affect sample quality.

Overall, increasing the concentration of CuEDTA results in progressively better sensitivity per unit time. To determine if this enhancement is relatively uniform throughout the proteins we performed a site-specific analysis of the signal-to-noise ratio of peaks which were resolved in all four NCA spectra, shown in Figure 4-2. We expect that if paramagnetic $T_1$ enhancement were not transferred uniformly throughout the protein through spin diffusion that we would see a pattern of reduced sensitivity enhancement in residues which are located far from the paramagnetic labels due to the fast recycling and slower magnetization recovery.

In agreement with the bulk $T_1$ estimates, moderate enhancement with an average value of 1.2±0.2 is observed in ASRCu compared to ASR, with no apparent correlation between the enhancement and residue proximity to the paramagnetic center, which in this case is attached to C148 in the E-F loop on the cytoplasmic side. For example, we see that residues A55, I56, and D57, at the extracellular flank of helix B and in the BC loop on the extracellular side, are as strong as F172 and T173, located in the middle of helix F close to the CuEDTA label. Thus, evaluation of the relative signal-to-noise ratio values obtained for resolved peaks shows that there is moderate, yet fairly uniform, enhancement across the entire protein in ASRCu, despite the paramagnetic probe having access only to the specific site in the EF loop of the protein.

With the addition of 40 mM CuEDTA we observe a significant overall increase in enhancement with an average value of 1.6±0.3. Site-specific analysis shown in Figure 4-3
suggests that the observed sensitivity enhancement is not homogeneous, and is larger in the loops (e.g., the BC loop) and in other solvent-accessible regions within α-helices, e.g., T170 [93, 123]. We also observe some transmembrane residues are only moderately enhanced (e.g., T79, T80, P81). Interestingly, helix C is well protected according to previously published hydrogen/deuterium exchange data [123, 126], suggesting that the proximity to solvent exposed fragments and possible CuEDTA binding sites appears to be important. Overall, however, despite this variation of $T_1$ in the sample, it appears that spin diffusion is not a limiting factor in ASR, and is sufficient to transfer the $^1$H spin-lattice PRE effect over the entire protein.

Due to the small decrease in the bulk $^1$H $T_1$ values between ASR40Cu and ASR80Cu and the generally low signal to noise ratios of peaks, we did not expect to observe much greater enhancement in ASR80Cu when compared to ASR40Cu. Indeed, as shown in Figure 4-3, signals from most residues have similar signal-to-noise ratios in the two spectra, and the average increase in enhancement in ASR80Cu is also 1.6±0.3. There are however, several residues, primarily in the exposed regions, which experience a larger than average sensitivity increase in ASR80Cu, which may be attributed to more even distribution of the paramagnetic enhancers on the protein surface.

It is instructive to directly compare the $^{13}$C-detected experiments performed on the 1.3-mm probe at 50 kHz MAS with those which we obtain on a fully packed rotor in a Bruker 3.2-mm E-free probe with high power decoupling at moderate spinning rates (14-15 kHz), and with a typical recycle delay of 1.7 s. Although under optimal conditions on the 1.3-mm probe (ASR40Cu or ASR80Cu) the recycle delay can be reduced by a factor of approximately 13, we still see approximately a 3x sensitivity decrease. The sensitivity per unit mass, however, is greatly increased (~5x). These results are summarized in Table 4-2. Thus, in situations when a
protein can be expressed in sufficient quantities, the 3.2-mm probe is a better choice, as it provides better overall $^{13}$C sensitivity. For mass-limited samples of proteins of 27 kDa investigated in this study, the sensitivity obtained using the 1.3-mm probe is sufficient for applications involving 2D spectroscopy, but is unlikely to be adequate for more complex 3D applications.

![Figure 4-3 Site-specific comparison of the relative S/N of ASRCu (blue), ASR40Cu (red), and ASR80Cu (green) for residues which are resolved in all four 2D NCA spectra. All S/N values are normalized to those in spectra of diamagnetic ASR considering the same sample quantities and experimental time. The secondary structure of ASR derived from solid-state NMR studies [93] is shown on top. In cases where sensitivity enhancement is not significantly different between ASR40Cu and ASR80Cu, no ASR80Cu bar is shown.](image)

**Table 4-2** Sensitivity comparison between the 3.2 mm E-free probe and the 1.3 mm probe under optimized PACC conditions

<table>
<thead>
<tr>
<th>Probe</th>
<th>$^{13}$C-detected experiments</th>
<th>$^1$H-detected experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.2-mm</td>
<td>1.3-mm</td>
</tr>
<tr>
<td>Spinning Frequency</td>
<td>14.3</td>
<td>50</td>
</tr>
<tr>
<td>Relative Volume</td>
<td>1</td>
<td>~0.067</td>
</tr>
<tr>
<td>Recycle Delay (s)</td>
<td>1.7</td>
<td>0.135</td>
</tr>
<tr>
<td>Relative Sensitivity/Scan$^1$</td>
<td>1</td>
<td>~0.11</td>
</tr>
<tr>
<td>Relative Sensitivity/Unit time$^2$</td>
<td>1</td>
<td>~0.33</td>
</tr>
<tr>
<td>Relative Sensitivity/Unit sample$^3$</td>
<td>1</td>
<td>~5</td>
</tr>
</tbody>
</table>

$^1$ relative sensitivity of a single scan with a recycle delay of at least $3^*T_1$

$^2$ relative sensitivity measured over a constant time

$^3$ sensitivity per unit sample measured over a constant time
4.3.3 Paramagnetically-Enhanced Proton-Detected Experiments

To investigate the potential of the PACC scheme for proton-detected spectroscopy of membrane proteins, we created two uniformly $^2$H,$^{13}$C,$^{15}$N-labeled samples of PR. One sample was prepared without CuEDTA as a control (DPR), and one with 80 mM CuEDTA (DPR80Cu). Both samples contained 40% H$_2$O in the solvent. The $^1$H $T_1$ of DPR was measured to be 460 ms, and with the addition of 80 mM CuEDTA this value was reduced to 189 ms. The longer $T_1$ values than those measured on UCN N148C ASR are likely due to the fact that the proton bath is severely diminished by deuteration, resulting in less efficient spin diffusion between protons, as has been observed previously in $\alpha$-spectrin SH3 domain [198].

To further investigate the effect of CuEDTA on our deuterated samples, we collected 2D $^1$H-$^{15}$N heteronuclear correlation spectra on both samples. In Figure 4-4, we compare these spectra with that obtained using a 3.2-mm TL2 Bruker probe at a spinning rate of 20.5 kHz.

Taking advantage of the relatively short proton $T_1$, faster recycling can be used even in diamagnetic samples. Remarkably, despite a ~15-fold reduction in the amount of sample packed in 1.3-mm rotor, it took only 80 minutes to collect a 2D $^1$H-$^{15}$N correlation spectrum with sensitivity similar to that from a 3.2-mm rotor (Figure 4-4A and 4-4B). We have previously established that the proton line width in PR is dominated by inhomogeneous contributions [59]. In accordance, although the bulk coherence life time, $T_2'$, is strongly dependent on the spinning rate – we determined a value of 4.4 ms at 50 kHz, compared to 2.4 ms at 20.5 kHz [59] – the proton line widths are the same, on the order of 0.15-0.25 ppm in both spectra.

The addition of 80 mM CuEDTA leads to a fairly uniform apparent reduction of $^1$H $T_1$, which allows for faster experimental recycling, enabling the collection of spectra of similar sensitivity in 35 minutes. However, CuEDTA appears to induce some spectral changes. The most
obvious one is the disappearance of several peaks in the spectrum in Figure 4-4C, namely the side chain signals of Q183 and N176, located in the E-F loop on the cytoplasmic side, which may be due to CuEDTA binding. Although to a much lesser extent, the intensity of G169, which is found in the same E-F loop, is also diminished, as can be seen in the 1D slices displayed in Figure 4-4. Apart from these differences the rest of the spectra are of similar intensity to that of the diamagnetic PR in Figure 4-4B.

Figure 4-4 $^1$H-$^{15}$N heteronuclear correlation spectra of U-$[^2$H-$^{15}$N-$^{13}$C]-PR. Diamagnetic PR at a spinning frequency of 20.5 kHz A), 50 kHz B), and C) 50 kHz with 80 mM CuEDTA added to the solvent. In all cases exchangeable sites are protonated at 40%. Dashed lines indicated the position of the slices shown above the spectra. Experimental times were 45 min, 80 min, and 35 minutes respectively. The lowest contour on each spectrum is cut at $10^{+\sigma}$. S/N values for individual peaks are noted beside 1D slices.
Proton detection is found to display an even better increase in efficiency than carbon detection. As summarized in Table 4-2, initial investigations of the sensitivity show that the experiments performed on the 1.3-mm probe have only ~2.5x less sensitivity than those collected on the 3.2-mm probe. As recycle delays during proton-detected experiments on the 3.2-mm probe are increased due to decoupling applied to the nitrogen channel during acquisition (D1≈3 s), despite the reduced enhancement of \(^1\)H \(T_1\) relaxation times, caused by the extensive deuteration of the sample, we are still able to recycle proton detected experiments 12x faster on the 1.3-mm probe. Therefore the 1.3-mm probe and the 3.2-mm probe can achieve comparable sensitivities in a given time. In proteins with higher structural homogeneity than PR, proton detection would further benefit from faster MAS rates, resulting in higher spectral resolution and sensitivity [62]. Thus, the development of proton-detected experiments for characterizing membrane proteins in the 1.3-mm probe would be beneficial, as the characterization of a protein would then require only ~1/15 of the sample required for the 3.2-mm probe, and still obtain the same sensitivity, either in protonated or per-deuterated samples in cases when such samples can be produced.

### 4.4 Summary and Conclusions

We have investigated the utility of the previously introduced paramagnetic relaxation-assisted condensed data collection (PACC) scheme for applications to relatively large 7TM proteins packed in a small 1.3-mm rotor. Between the two studied paramagnetic doping schemes, the covalent introduction of CuEDTA in one of the loops, and the addition of CuEDTA in solution, the latter provided much better paramagnetic relaxation enhancements and, in combination with low power decoupling and magic angle spinning of 50 kHz, permitted fast
recycle delays of 135 ms ($1.3T_i$) in samples containing 80 mM CuEDTA (approximately 8 Cu$^{2+}$ ions per ASR molecule). This is approximately 13 times faster compared to traditional methods used in larger 3.2-mm probes. However, due to the small sample volume, the overall sensitivity per unit time in carbon-detected experiments was still lower than that seen in the 3.2-mm probe.

In contrast, the sensitivity of proton detection in a 1.3-mm probe is comparable to that observed earlier in the TL2 Bruker 3.2-mm probe [59]. Furthermore, the 1.3-mm probe offers the benefit of faster spinning rates, which in samples with homogeneous proton line widths provides additional sensitivity enhancements [62]. This is an attractive prospect for studies of samples that can only be produced in limited quantity.
Chapter 5: Proton detection for signal enhancement in solid-state NMR experiments on mobile species in membrane proteins

This work has been published:

Direct proton detection is becoming an increasingly popular method for enhancing sensitivity in SSNMR spectroscopy. Generally, these experiments require extensive deuteration of the protein, fast magic angle spinning (MAS), or a combination of both. Here, we implement direct proton detection to selectively observe the mobile entities in fully-protonated membrane proteins at moderate MAS frequencies. We demonstrate this method on two proteins that exhibit different motional regimes. Myelin basic protein (MBP) is an intrinsically-disordered, peripherally membrane-associated protein that is highly flexible, whereas *Anaibaena* sensory rhodopsin (ASR) is composed of seven rigid transmembrane α-helices connected by mobile loop regions. In both cases, we observe narrow proton linewidths and, on average, a 10 x increase in sensitivity in 2D INEPT (insensitive nuclear enhancement of polarization transfer)-based HSQC experiments when proton detection is compared to carbon detection. We further show that our proton-detected experiments can be easily extended to three dimensions and used to build complete amino acid systems, including sidechain protons, and obtain inter-residue correlations. Additionally, we detect signals which do not correspond to amino acids, but rather to lipids, and/or carbohydrates which interact strongly with membrane proteins.

5.1 Introduction

As molecular tumbling is suppressed in solid samples, their NMR spectra are broadened by strong anisotropic interactions. Most commonly, magic angle spinning (MAS) [39, 40] in combination with high power decoupling [41] is utilized to average out these interactions and to re-establish high resolution for low gamma nuclei such as $^{13}$C or $^{15}$N. In contrast, the linewidths of protons remain prohibitively broad at moderate MAS frequencies (10-20 kHz) due to the strong inter-proton dipolar couplings.
The implementation of proton detection provides an increase in sensitivity, additional chemical shift data, and more sensitive probes of conformational and environmental changes [42, 43]. Recent advances in high magnetic field and fast MAS probe technologies [44–48] and sample preparation protocols [199–202] have enabled the use of proton detection in many systems with limited mobility. Through the use of dipolar-based correlation experiments, linewidths of ~0.2 ppm (160-180 Hz at 800 MHz field strength) could be achieved on fully-protonated samples at MAS rates of 40-60 kHz [48–53], and could be significantly improved by combining high spinning frequencies with perdeuteration of the sample and the re-introduction of protons at exchangeable sites through back-exchange with protonated buffer [48, 54–61]. Although the perdeuteration of proteins is necessary to achieve sufficiently narrow proton linewidths, this method is limited as only exchangeable protons, such as amide protons and exchangeable sidechain protons, are detectable. Thus, SSNMR experiments on fully-deuterated samples are restricted to those which specifically detect these atoms, and sidechain proton assignments are difficult to obtain.

Proton-proton dipolar interactions are averaged not only by MAS, but also by local motions. For example, resolution approaching that obtained in solution NMR samples can be observed in high resolution (HR)–MAS spectra. In these experiments semi-solid materials, such as tissue samples and swollen resins, are spun about the magic angle at low frequencies (~2-5 kHz) which, due to the high mobility of the samples, is sufficient to completely average the dipolar couplings [203–206].

Typically, structured proteins consist of rigid secondary structural elements which are linked by more flexible, and often less structured, loops and turns. Within these less structured regions, local, sub-microsecond molecular motions of sufficiently large amplitude can lead to an
averaging of the dipolar interaction. This can result in a reduction of the linewidths, but also often leads to reduced sensitivity for these residues in the dipolar-based two- and three-dimensional chemical shift correlation experiments which are used to study the well-structured regions of proteins. However, the averaging of the strong dipolar interactions facilitates through-bond polarization transfers, which are based on the comparatively weak J-interaction. Thus, through the utilization of polarization transfer methods such as insensitive nuclear enhancement of polarization transfer (INEPT) [207], the mobile regions of a protein can be selectively excited [208, 209] and excellent resolution can be obtained indirectly in the proton dimension, even without the use of extensive deuteration or fast MAS. This has been demonstrated previously on the mobile regions of integral and peripheral membrane proteins [82–85], histones [86], and amyloid fibrils [87, 88].

Here, we establish the utility of high-sensitivity direct proton detection of mobile fragments in fully-protonated proteins at moderate spinning frequencies. INEPT-based experiments were carried out on two proteins which exhibit very different motional regimes. The first of the two, myelin basic protein (MBP, 18.5-kDa splice isoform), depicted in Figure 5-1A, belongs to the class of intrinsically-disordered proteins. As a major component of the myelin sheath in the brain and spinal cord, the primary function of MBP is believed to be maintenance of the compaction of the myelin sheath through association with the cytoplasmic faces of the oligodendrocyte membrane [210–212]. In doing so, MBP peripherally interacts with lipid membranes, forming three surface-seeking amphipathic α-helices [213, 214] which are linked by highly flexible, unstructured regions [85]. MBP has also been demonstrated to associate with many other proteins, gaining partial secondary structure upon these interactions as well [215–220]. Previous SSNMR studies of MBP in a myelin-mimetic lipid environment have shown that
the unstructured or extended regions can be observed by $^{13}$C-detected INEPT-based spectroscopy with typical proton linewidths of 0.15-0.2 ppm at 600 MHz field strength [85], whereas one of the less mobile peripheral $\alpha$-helices, the immunodominant epitope comprising residues N81-S99, exhibited proton line widths on the order of 0.3-0.7 ppm at 800 MHz field strength and has been observed through dipolar-based correlation spectroscopy [221].

The second example, a seven transmembrane (TM) $\alpha$-helical protein, *Anabaena* Sensory Rhodopsin (ASR) [118], depicted in Figure 5-1B, is largely composed of rigid TM $\alpha$-helices which are joined by less structured, shorter loop regions, and has been found to form a hexagonal lattice of trimers in our samples [121, 122]. Despite the increased mobility in some of the loop regions, many of these residues can be observed in dipolar-based experiments, although frequently with lower sensitivity [125], indicating that the motions in these loops are much more restricted than those present in the unstructured regions of MBP.

![Figure 5-1 A) Schematic representation of 18.5-kDa MBP. Upon interaction with lipids, MBP forms three short amphipathic $\alpha$-helices, whereas the majority of the protein remains unstructured. The arrangement shown represents one of many possible arrangements of MBP in the compact myelin environment, as the mechanisms through which MBP-lipid interactions across apposing membrane leaflets occur are unknown. B) Schematic representation of lipid reconstituted ASR. In our SSNMR samples, ASR has been shown to form a hexagonal lattice of trimers. The majority of the protein is well-structured as an intramembrane $\alpha$-helical bundle, although the loop regions which join these $\alpha$-helices may experience significant motions.](image-url)
Through the use of INEPT-based spectroscopy, we can selectively observe many protein resonances with intrinsically narrow linewidths in both ASR and MBP. In addition to protein resonances, we observe several peaks in the ASR spectra which likely correspond to isotopically-labeled lipids or carbohydrates which originate from the *E. coli* membranes, co-purify with the protein, and are tightly bound to it. With the implementation of direct proton detection, sensitivity enhancements of up to 10-fold can be achieved for the observed regions of both MBP and ASR in two-dimensional $^1$H-$^{13}$C INEPT HSQC experiments. To further study the mobile regions of MBP and ASR, this two-dimensional proton-detected experiment was extended to three dimensions through the incorporation of proton-proton mixing. In MBP, this step facilitates the detection of sidechain protons, and the determination of inter-residue correlations. In ASR, proton chemical shifts can be added to many of the spin systems corresponding to the mobile fragments of molecules which interact strongly with the protein.

5.2 Methodology

5.2.1 Expression, Purification, and Reconstitution of Myelin Basic Protein

The MBP sample was prepared and provided by Dr. Mumdooh Ahmed using the following protocol. Uniformly $^{13}$C- and $^{15}$N-labeled wild-type (i.e., unmodified), 18.5-kDa recombinant murine myelin basic protein (UCN MBP) was expressed in *E. coli* BL21-CodonPlus(DE3)-RP cells (Stratagene, La Jolla, CA) grown in M9 minimal media using $^{15}$NH$_4$Cl and $^{13}$C$_6$-glucose as the sole nitrogen and carbon sources. The protein was purified by nickel-affinity chromatography as described previously [222, 223]. To remove minor contaminating material, an additional step of ion exchange chromatography was performed [224]. The protein was then dialysed twice against buffer (50 mM Tris-HCl, 20 mM NaCl, pH
7.4), twice against 100 mM NaCl, and four times against ddH₂O. All dialysis was done against 2 L of solvent utilizing tubing with a molecular mass cutoff of 6000–8000 Da. The protein concentration was determined by measuring the absorbance at 280 nm and the purity of the protein preparation was assayed by SDS-polyacrylamide gel electrophoresis and high-performance liquid chromatography (HPLC), as described previously [222, 223].

For solid-state NMR (SSNMR) studies, proteins were reconstituted into large unilamellar vesicles (LUVs) composed of 1:1 DMPG/DMPC (1,2-dimyristoyl-sn-glycero-3-[phospho-rac-(1-glycerol)] and 1,2-dimyristoyl-sn-glycero-3-phosphocholine, respectively) at an initial protein:lipid mass ratio of 1:2, as has been described previously [221]. The LUVs and protein were diluted to a concentration of 0.5 mg/mL in the same buffer (2 mM HEPES-NaOH, 100 mM NaCl, and 1 mM EDTA at pH 7.6) and mixed by inversion before being incubated at room temperature for 2 hours. The protein-lipid complexes were then collected by centrifugation. The final protein and lipid content of the pellet was estimated using standard bicinchoninic acid and phosphorus assays [225–228], respectively, as described previously [85].

### 5.2.2 Magic Angle Spinning Solid-State NMR Spectroscopy

All NMR experiments on MBP were performed on a Bruker Biospin Avance III spectrometer operating at 600.13 MHz using a Bruker 3.2-mm TL2 ¹H-¹³C-¹⁵N MAS probe. Approximately 4 mg of UCN MBP was center-packed in a 3.2-mm rotor. The effective temperature was kept at 30 °C in all experiments with the spinning frequency set to 10 kHz.

All NMR experiments on ASR were performed on a Bruker Biospin Avance III spectrometer operating at 800.230 MHz using a Bruker 3.2-mm TL2 ¹H-¹³C-¹⁵N MAS probe. Approximately 4 mg of UCN ASR was center-packed in a 3.2-mm rotor. For consistency with
our previous measurements on ASR [93, 121, 123, 124, 126], the effective temperature was kept at 5°C in all experiments, with the spinning frequency set to 14.3 kHz.

Sample temperatures were calibrated using an external reference of methanol [229], and were confirmed by the position of the water peak [230]. The pulse sequences used to collect data and additional experimental parameters are provided in Figure 2-6. To further investigate new peaks observed in the 1D INEPT spectrum of ASR a 2D INEPT-TOBSY [160, 161] experiment was performed and compared to a 2D $^{13}$C-$^{13}$C CPMAS-DARR [153, 154] experiment. These spectra were recorded using the pulse sequences shown in Figure 2-6A and Figure 2-4A, respectively. To investigate the sensitivity increase available from proton detection in INEPT-based experiments, we recorded $^{13}$C- and $^1$H-detected $^1$H-$^{13}$C INEPT HSQC spectra on MBP and ASR, collected with the standard pulse sequences shown in Figure 2-6B, C, respectively [167–169], with acquisition parameters set such that the two experiments take approximately the same amount of time.

For the $^{13}$C-detected $^1$H-$^{13}$C INEPT HSQC spectrum collected on MBP the acquisition lengths in $t_1$ and $t_2$ were 16.5 ms and 30 ms, respectively and the spectrum was collected with 44 scans. For the $^1$H-detected spectrum collected on MBP the acquisition lengths in $t_1$ and $t_2$ were 12 ms and 23 ms, respectively and the spectrum was collected with 16 scans. Both experiments took approximately 4 ½ hours.

For the $^{13}$C-detected $^1$H-$^{13}$C INEPT HSQC spectrum collected on ASR the acquisition lengths in $t_1$ and $t_2$ were 8 ms and 20 ms, respectively. The spectrum was collected with 48 scans. For the $^1$H-detected spectrum collected on ASR the acquisition lengths in $t_1$ and $t_2$ were 10 ms and 20 ms, respectively. The spectrum was collected with 20 scans. Both experiments took approximately 4 ½ hours. All experiments above were collected with a recycle delay of 1.7 s.
Two three-dimensional (H)CHH spectra, with proton-proton mixing times of 50 and 150 ms, were collected on both MBP and ASR using the pulse sequence shown in Figure 2-6D. On MBP, a total of 168, 96, and 4096 points were acquired in $t_1$, $t_2$, and $t_3$, leading to the acquisition lengths of 7, 8, and 23 ms, respectively. For ASR, a total of 312, 96, and 4096 points were acquired in $t_1$, $t_2$, and $t_3$, leading to the acquisition lengths of ~7, 6, and 25 ms, respectively. Each of the 3D experiments on MBP took about 1 day whereas the experiments run on ASR took about 1.5 days. All experiments were run in duplicate with proton-proton mixing times of 50 and 150 ms. Due to the lack of high-power decoupling during indirect and direct acquisition times, it is possible to use a shorter recycle delay than those which are necessary for experiments which detect carbon and use dipolar-based, continuous pulse excitation methods. Therefore, our 3D proton-detected experiments were all run with a recycle delay of 1 s.

5.2.3 Water Suppression

As these INEPT-based experiments focus on the detection of non-exchangeable aliphatic protons, water suppression can be achieved through chemical replacement of the buffer protons with deuterons. This exchange can be achieved by soaking the compacted samples in D$_2$O buffer. All D$_2$O buffers were prepared with 99.99 atom% D$_2$O (Sigma-Aldrich, Oakville, Ontario, Canada). To facilitate buffer exchange, the open rotor was soaked in ~1 mL of D$_2$O buffer. The buffer was applied directly into the rotor and then mixed several times by pulling the buffer into and out of the pipette tip. The samples were incubated overnight at 4 °C before the buffer was removed, the rotor blotted dry, and the spacer and cap were replaced. Small amounts of H$_2$O were re-introduced into the D$_2$O buffers upon the addition of concentrated solutes (prepared in H$_2$O) and during pH adjustment, and were estimated not to exceed 2-3%.
5.3 Results and Discussion

5.3.1 General Characterization of Samples by Carbon–Detected SSNMR

We begin our discussion by describing the one-dimensional SSNMR spectra of MBP and ASR. Magic angle spinning cross-polarization [151] (CPMAS) and INEPT [207, 231] experiments utilize different polarization transfer mechanisms, and thereby allow for the selective excitation of protein regions with differing dynamic characteristics. Whereas CPMAS transfers are based on dipolar couplings and selectively excite rigid regions of the protein, INEPT experiments are based on J-couplings and favor the relatively mobile regions [82, 83, 85, 159].

Figure 5-2 1D $^{13}$C A) INEPT- and B) CPMAS-based spectra of UCN MBP. TPPM decoupling of 71.4 kHz was applied during acquisition, the length of which was 25 ms and 40 ms in the CPMAS and INEPT experiments, respectively. C) 1D $^{13}$C INEPT- and D) 1D $^{13}$C CPMAS-based spectra of UCN ASR. SPINAL64 [232] decoupling of ~83 kHz was applied during acquisition, the length of which was 20 ms and 25 ms in the CPMAS and INEPT experiments, respectively. All spectra were acquired with 2048 scans and are presented without apodization.
In the presence of negatively-charged lipids, MBP forms three amphipathic α-helices which anchor the protein onto the membrane, whereas most of the protein remains unstructured and highly flexible [85, 213, 214, 221]. In line with this disposition of the protein, Figure 5-2A,B shows that INEPT excitation is much more effective than CPMAS on MBP. The linewidths of the peaks observed in the INEPT spectra are much narrower than those in the CPMAS spectra, with J-splitting often being observable in the INEPT spectrum, as shown in the inset. Both spectra were obtained at 30 °C and at a MAS frequency of 10 kHz, and the dependence of MBP spectra on temperature and spinning frequency has been previously investigated [85]. We observed a significant decrease in signal intensity at higher MAS frequencies (15-22 kHz), probably due to a reversible water:lipid phase separation and protein dehydration, as has been observed previously in MBP [85]. At MAS frequencies of 10 kHz, the molecular motions which facilitate INEPT excitation can be slowed through a decrease in temperature, leading to a decrease in INEPT efficiency and an increase in CPMAS efficiency. However, membrane-bound MBP still exhibits a very high degree of mobility at low temperatures above 0 °C, and below the freezing point of water, although the CPMAS efficiency is drastically increased, spectra become inhomogeneously broadened. This indicates that the protein takes on multiple conformations, and that CPMAS approaches are inapplicable.

In stark contrast to MBP, ASR is a rigid integral α-helical membrane protein with short, often structured, loop regions joining the α-helices. Figure 5-2C, D compare INEPT and CPMAS spectra collected at a MAS frequency of 14.3 kHz and a temperature of 5 °C, which were chosen so as to be consistent with previous data collected on ASR [93, 121, 123, 124, 126]. Consistent with the rigidity of the protein, the dipolar-based CPMAS excitation is much more efficient in exciting ASR resonances than INEPT. Whereas the CPMAS spectrum is very
crowded and only very few individual peaks can be resolved, the INEPT spectrum is significantly less populated and several very well-resolved peaks, often with J-splitting evident, are present (Figure 5-2C, inset). Although an increase in the sample temperature to 30 ºC was found to increase the efficiency of INEPT excitation by approximately 2- to 8-fold (data not shown), the structure of the INEPT spectrum is largely unchanged, and CPMAS excitation remains significantly more effective. Furthermore, we found that an increase of the MAS frequency to 20.5 kHz did not significantly improve sensitivity or carbon line widths of the ¹³C INEPT spectra, implying that the dipolar interactions are sufficiently averaged even at lower frequency of magic angle spinning of 14.3 kHz.

The few resonances observed in the ASR ¹³C INEPT spectrum may originate from amino acids in the mobile regions of the protein, as well as from other mobile species which associate with the protein. Indeed, in addition to several peaks which reside in the aliphatic and sidechain carbon regions (~15-70 ppm), there are several peaks in the 70-80 ppm range which are not present or are below detection in the CPMAS spectrum. Similar peaks have previously been reported in SSNMR spectra of ASR and PaGL in the E. coli membrane [100, 122], as well as in the NMR spectra of bacteriorhodopsin (BR) [33], and of green proteorhodopsin (GPR) [114], and have been attributed to carbohydrates that are tightly associated and co-purify with GPR or ASR, or to glycolipids of BR.

We have previously shown that ASR trimer formation occurs in the E. coli membrane [122]. It is possible that the molecules we observe could be trapped between monomers, potentially mediating and contributing to the intermonomer interactions, or be bound in the inner pocket of the trimer. The ASR trimers are stable throughout solubilization and reconstitution
[121], and any molecules which are associated with the trimer in the *E. coli* membrane can easily be retained in proteoliposomes.

The co-purification of lipids with membrane proteins is a fairly common occurrence, with the majority of observed protein-lipid contacts being due to hydrophobic interactions between the protein and lipid-acyl chains. Indeed, it is not unusual for lipid-acyl chains to be resolvable in crystallographic structures, indicating a high degree of order and rigidity in these regions. For example, the X-ray structure of ASR includes several acyl chains [29], and the acyl chains seen in the crystal structure of BR [30–32] have been well studied and linked to headgroups through neutron diffraction, with the lipid molecules being further characterized by mass spectrometry [233]. Although acyl chains are commonly visible, the remainder of the molecule most often cannot be resolved in X-ray structures, implying the presence of significant motions or disorder in these regions. It is possible that lipid molecules bind tightly to ASR through the acyl chains, which are rigid and invisible in the INEPT spectra, whereas the remainder of the molecule remains mobile, and thus is visible in our INEPT spectra. For example, while the acyl chains of the glycolipids are rigid and are visible in the X-ray structure of BR [30–32], the headgroups are mobile and are visible in solution-NMR spectra [33].

The nature of the additional resonances in our $^{13}$C INEPT spectrum of ASR has been investigated further by a two-dimensional $^{13}$C-$^{13}$C experiment in which $^{13}$C polarization is excited through INEPT, and TOBSY (total through-bond correlation spectroscopy) is utilized to establish correlations between bonded carbon atoms ([Figure 5-3A black]) [160, 161]. As both of these methods are based on through-bond transfers, this spectrum differs significantly from that excited using CPMAS and utilizing DARR (dipolar-assisted rotational resonance) for $^{13}$C-$^{13}$C mixing ([Figure 5-3A red]) [153, 154]. Whereas the INEPT-TOBSY spectrum contains only a
Figure 5-3 A) 2D $^{13}$C-$^{13}$C spectra of UCN ASR obtained with INEPT excitation and TOBSY mixing (black) and CPMAS excitation and DARR mixing (red). B) Enlarged regions of the same spectra showing an extended spin system corresponding to a carbohydrate moiety built from the TOBSY correlations available in this spectrum. Resonances are numbered C1-C5 based only on their hypothesized order in the molecule. Dashed lines represent spin system connectivities, and the intersections of these lines indicate expected cross-peak positions. As only 1-2 bond correlations are present, not every intersection contains a cross-peak. Many of the spin systems which can be built from the INEPT-TOBSY spectrum are inconsistent with the typical chemical shifts of amino acids, but rather agree with those of carbohydrates and phospholipid headgroups and backbones. Assigned spin systems are summarized in Table 5-1.

Few peaks which may correspond to amino acid spin systems, many peaks which are not present in the CPMAS-DARR spectrum are present, particularly in the 70-80 ppm and ~100 ppm ranges. As TOBSY mixing (~7 ms) provides 1-2 bond correlations under our conditions, extended spin systems can be built when multiple cross-peaks are present for a given chemical shift. For example, in Figure 5-3B the resonance labeled C5 has cross-peaks with both the C3 and C4 resonances. All three shifts can be assigned to the same spin system if there are also cross-peaks between the C3 and C4 resonances, which is the case in this instance. In Figure 5-3B we show that a total of five carbon atoms, which we have numbered based only on their hypothesized
order in the molecule, can be incorporated into a single spin system which contains multiple resonances within the 70-80 ppm region, as well as correlations with peaks at ~100 ppm (typical for C1 carbon of sugars), and at ~ 56 ppm (typical for an N-acetylated carbon). Thus, this spin system likely represents a carbohydrate moiety [234], which may be associated with rigid acyl chains that are tightly bound to ASR. As the *E. coli* membrane does not contain glycolipids, it is likely that this molecule is a lipopolysaccharide precursor. In addition to several such spin systems, we also observe spin systems which resemble the headgroups and backbones of phospholipids [235]. Phospholipid headgroups generally contain 2–3 carbon atoms which resonate in the 40-70 ppm range, and cannot always be differentiated from amino acids in this spectrum. The phospholipid backbones (most commonly glycerol, though more complicated alcohols are possible), however, generally consists of three resonances in the 55-75 ppm range [235], and several such systems can be identified in this spectrum, although these systems could also represent fragments of sugar moieties or lipid head-groups. A full list of assigned spin systems obtained from the two-dimensional $^{13}$C-$^{13}$C INEPT-TOBSY experiment is provided in Table 5-1.
<table>
<thead>
<tr>
<th>Amino Acids</th>
<th>C</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1/A</td>
<td>2/B</td>
</tr>
<tr>
<td>1</td>
<td>GLN</td>
<td>57.5</td>
</tr>
<tr>
<td>2</td>
<td>ILE</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ILE</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>VAL</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>LYS</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>GLU</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>THR</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>ASN</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>LEU</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>SER</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>HIS</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>MET</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Carbohydrates</th>
<th>C</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>68.2*</td>
<td>53.0*</td>
</tr>
<tr>
<td>4</td>
<td>101.5*</td>
<td>56.1*</td>
</tr>
<tr>
<td>5</td>
<td>100.6</td>
<td>74.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lipid headgroup/backbone</th>
<th>C</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>81.6*</td>
<td>72.9*</td>
</tr>
<tr>
<td>7</td>
<td>69.8</td>
<td>19.3</td>
</tr>
<tr>
<td>8</td>
<td>74.7</td>
<td>73.3</td>
</tr>
<tr>
<td>9</td>
<td>69.8</td>
<td>59.2</td>
</tr>
<tr>
<td>10</td>
<td>64.6</td>
<td>42.9</td>
</tr>
<tr>
<td>11</td>
<td>75.1</td>
<td>65.3</td>
</tr>
<tr>
<td>12</td>
<td>81.5</td>
<td>72.8</td>
</tr>
<tr>
<td>13</td>
<td>65.3</td>
<td>70.7</td>
</tr>
<tr>
<td>14</td>
<td>60.2</td>
<td>43.9</td>
</tr>
<tr>
<td>15</td>
<td>71.9</td>
<td>70.8</td>
</tr>
</tbody>
</table>

*Resonances which are observed in both spectra.
5.3.2 Water Suppression in Proton-Detected SSNMR Spectra

Proton detection in SSNMR is often hindered by the presence of an overpowering water signal which dwarfs all other signals and obscures useful data. In fully deuterated samples, protons are only present at exchangeable sites, necessitating the incorporation of some percentage of H$_2$O in the buffer, and the implementation of more complicated water suppression methods [50, 201, 202, 236]. In contrast, our INEPT-based experiments focus on the detection of non-exchangeable aliphatic protons, and water suppression can be achieved through chemical replacement of the buffer protons with deuterons. This exchange can be achieved by soaking the compacted samples in D$_2$O buffer. In addition to reducing the strong water signal, this process also results in the replacement of all exchangeable protons in the proteins with deuterons (H/D exchange). Thus, a relatively large volume of D$_2$O buffer (~1 mL) is necessary to ensure the exchanged protons, as well as the original H$_2$O buffer, are sufficiently diluted (at least 100 x).

Figure 5-4 1D $^1$H spectra of A) UCN MBP and B) UCN ASR in 100% H$_2$O buffer (black) and D$_2$O buffer (red). The significant decrease in the water signal (as illustrated in the insets) seen in the proteins in D$_2$O buffer facilitates proton-detected experiments. DMPC assignments are labeled according to Hong et. al. [237]. All spectra are presented without apodization.
Whereas a large portion of MBP is solvent-accessible and amenable to quick back-exchange, many ASR residues are protected from H/D exchange by the tight packing of α-helices and the membrane environment [123, 126], and studies on a wide range of proteins indicate that such protected regions may experience exchange on a slower time scale (> 12 hours) [71–75, 77]. In order to ensure complete H/D exchange, our samples were soaked in D₂O buffer for a minimum of 15 hours.

Figure 5-4 shows directly-detected ¹H spectra of both MBP and ASR before and after H/D exchange. Several notable differences between the spectra are evident. To begin, the sidechain proton regions (CH₂ and CH₃ groups) of MBP are much narrower than those of ASR, which is consistent with MBP being more mobile than ASR. There are also striking differences in the water content of the two samples. Relative to the protein content, it appears that the ASR sample contains 4-8× more residual water than MBP, which suggests that the ASR proteoliposomes retain significantly more water than those with MBP. Whereas the water peak in the ASR spectrum is reduced 30-fold after exchange with D₂O buffer, the water peak in the MBP spectrum is reduced by only ~8-fold. This difference could be due to differing proton content of the buffers used for ASR and MBP, which could easily be a result of the buffer preparation process (i.e. pH adjustment, release of protons from hydrated solutes). Alternatively, the overall buffer content of the MBP sample could be increased after buffer exchange due to less efficient removal of residual buffer. Although the water is not completely removed in either sample, the drastic reduction of this signal makes two- and three-dimensional proton-detected INEPT-based experiments possible.
5.3.3 Two-dimensional Proton-Detected SSNMR Experiments on Myelin Basic Protein

To investigate the sensitivity enhancement available from proton detection in INEPT-based experiments, we recorded $^{13}\text{C}$- and $^1\text{H}$-detected $^1\text{H}$-$^{13}\text{C}$ INEPT HSQC spectra on MBP, as shown in Figures 5-5A and B, respectively. The chemical shifts of the observed cross-peaks in both spectra agree well with assignments which have been made previously on MBP in similar lipid and buffer environments [85], and the type of many peaks can be identified. For example, the peaks located in the carbon ~41-44 ppm and proton ~2.8-3.5 ppm region belong to the C\-H\-E atoms of lysine residues, or to the C\-D\-H\-D atoms of arginines. Furthermore, the peaks seen at ~ 0.8-2 ppm in the proton dimension and ~16-17 ppm in the carbon dimension likely belong to sidechain methyl groups. Although there are many peaks which represent sidechain atoms (carbon shift of ~20-35 ppm), there are noticeably fewer resolved CA resonances (carbon shift of ~45-65 ppm), likely due to the high level of spectral overlap [85]. As the mobile regions of MBP which are excited by INEPT are unstructured [85], the chemical shifts are poorly dispersed and many overlapping signals are present, as is characteristic of intrinsically-disordered proteins [238, 239]. For example, based on previous assignments the peak located at $^{13}\text{C}$ and $^1\text{H}$ values of ~69.6 ppm and ~4.0 ppm, respectively, is composed of five overlapping threonine CB/HB peaks, which all have identical proton chemical shifts (within 0.01 ppm) and $^{13}\text{C}$ chemical shifts which vary by only ~0.2 ppm [85].

In the few non-overlapping peaks present in Figure 5-5, the average MBP linewidths range from 0.13-0.20 ppm and 0.35-0.65 ppm in the proton and carbon dimensions, respectively. Coherence life-time, $T_2'$ [182], measurements indicate values of 4-6 ms for carbons and 5-10 ms for protons. These values correspond to homogeneous proton linewidths of approximately 0.05-0.11 ppm, and carbon linewidths of 0.35-0.55 ppm, which indicates that inhomogeneous
Figure 5-5 Comparison of the aliphatic region of the $^1$H-$^{13}$C INEPT HSQC spectra of MBP and ASR. A) $^{13}$C-detected and B) $^1$H-detected HSQC spectrum of MBP. D) $^{13}$C-detected and E) $^1$H-detected HSQC spectrum of ASR. All spectra were collected with a recycle delay of 1.7 s and took approximately 4½ hours to acquire. In panels C) and F), we show 1D traces extracted from the 2D $^1$H-detected (black) and $^{13}$C-detected spectra (grey) of MBP and ASR, respectively, at the positions indicated by the dashed lines and scaled such that the noise amplitudes are equal.
broadening, most likely due to conformational heterogeneity which is not completely averaged by fast local molecular motions [85], has only a minimal effect on the linewidths.

As expected, proton detection results in a significant increase in sensitivity, with peaks in the $^1$H-detected spectrum being ~10-fold more sensitive than those in the $^{13}$C-detected spectrum, and values generally ranging from 7-12. These values are larger than the increase expected from comparing gamma ratios alone, indicating that other factors, such as the quality factors of the coil, the polarization efficiency, the length of the acquired time domain signal, and the linewidths are responsible for some portion of the enhancement [240, 241].

5.3.4 Two-dimensional Proton-Detected SSNMR Experiments on Anabaena Sensory Rhodopsin

ASR represents a more challenging case than MBP because both the transmembrane regions and some of the loops are structured and more rigid, experiencing only small amplitude motions, as is evident from the order parameter measurements [125]. To investigate the applicability of proton detection of the mobile fragments, we have compared $^{13}$C- and $^1$H-detected $^1$H-$^{13}$C INEPT HSQC spectra of ASR (Figure 5-5D, E, respectively). The overall shape of these spectra is very similar to that of MBP, and many cross-peaks are observed in regions which can tentatively be assigned to CA-HA and CB-HB correlations, as well as to the sidechains of lysines and arginines (in the carbon ~41-44 ppm and proton ~2.8-3.5 ppm regions). We also observe many peaks in the ~70-80 ppm region in the carbon dimension, which we can tentatively assign to the isotopically-labeled lipids or carbohydrates which co-purify with ASR.

Similar to what we have observed in MBP, narrow linewidths are obtained in the directly-detected $^{13}$C and $^1$H dimensions with values ranging from 0.4-0.8 ppm and 0.10-0.22...
ppm, respectively. The $T_2'$ measurements on ASR reveal values which are generally shorter than those obtained on MBP (3-5 ms in $^{13}$C and 3-7 ms in $^1$H), indicating that once again inhomogeneous broadening does not have a large effect on the linewidths. In line with our observations on MBP, the sensitivity of a given peak is, on average, ~10-fold greater in the $^1$H-detected spectrum, with values ranging from 7-12.

5.3.5 Three-dimensional Proton-Detected SSNMR Experiments on Myelin Basic Protein

The limited resolution of the two-dimensional INEPT-HSQC spectra can be improved by adding the third dimension. In Figure 2-6D, we show a pulse sequence for a three-dimensional (H)CHH experiment, which incorporates an additional proton mixing step to facilitate inter-proton polarization transfer. Active recoupling methods, such as those based on DREAM [157], and RFDR [242], were initially tested on MBP and found to be inefficient. Conversely, mixing based on the nuclear Overhauser effect (NOE) [170–172] was found to result in effective polarization transfer, with the additional benefits of ease of set-up and lack of high-power irradiation, therefore allowing for faster experimental recycling. An additional echo period is added at the end of the pulse sequence to suppress broad proton signals which arise from immobile residues as a result of the final mixing step.

The 3D (H)CHH experiment was performed on MBP with proton mixing times of 50 and 150 ms. Figure 5-6A shows the two-dimensional $^{13}$C-$^1$H projection plane of this experiment acquired with 50 ms of proton-proton mixing and with the $t_2$ evolution time set to zero (2D (H)C(H)H, spectrum shown in black). Compared to the two-dimensional $^1$H-$^{13}$C INEPT HSQC (shown in blue), additional cross-peaks, resulting from the proton-proton mixing, are seen. The majority of cross-peaks represent intra-residue correlations between neighboring sidechain
protons. For example, the CE-HE peaks of lysine and the CD-HD peaks of arginine which were previously identified in the two-dimensional $^1$H-$^1$C spectra both have cross-peaks at proton shifts of ~1.5-2.0 ppm, which is the expected chemical shift of both HG of arginine and the HD of lysine. In addition to intra-residue correlations, many cross-peaks appear between the amino acid peaks and water, indicating a close association between the amino acids and water. Similar correlations have been observed previously in MBP [85], and are consistent with these amino acids being located either outside of the membrane or in the hydrophilic region of the phospholipid headgroups.

In Figure 5-6B, C we show several two-dimensional $^1$H-$^1$H planes from the (H)CHH experiments. Amino acid systems can be built from data obtained from this spectrum when it is collected with 50 ms of proton mixing by matching proton-proton cross-peaks to obtain the chemical shift values of neighboring carbons. Specifically, cross-peaks provide $C_x[i]-H_x[i]-H_y[i]$ and $C_y[i]-H_y[i]-H_x[i]$ correlations (where X,Y=A,B,G, etc.). By matching the $H_y-H_x$ shifts of one system to the $H_x-H_y$ shifts of another, extended $C_x[i]-H_x[i]-C_y[i]-H_y[i]$ systems can be built, as is illustrated in Figure 5-6B. Often, several proton-proton cross-peaks are present for a given $C_x[i]-H_x[i]$ pair, as shown in Figure 5-6C(1), and the amino acid spin system can be extended by repeating the process described above for multiple $H_x[i]-H_y[i]$ pairs to create complete $C_x[i]-H_x[i]-C_y[i]-H_y[i]-C_z[i]-H_z[i]$... systems. The amino acid type can then be determined by comparing the chemical shift values to known values for amino acids [243, 244]. For example, the $C_x[i]-H_x[i]-C_y[i]-H_y[i]$ system displayed in Figure 5-6A can be identified as the CA-HA-CB-HB of a serine residue, and the spin system displayed in Figure 5-6B(1) can be identified as a lysine. Using this method, 15 amino acid spin systems can be identified in the three-dimensional (H)CHH experiment. The amino-acid type-specific assignments are summarized in Table 5-2.
The use of a longer proton mixing period of 150 ms results in additional intra-residue correlations for longer sidechains as well as inter-residue correlations between the amino acids which are spatially close. As can be seen in Figure 5-6C(2), such peaks, shown in red, are clearly visible in the spectrum of the lysine spin system acquired with 150 ms mixing time. Peaks such as these can be seen for many other amino acid spin systems in our spectra, and we display several examples in Figure 5-6C(3-5). The acquisition of additional intra-residue correlations strengthens and extends amino acid spin system assignments. Inter-residue correlations, which are expected to correspond mostly to sequential correlations due to the unstructured nature of the mobile regions of MBP [85], assist in the assignment process. For example, in Figure 5-6C(3) we show two overlapping glutamine systems (as evidenced by the presence of two CB peaks) which are indicated to be spatially close to the lysine shown in Figure 5-6C(4), as well as to the methyl group of either a leucine, isoleucine, or valine residue. The 18.5-kDa isoform of MBP contains only 8 glutamine residues and, of these, only Q70 is within 1-2 residues of a leucine, isoleucine, or valine residue. Therefore, one of the glutamine systems shown can be tentatively assigned to Q70, and the methyl group to L68. Although a lysine residue also exists in this region (K71), it is not obvious from the data which of the overlapping glutamines is responsible for the correlations seen in Figure 5-6C(3).

The identification of additional spin systems in MBP and the acquisition of further sequential assignments are prevented by the large, relative to the spectral dispersion, linewidths obtained in this spectrum and the high degree of spectral overlap, which is characteristic of unstructured proteins. Previously, N and CO assignment data, as well as the combination of several three-dimensional experiments, have been necessary to obtain sequential assignments of MBP [85]. The detection of protons will facilitate this process not only through gains in
sensitivity, but also through the incorporation of further chemical shift data and an additional dimension in which to resolve systems.

**Figure 5-6A)** 2D spectrum of an (H)C(H)H experiment on MBP (black) overlaid with the 2D 1H-13C INEPT HSQC (blue). Additional cross-peaks which are not present in the 2D 1H-13C INEPT HSQC appear in the 2D (H)C(H)H experiment as a result of the proton-proton mixing (50 ms) and represent intra-residue correlations, as well as correlations between amino acids and water. In panels **B** and **C**), 2D 1H-1H planes of the 3D (H)CHH experiments are shown. In panel **B**), we demonstrate how a simple amino acid system can be built by matching cross-peaks in the spectrum collected with 50 ms of proton-proton mixing. In panel **C**(1), we show that proton information extending HA and HB can be obtained from this experiment by displaying a lysine spin system at the CE 13C position. When the proton-proton mixing period is extended to 150 ms, as shown in panel **C**(2-5), additional cross-peaks (red) are observed which correspond to additional intra-residue correlations, as well as inter-residue correlations between protons of neighboring amino acids.
5.3.6 Three-dimensional Proton-Detected SSNMR Experiments on Anabaena Sensory Rhodopsin

To determine the applicability of our three-dimensional proton-detected experiments to ASR, we have also performed the 3D (H)CHH experiment on this sample with proton mixing times of 50 and 150 ms. Figure 5-7A shows the two-dimensional $^{13}$C-$^1$H projection plane of this experiment acquired with 50 ms of proton-proton mixing and with the $t_2$ evolution time set to zero (2D (H)C(H)H, spectrum shown in black). Although many peaks which can be tentatively assigned to amino acid systems are present in the two-dimensional $^1$H-$^{13}$C INEPT HSQC spectra (spectrum shown in blue), few proton-proton cross-peaks are observed in these regions in the three-dimensional (H)CHH experiment, even with long proton-proton mixing (150 ms), likely due to the overall rigidity of ASR. However, several more complex systems are observed, which originate from the resonances seen previously in the $^{13}$C 70-80 ppm range of the one-dimensional $^{13}$C INEPT spectrum, and which have been identified through a 2D $^{13}$C-$^{13}$C INEPT-TOBSY experiment as carbohydrates, as well as lipids. Spin systems which incorporate proton chemical shifts can be built for these molecules in the same manner in which amino acid systems in MBP were created. Figure 5-7B shows an example of such a spin system. In addition to the five carbon resonances acquired from the $^{13}$C-$^{13}$C experiment (Figure 5-3B), full proton chemical shifts for this spin system have been obtained. Several other such systems could be detected in the 3D (H)CHH spectrum, and a complete list is provided in Table 5-1. The proton chemical shifts provide additional information for the categorization of these molecules, and allow for the implementation of further proton-detection based experiments which could be used to obtain $^1$H-$^1$H intermolecular distance restraints, allowing for the determination of protein binding sites.
Figure 5.7 A) 2D spectrum of an (H)C(H)H experiment on ASR (black) overlaid with the 2D \(^1\)H-\(^13\)C INEPT HSQC (blue). Additional cross-peaks which are not present in the HSQC appear in the 2D (H)C(H)H experiment as a result of the proton-proton mixing (50 ms). The chemical shifts of these cross-peaks indicate that they do not correspond to amino acids, but rather to carbohydrates or lipids which interact strongly with the protein. In panel B), we show 2D \(^1\)H-\(^1\)H planes of the 3D (H)CHH experiment which display such a system, built using the same method with which amino acid systems are built. This system is also observed in the 2D \(^13\)C-\(^13\)C INEPT-TOBSY experiment shown in Figure 5.3B.

<table>
<thead>
<tr>
<th>Amino Acids</th>
<th>CA</th>
<th>CB</th>
<th>CG</th>
<th>CD</th>
<th>CE</th>
<th>HA</th>
<th>HB</th>
<th>HG</th>
<th>HD</th>
<th>HE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>52.0</td>
<td>19.8</td>
<td></td>
<td></td>
<td>4.0</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>19.1</td>
<td></td>
<td></td>
<td></td>
<td>4.2</td>
<td>1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>69.7</td>
<td>21.6</td>
<td></td>
<td></td>
<td></td>
<td>4.1</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>58.4</td>
<td>63.8</td>
<td></td>
<td></td>
<td>4.4</td>
<td>3.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>56.3</td>
<td>33.0</td>
<td>24.7</td>
<td>28.8</td>
<td>41.9</td>
<td>4.2</td>
<td>1.7</td>
<td>1.4</td>
<td>1.6</td>
<td>2.9</td>
</tr>
<tr>
<td>Lys</td>
<td>26.8</td>
<td>43.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.2</td>
<td>1.6</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>30.7</td>
<td>26.8</td>
<td>43.2</td>
<td></td>
<td>4.2</td>
<td>1.8</td>
<td>1.6</td>
<td>3.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>32.1</td>
<td>27.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.9</td>
</tr>
<tr>
<td>Gln</td>
<td>55.9</td>
<td>29.4</td>
<td>33.7</td>
<td></td>
<td>4.2</td>
<td>2.0</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>55.9</td>
<td>29.4</td>
<td>33.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>Trp/His/Cys</td>
<td>56.3</td>
<td>31.0</td>
<td></td>
<td></td>
<td></td>
<td>4.4</td>
<td>2.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trp/His/Cys</td>
<td>57.6</td>
<td>31.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Met</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16.9</td>
</tr>
<tr>
<td>Leu/Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*In the case of multiple amino acid type, possible atoms are labelled according to the first amino acid listed.
5.4 Summary and Conclusions

Proton detection is an important method for increasing the sensitivity of SSNMR experiments. We have shown here that it can be implemented even at moderate spinning frequencies on a fully-protonated sample to detect mobile entities, including non-protein binding partners, selectively in membrane proteins. These experiments are demonstrated on two proteins which generally exhibit very different motional regimes. In both cases, narrow lines in both the proton and carbon dimensions are maintained under direct proton detection. When compared to $^{13}$C-detection, we observe an $\sim 10 \times$ sensitivity increase in two-dimensional $^{13}$C-$^1$H HSQC experiments, and show that these experiments can be easily extended to three dimensions through the incorporation of a proton-proton mixing period.
Chapter 6: 

*In Situ* Structural Studies of *Anabaena* Sensory Rhodopsin in the *E. coli* Membrane

This work has been published:


The natural cellular membrane is invariably more complex than the proteoliposomes most often used for solid-state NMR (SSNMR) studies, and differences between these two environments may affect the structure and dynamics of the proteins under examination. In this work we use SSNMR and other biochemical and biophysical methods to probe the structure of a seven transmembrane (7TM) helical photoreceptor, *Anabaena* sensory rhodopsin (ASR), prepared in the *E.coli* inner membrane, and compare it to that in a bilayer formed by DMPC/DMPA lipids. We find that ASR is organized into trimers in both environments but forms 2D crystal lattices of different symmetries. It favors hexagonal packing in liposomes, but may form a square lattice in the *E.coli* membrane. To examine possible changes in structure site-specifically, we perform two and three-dimensional SSNMR experiments and analyze the differences in chemical shifts and peak intensities. Overall, this analysis reveals that the structure of ASR is largely conserved in the inner membrane of *E. coli*, with many of the important structural features of rhodopsins previously observed in ASR in proteoliposomes being preserved in the *E.coli* membrane environment. Small, site-specific perturbations in protein structure which occur as a result of the membrane environment changes indicate that the protein can subtly adapt to its environment without large structural rearrangement.

**6.1 Introduction**

Interaction with the membrane environment, e.g., lipids, other proteins, etc., is one of the important factors which often influences both the conformation and function of membrane proteins [6, 9, 11]. However, lipid membranes are generally not suitable for structural studies by the primary structural methods of X-ray crystallography and solution NMR, due to the need for 3D crystals or rapid sample tumbling, respectively. Thus, most structural studies undertaken to
date have involved purification of the protein of interest through removal from its native membrane, isolation, and subsequent crystallization, or solubilization in suitable detergents.

Recent advances in SSNMR methodologies have paved the way for detailed investigations of structural perturbations caused by non-physiological environments. For example, comparison of structures of the bitopic oligomeric influenza A M2 proton channel in lipid bilayers, detergent micelles, and 3D crystals reveals significant changes in helical packing [19, 245]. Furthermore, the trimeric oligomerization state of the polytopic α-helical Anabaena sensory rhodopsin in lipids [121] was found to be significantly different from the dimer found in 3D crystals [29]. In bacteriorhodopsin (BR), which has been well studied using a variety of methods, changes in the protein structure and oligomeric organization have also been observed between membrane mimetic environments [246–249].

While the lipid bilayer represents a good experimentally tractable model environment for structural and functional studies of membrane proteins, the natural cellular membrane is invariably more complex and includes a diverse array of lipids as well as peripheral and transmembrane proteins. The hydrophobic thickness and fluidity of the lipid bilayer, as well as interactions with other proteins present in the membrane, may affect the tilts and length of individual helices, interhelical packing, side chain conformations and protein dynamics, and through these effects influence the functional properties of proteins [6, 9, 11, 18, 19].

Several recent studies have demonstrated the feasibility of studying over-expressed peptides and proteins in the native E. coli membrane using solid-state NMR (SSNMR), often enhanced by Dynamic Nuclear Polarization (DNP) [99–104]. For example, as a result of high expression levels (70-80% of labelled protein), 12 of the 23 residues of the single-spanning transmembrane domain of LR11 could be assigned in native E. coli membranes using a 2D $^{13}$C-
$^{13}$C experiment [99]. Studies of full-length M2 in *E. coli* membranes have been used to validate the protein structure observed in liposomes [101], while in the outer membrane protein PagL small structural perturbations were observed in some of the 13 detected residues in the loop and membrane-water interface regions, particularly at the level of amino acid side chains [100].

Here, we extend these studies to polytopic α-helical proteins, and use multidimensional SSNMR to characterize the structure of a seven transmembrane helical (7TM) microbial photosensor *Anabaena* Sensory Rhodopsin (ASR) [118] in the membrane environment of *E. coli*, in which ASR is expected to partition into the inner membrane. The wealth of structural and spectroscopic data available for ASR reveals a number of structural variations which are likely related to the lipid bilayer mimetic used in structural studies. For example, spectroscopic NMR assignments, which are available for 205 of the 229 residues of ASR reconstituted in proteoliposomes [123, 124], reveal a single all-trans conformation of the retinal chromophore and of the retinal binding pocket in the dark state, yet double, structurally very similar, conformations for a number of residues located on the cytoplasmic side [124]. In contrast, the all-trans and 13-cis conformations of retinal coexist in crystals [29], which results in double conformations for some of the residues in the vicinity of retinal (e.g., K210, S86), but does not cause perturbations in the rest of the protein. We have pointed out above that the oligomeric organization of ASR is different in the lipid-reconstituted and crystalline states [93]. Furthermore, there are also differences in the structure of some of the peripheral loops, as well as looser helical packing on the cytoplasmic side of ASR in proteoliposomes [93]. The latter observation is consistent with the intermediate time scale motions of helices (tens of nanoseconds) estimated from SSNMR relaxation measurements [125], as well as with the higher solvent accessibility of the cytoplasmic side of the protein, both in the dark state [123] and under
illumination [126]. This increased conformational adaptability on the cytoplasmic side may be related to the function of ASR, in particular to its ability to interact with its transducer.

The apparent sensitivity of the ASR structure to the surrounding environment and the available structural and spectroscopic SSNMR data make this protein an interesting target for the examination of the subtle effects of membranes of different composition. Here, we combine biochemical approaches, which allow us to enhance the relative NMR signal contribution from ASR, with the biophysical techniques of circular dichroism in the visible range (CD), small angle X-ray scattering (SAXS), Fourier transform infrared spectroscopy (FTIR), and SSNMR to probe the conformation of ASR in the E.coli inner membrane (IM) environment and to compare it with the recently published data for ASR obtained in proteoliposomes [123, 124]. We show that ASR forms similarly structured trimers in both environments, which, however, are arranged in 2D crystals of different symmetry: a hexagonal bacteriorhodopsin-like lattice in the DMPC/DMPA proteoliposomes, versus a more loosely packed square 2D lattice in the E.coli IM. Site-specific analysis of the chemical shifts obtained for ~40% of ASR residues in the E.coli IM indicates a high degree of overall structural conservation. Many functionally important regions of ASR (the retinal binding pocket, the trimer interface, β-hairpin in the extracellular B-C loop) are unperturbed and represent very structurally stable elements. Outside of the conserved regions, small yet noticeable structural changes corresponding to chemical shift perturbations of up to 0.5 ppm for carbon and 1.0 ppm for nitrogen are observed in both transmembrane helices and the peripheral parts of the protein. Analysis of cross peak intensities provides an approximate measure of the proteins mobility and indicates stabilization of the peripheral helix E’ on the cytoplasmic side, which may occur because of better matching between the hydrophobic
thicknesses of ASR and the \textit{E.coli} lipids compared to DMPC/DMPA, or due to the different lattice packing of trimers.

\subsection*{6.2 Methodology}

\subsubsection*{6.2.1 Sample Preparation}

In order to facilitate the study of ASR in the \textit{E. coli} membrane, it is desirable to remove as much background signal originating from areas of the cell which do not contain ASR, as possible. This is achieved both through physically separating segments of the membrane containing ASR, and through isotopic labelling strategies which strategically limit isotopic incorporation into background proteins. In \textbf{Figure 6-1} the steps taken in order to achieve this are summarized. We denote samples of ASR which are isolated in the \textit{E. coli} membrane using these methods as EM-ASR.

\textit{Expression of Anabaena Sensory Rhodopsin with reduced isotopic labelling of background proteins}

To reduce the background NMR signal from non-ASR cell components, isotopes can be strategically incorporated preferentially into the recombinant protein and co-expressed molecular components. This was achieved through allowing cells to grow to the exponential phase in unlabeled media (\textbf{Figure 6-1}), before resuspending them in a lower volume of isotopically enriched media for protein expression [250]. In this way, the majority of cellular components are produced in unlabeled media, and therefore do not contain isotopic labels, while the recombinant protein is produced in isotopically labelled media, and therefore contains a higher proportion of isotope labels.
Figure 6-1 Overview of isotopic labelling and sample preparation strategies. To reduce NMR signal from the background proteins, cells are grown to the exponential phase in NA media, and then resuspended at a high concentration in labeled media before protein expression is induced (denoted as rbUN or rbUCN labeling). The inner and outer membranes are then separated through a sucrose gradient. IM vesicles not containing ASR are then separated using a 2-phase system which partitions ASR-containing vesicles to the heavy phase through resin binding.

In order to achieve this in our EM-ASR samples, cells were first grown to an $A_{600}$ of ~0.7 in unlabeled LB media before being collected through gentle centrifugation [250]. The cell pellet was then washed with minimal media salts before being resuspended in UN or UCN M9 minimal media at 4x their original concentration, resulting in a starting $A_{600}$ of ~2.8. After ~1 hr of growth, expression was induced with IPTG at a concentration of 1 mM and retinal was added at a concentration of 7.5 μM. At ~4 h after induction, the cells were collected by centrifugation. Samples created from cells grown using this labelling method are denoted by rbUN and rbUCN.
Isolation of Anabaena Sensory Rhodopsin in the E. coli membrane

In order to physically remove background proteins and other material from the sample while maintaining the inner membrane environment around ASR several steps were taken. To begin, cell pellets were pre-treated with lysozyme (12 mg/L of culture) and DNase I (600 units/L culture) before being broken by sonication. Low speed centrifugation was used to remove unbroken cells, and the supernatant was pelleted to remove soluble debris. The membrane pellet was then resuspended and the inner and outer membranes were separated by sucrose gradient [251]. The raw membranes were first applied on top of a two-step sucrose gradient composed of 55 % sucrose, and 9 % sucrose (all sucrose percentages are reported as w/w). These gradients were centrifuged for 2.5 h at 210,000×g in a swinging bucket rotor and the membrane fraction was collected from the top of the 55% sucrose layer. The membrane fraction was then diluted by at least 3x and subjected to a six-step sucrose gradient which was composed of discrete (from top to bottom) 30 %, 35 %, 40 %, 45 %, 50 %, and 55 % sucrose fractions. The gradients were centrifuged for 15 h at 210,000×g and the IM fraction was collected from the top of the 40% sucrose layer. Both sucrose gradients were prepared in buffer containing 50 mM TEA, 1 mM EDTA, and 1 mM DTT at pH 7.5.

The IM fraction was then resuspended in buffer containing 5 mM Tris (pH 8) before being added to a two-phase affinity system [252, 253] (10 mM Tris (pH 8), 0.3 M NaCl, 6.45% (w/w) PEG 3350, 8.45% (w/w) dextran, and 320 μl of Ni–NTA–agarose slurry per ml of system) which utilizes the C-terminal 6xHis-tag on recombinant ASR to bind ASR-containing membrane fragments. After incubation with gentle mixing in the 2-phase system for ~30 min, the phases were separated by low-speed centrifugation (1,600×g) for 5 min. The top phase, which contains IM vesicles not containing ASR, was then collected with a Pasteur pipette and replaced with an
equal volume of buffer (6.45% (w/w) PEG 3350, 10 mM Tris and 40 mM imidazole). The previous two steps were repeated until the absorption spectra showed a lack of protein in the top phase. Membranes were then eluted from the affinity beads in the bottom phase by incubation in 500 mM imidazole for 15 min, before low-speed centrifugation (1,000×g) was used to sediment the affinity beads.

As we are likely to still see a strong presence of background proteins in our SSNMR spectrum, a control sample was produced in which all steps of protein expression and IM-isolation were identical to EM-ASR preparations, including rbUCN labelling, but without the induction of ASR expression. As a result of this method, only the non-ASR proteins seen in rbUN or rbUCN labelled samples will be seen in the samples created from non-induced cells (referred to as EM-NIC in the following), and a direct comparison of the spectra obtained on EM-NIC with the spectra obtained on EM-ASR samples can be used to confirm the ASR peaks in the 2D and 3D NMR spectra.

In Figure 6-2 we briefly summarize the samples used in this study. ASR samples grown entirely in natural abundance, [U-15N], and [U-13C,15N]-labeled media are referred to as NA, UN, and UCN, respectively. Samples of ASR used in this study were: (i) isotopically labeled, solubilized, purified ASR reconstituted into liposomes, referred to as proteoliposomes or PL-ASR in the following; (ii) ASR isolated in the E. coli membrane (EM), denoted EM-ASR; (iii) samples of the E. coli membrane from non-induced cells subjected to isolation procedures identical to EM-ASR and used as a control for NMR assignments, referred to as EM-NIC; (iv) samples with reduced background signal (rb) in which cells are grown to the exponential phase in natural abundance media and then resuspended at a high concentration in UN or UCN media before protein expression is induced, denoted as rbUN and rbUCN, respectively.
Figure 6-2 Summary of isotopic labelling, sample preparation strategies, and sample compositions. NA, UN, and UCN denote natural abundance, uniformly $^{15}\text{N}$-labeled and uniformly $^{13}\text{C},^{15}\text{N}$-labeled media or samples, respectively. A sample of purified ASR reconstituted in DMPC/DMPA lipids (PL-ASR, leftmost column) was used as reference. Two types of ASR samples expressed in the E.coli membrane (EM-ASR) were prepared. For the first type, ASR and background proteins were labeled at the same level, and outer membrane and inner membrane fractions not containing ASR were minimized using a sucrose gradient and His-tag binding (second left column). To reduce NMR signal from the background proteins, cells were grown to the exponential phase in NA media, and then resuspended at a high concentration in labeled media before protein expression was induced (denoted as rbUN or rbUCN, second right column). An additional control sample was produced by following the expression and isolation procedures for the rbUCN EM-ASR sample, but without induction (EM-NIC, rightmost column). Sample compositions are given at the bottom of the figure and are explained in the text.

### 6.2.2 Magic Angle Spinning Solid-State NMR

All experiments on PL-ASR and EM-NIC, as well as initial one and two-dimensional NMR experiments on EM-ASR were performed on a Bruker Biospin Avance III spectrometer operating at 800.230 MHz using a 3.2-mm $^1\text{H},^\text{13}\text{C},^\text{15}\text{N}$ E-FREE magic angle spinning probe. All pulse sequences and their descriptions are given in Figure 2-4. Samples were packed into 3.2-mm thin-wall rotors (Bruker Biospin). All experiments at 800 MHz were performed at a spinning rate of 14.3 kHz, and at an effective temperature of 5 °C, with a recycle delay of 1.7 s. Typical π/2 pulses were 2.5 μs for $^1\text{H}$, 4 μs for $^\text{13}\text{C}$, and 7 μs for $^\text{15}\text{N}$. 
All 3D experiments on EM-ASR were performed on a 900 MHz Bruker Avance III NMR spectrometer equipped with a low-E triple resonance $^1$H-$^{13}$C-$^{15}$N 3.2 mm magic angle spinning probe designed and built at the National High Magnetic Field Laboratory, Tallahassee, FL [181]. All experiments at 900 MHz were performed with spinning at 15 kHz and at an effective temperature of 5 °C with recycle delays of 1 or 1.2 s. The high sensitivity and fast recycle delays afforded by the 900 MHz probehead were critical for the detection of small amounts of ASR. Typical π/2 pulses were 2.5 μs for $^1$H, 3.5 μs for $^{13}$C, and 5.6 μs for $^{15}$N.

For both spectrometers, the $^1$H/X (where X is $^{15}$N or $^{13}$C) cross-polarization (CP) [151] contact times were 2 ms, with a constant RF field of 35 and 50 kHz on nitrogen and carbon, respectively, while the proton lock field was ramped linearly around the n = 1 Hartmann/Hahn condition [149]. $^{15}$N/$^{13}$CA and $^{15}$N/$^{13}$CO band-selective transfers [166] were implemented with a contact time of 6 ms. For the $^{13}$CA/$^{15}$N CP, a constant lock field of $2.5\times v_r$ ($v_r = \omega_r/2\pi$, spinning frequency) strength was applied on $^{15}$N, while the $^{13}$C field was ramped linearly (10-12% ramp) around $1.5\times v_r$. For the $^{13}$CO/$^{15}$N transfer, a constant lock field of $3.5\times v_r$ field strength was applied on $^{13}$C, while the $^{15}$N field was ramped linearly (10% ramp) around $2.5\times v_r$. DREAM [157] recoupling with tangential sweep around HORROR [156] recoupling condition was used to accomplish band-selective carbon-carbon polarization transfer in the NCACB experiment. CW proton decoupling at 100 kHz was used during $^{15}$N/$^{13}$C CP and DREAM [157]. SPINAL64 [232] decoupling optimized around 83 kHz was used during $^{15}$N and $^{13}$C direct and indirect chemical shift evolutions.
6.2.3 Fourier Transform Infrared Spectroscopy

Fourier transform infrared spectroscopy (FTIR) was used to assess the protein to lipid ratio as well as the relative amounts of α-helical and β-barrel proteins in membrane protein samples. Rapid-scan FTIR was performed by using a Bruker IFS66vs machine. Dehydrated films of proteoliposomes were compressed between two CaF$_2$ windows separated with a Teflon™ spacer, and data acquisition was controlled by OPUS software (Bruker).

6.2.4 Circular Dichroism Spectroscopy

Circular Dichroism (CD) in the visible range can be used to determine the oligomeric organization of retinal-containing proteins. CD spectra were recorded on a Jasco J-810 spectropolarimeter (Jasco Inc., Easton, MD). The samples were scanned at room temperature over the 350-700 nm range at a rate of 100 nm/min. Spectra taken on buffers were used for baseline correction. UV-visible absorption spectra were collected simultaneously with the CD spectra. To collect the CD spectra of ASR in the proteoliposome and IM environment, the samples were suspended in 70% glycerol (v/v) to reduce light scattering via refraction index matching [254]. The spectra were collected using a 0.1 cm path length quartz microcell.

6.2.5 Small Angle X-Ray Scattering

Synchrotron SAXS spectra, measured at the Stanford Synchrotron Radiation Lightsource (SSRL), were obtained as has been described previously [255]. Briefly, the membranes were sealed in quartz capillaries (diameter ~1.5 mm, Hilgenberg GmbH, Germany) and irradiated for ~5 s with incident synchrotron X-rays from an eight-pole Wiggler, which were monochromatized ($\lambda=1.37776$ Å) and focused using a cylindrical mirror. The scattered radiation
was collected using a Rayonix MX225-HE detector, with each sample being measured six times. Final spectra are averages of these measurements. No radiation damage was observed for all measurements. The 2D SAXS powder patterns were integrated using FIT2D (www.esrf.eu/computing/scientific/FIT2D/), and the sample-to-detector distance was calibrated using silver behenate as a standard. Data were fitted by the nonlinear least-squares method using Igor Pro 6 (WaveMetrics, Lake Oswego, OR, USA).

6.3 Results and Discussion

6.3.1 Sample Optimization and the State of ASR in E. coli Membranes

In Figure 6-3 we show SDS/PAGE analysis of ASR samples in the E. coli membrane and proteoliposome environment. We show gels stained with Sypro Ruby Protein stain, which displays all proteins in the sample, as well as Invision His6-tag stain, which selectively stains proteins, like recombinant ASR, which contain a His6-tag. SDS/PAGE analysis of the cell pellet indicates that, despite over-expression, the ASR content of whole cells is low, with only a small amount of ASR being detected after the cells have been induced. The gels show the strong presence of a ~37 kDa protein, which likely corresponds to the major outer membrane proteins OmpA and OmpF [256–258], along with many other background proteins. This indicates that ASR would not be the dominant component in the sample and therefore additional measures are necessary in order to facilitate studies of ASR in the E. coli inner membrane.

To this end, it is desirable to remove as much background signal originating from areas of the cell which do not contain ASR, as possible. Our goal was to create a sample of ASR in which the E. coli membrane environment is maintained, while fewer background proteins and other molecules from the E. coli cells are present. This is achieved both through physically separating
segments of the membrane containing ASR (through sucrose gradient and two phase system), and through isotopic labelling strategies which strategically limit isotopic incorporation into background proteins, as described above.

Although the gels show a significant decrease in the amount of background proteins after the sucrose gradient and two-phase system, there is still a strong presence of the ~37 kDa protein. To confirm the identity of this protein, the gel band from the EM-ASR lane was subjected to cysteine reduction and carbamidomethylation, followed by trypsin digestion. The extracted peptides were analyzed by an Agilent UHD 6530 Q-TOF Mass spectrometer. Using Peaks 7 software (Bioinformatics Solutions Inc.) the predominant proteins in this band were identified as OmpF and OmpA, with 27 peptides which matched OmpF from *E. coli* and covered 58% of the OmpF sequence and 20 peptides which matched OmpA from *E. coli* and covered 53% of the OmpA sequence, being identified. Therefore the band at ~37 kDa is most likely due to OmpF and OmpA, which likely come from the residual outer membrane fraction.

![Figure 6-3 A) Sypro Ruby Protein and B) Invision His6-tag stained SDS/PAGE analysis of *E. coli* membranes from BL21-Codon-plus-RIL cells containing the plasmid for ASR which have been induced (ASR) and not induced (NIC). Samples were taken pre-induction (PI) as well as before (cell pellet) and after EM purification. Membranes were boiled in SDS before electrophoresis. The positions of PL-ASR and the high abundance 37 kDa proteins are indicated with white and black dashed boxes respectively.](image-url)
As the gel is crowded and does not give information on the isotopic labelling of proteins, we use a combination of FTIR and 1D $^{15}$N NMR to further analyze the compositions of our samples. To more accurately monitor the inner membrane isolation process and the protein to lipid ratio of samples, FTIR was performed after each step (Figure 6-4A). The Amide I band is sensitive to protein secondary structure and can be used to monitor the relative amounts of α-helical, β-strand, and random coil secondary structures in the samples. Through comparing the integrated area of the Amide I band to the lipid-ester band in FTIR spectra of samples we can determine the protein to lipid ratio. Narrowing of the Amide I band around the frequencies corresponding to α-helical structure indicates that the relative amount of β-barrel proteins is indeed being reduced during the purification. However, the greater width of the final EM-ASR peak as compared to the PL-ASR peak points to the presence of other proteins in the final sample.

To evaluate the effectiveness of rbUN and rbUCN labelling, we compare UN EM-ASR and rbUN EM-ASR samples. The 1D $^{15}$N spectrum of UN EM-ASR (Figure 6-4B) has an overall reduced intensity (~40%) compared to that of PL-ASR, likely due to the increased lipid and water content of EM-ASR. In addition, it shows a significantly shifted and broadened maximum as compared to PL-ASR as well as a secondary maximum peak which aligns with the maximum of the $^{15}$N spectrum of PL-ASR. The first maximum peak is up-shifted by ~ 1 ppm relative to the PL-ASR maximum, and the bulk centroid of the spectrum is up-shifted by much more, indicating the presence of isotopically labelled proteins with high β-strand content [243]. Resolved residues seen in the 1D $^{15}$N spectrum of PL-ASR (His8, His69, Schiff base (SB)) cannot be seen in UN EM-ASR, even after long acquisition, indicating a small relative amount of isotopically labelled ASR in the sample.
Figure 6-4  A) FTIR spectra of NA EM-ASR throughout the inner membrane isolation process. By observing the narrowing of the Amide I peak and the reduction of its β-shoulder, the reduction of outer membrane content (β-barrel proteins) can be monitored. B) One-dimensional 15N spectra of i) UCN PL-ASR, ii) UN EM-ASR, and iii) rbUN EM-ASR. The relative increase in UN-labeled α-helical proteins in rbUN EM-ASR is evident from the relative increase in the down-shifted maximum of the spectra (dashed line). All spectra are scaled to approximately match intensities, with scaling factors, corrected for the number of scans taken in each experiment (noted to the left). The main peak is presented with no window function, whereas 60 Hz of line broadening is applied to the regions of the three resolved peaks. C) One-dimensional 15N spectra of UCN PL-ASR and rbUN EM-ASR with the entire spectra processed with the stronger exponential window function of 60 Hz, and scaled such that the three resolved peaks (SB, His8, and His69) approximately match in intensity, showing that ASR accounts for only ~1/4 of the labeled protein content in EM-ASR.

The spectrum of EM-ASR with a reduced amount of isotopic labelling of background proteins (rbUN EM-ASR) shows a further decrease (55% when compared to the UN labelled sample) in overall signal intensity due to the increased presence of unlabelled protein. However, a noticeable relative increase in the ASR backbone peak, when compared to UN EM-ASR, is evident and some unique ASR residues can be resolved. In particular, the peaks at 177.2 ppm, 169.1 ppm, and 162.4 ppm correspond to the SB nitrogen of Lys210, and His8 and His69 Nδ1 resonances, respectively. All three peaks appear to be unshifted in the EM-ASR sample,
indicating that the retinal remains in an all-trans conformation and the conformations and protonation states of the His8 and His69 sidechains are unaltered in the E. coli IM.

As the UN and the rbUN EM-ASR samples were isolated in the same way, the protein content of these cells is identical, with only the isotopic labelling of the samples being different. As we show in Figure 6-4B, we see a reduction in signal in the rbUN EM-ASR sample of about 55% when compared to the UN EM-ASR sample. We can therefore estimate that ~45% of the total protein in our rbUN EM-ASR sample is isotopically labelled. Through comparing this spectrum to the PL-ASR spectrum and normalizing for the intensity of selected resolved peaks (Figure 6-4C) we find that ASR comprises ~1/4 of the total labelled protein content. Therefore, we estimate the relative amount of ASR in the EM-ASR samples to be ~10%, which leads to approximately ~ 1 mg (800 MHz, Bruker 3.2-mm thin wall rotor) and 1.6 mg (900 MHz, Revolution NMR thin wall rotor) of ASR in our SSNMR rotors.

The relative content of UCN labelled α-helical proteins in the rbUCN EM-ASR and EM-NIC samples was further analyzed using FTIR through observation of the Amide I band, which is dominated by CO stretching and is therefore sensitive to $^{13}$C labelling, and the Amide II band, which is sensitive to both $^{13}$C and $^{15}$N labelling. Isotopic labelling causes a downshift of the vibrational frequency of the amino acids containing the heavier atoms by ~20-50 cm$^{-1}$, and therefore the peaks representing the $^{13}$C and/or $^{15}$N-labelled proteins are downshifted from the frequencies observed in unlabeled proteins. As seen in Figure 6-5A, even in the cell pellet spectrum the Amide II band is clearly split into two peaks representing isotopically labelled and natural abundance proteins. The Amide II peak representing isotopically labelled proteins shows a relative increase in the EM-ASR spectrum, indicating the removal of unlabeled proteins. Amide I peaks representing the NA α-helical and β-sheet content are present between ~1658-
1647 cm\(^{-1}\) and \(\sim 1638-1632 \text{ cm}^{-1}\) respectively, with their isotopically labelled counterparts being significantly downshifted by \(\sim 40 \text{ cm}^{-1}\) from these wavelengths. Although all 4 peaks have similar intensities in the cell pellet spectrum, the increase in the relative content of isotopically labelled \(\alpha\)-helical proteins is evident in the EM-ASR spectrum.

The \(^{13}\text{C}\) spectrum of EM-ASR (Figure 6-5C) displays an amount of fine resolution in the CA region similar to that of PL-ASR (Figure 6-5B). Being heterogeneous, EM-ASR samples show a number of additional resonances that are not observed in the PL-ASR sample. Notably, three intense peaks located at 132.0, 34.7, and 32.9 ppm likely originate from the fatty acyl chains of lipids in the sample. Additionally, four well-resolved peaks can be observed in the spectral region of 70–80 ppm (Figure 6-5C insert) which are not seen in PL-ASR but have been reported before in solution NMR spectra of BR [33] and in SSNMR spectra of green proteorhodopsin (GPR) [114]. These signals have been attributed to carbohydrates which are tightly associated and co-purify with GPR or to glycolipids of BR.

Figure 6-5 A) FTIR spectra of rbUCN ASR before (cell pellet) and after (EM-ASR) the IM isolation process. Peaks corresponding to the natural abundance (NA) \(\alpha\)-helical and \(\beta\)-strand (\(\alpha\text{NA}\) and \(\beta\text{NA}\)) peaks as well as their isotopically labelled counterparts (\(\alpha\text{UCN}\) and \(\beta\text{UCN}\)) are indicated with dashed lines. A noticeable increase in the UCN \(\alpha\)-helical peak in EM-ASR as compared to the cell pellet is observed. 1D \(^{13}\text{C}\) NMR spectra of B) UCN PL-ASR and C) rbUCN EM-ASR. All NMR spectra are scaled to approximately match intensities.
We used absorption spectroscopy, CD spectroscopy in the visible range, and SAXS to further evaluate the functionality, oligomeric state and possible 2D crystallinity of ASR in the \textit{E. coli} membrane. As can be seen in Figure 6-6B, although a strong background is present, the absorption spectrum of EM-ASR can be observed with a maximum at \sim 540 nm, the wavelength at which ASR absorbs in the PL-ASR sample. Such an absorption peak is not observed in the EM-NIC sample (Figure 6-7B), confirming this spectral range to be ASR-specific. This also indicates that ASR remains well-folded in EM-ASR. The appearance of a bilobed shape in the visible CD spectra of oligomers formed by microbial rhodopsins is related to the excitonic coupling of retinal chromophores of the monomers, with the exact shape of the spectra depending on the relative orientation of retinals, as well as additional contributions from aromatic sidechains \cite{175–177}. CD spectra of ASR solubilized in DDM and in proteoliposomes show a bilobed shape which corresponds to a trimeric arrangement of ASR monomers, with the intermonomer interface being formed between helix B of one monomer, and helices E and D of another \cite{121}. In Figure 6-6A we compare the CD spectra measured in both PL-ASR, as published previously \cite{121}, and EM-ASR. Although there is a strong membrane background signal at shorter wavelengths in EM-ASR, similar bilobed shapes are clearly visible at higher wavelengths in both samples, pointing to the presence of ASR trimers in \textit{E. coli} membranes. The conservation of the trimeric structure is further supported by the analysis of SSNMR chemical shifts as discussed below.
Figure 6-6 A) Visible range CD and B) absorption spectra of PL-ASR and EM-ASR. In A), the bilobe shape is clearly present in both PL-ASR and EM-ASR, indicating the presence of trimers. Small angle X-ray scattering spectra of C) PL-ASR and D) EM-ASR. Bragg diffraction peaks in the SAXS spectrum of PL-ASR indicate the presence of an ordered 2D hexagonal lattice, while the peaks present in the EM-ASR spectrum indicate the presence of a 2D tetragonal lattice. See text for details.

Figure 6-7 A) Small angle x-ray scattering of green proteorhodopsin (PR) in the E. coli membrane. This sample was prepared by applying the protocol used to create the EM-ASR and EM-NIC samples to cells grown in the same E. coli strain as our ASR samples, with the expression optimized for PR. The lack of distinct scattering peaks indicates that peaks seen in the EM-ASR spectra are due to the presence of ASR in the sample. B) Absorption spectra of EM-NIC. Though the same scattering and absorption backgrounds as observed in the EM-ASR sample are present, there is no discernible peak at ~540 nm, indicating that the absorption at this wavelength in EM-ASR is due to the presence of ASR.
Given the high protein:lipid ratio in our SSNMR samples, it is possible that ASR may form a two-dimensional lattice in PL-ASR and/or EM-ASR. Under native conditions, BR trimers assemble into a two-dimensional hexagonal lattice called the purple membrane [259]. The synchrotron SAXS spectrum of PL-ASR shows a series of scattering peaks (marked by black arrows in Figure 6-6C). These peaks can be fit nicely to reveal a liposome form factor background (grey trace) [260, 261] and a series of scattering peaks (blue dotted traces) centered at 0.109, 0.190, and 0.218 Å⁻¹, respectively. These correlations can be indexed as the q₁₀, q₁₁, and q₂₀ scatterings of a 2D hexagonal lattice with a lattice parameter of ~66.4 Å, reminiscent of the purple membrane structure of BR, which has a lattice parameter of 62.7 Å [107]. The width of these scattering peaks suggests a relatively small crystalline domain size, i.e., ~500 Å as estimated using Sherrer’s equation [262].

In contrast, the synchrotron SAXS spectrum of EM-ASR samples, shown in Figure 6-6D, reveals a completely different 2D lattice arrangement. There are two relatively strong scattering peaks centered at 0.085 and 0.120 Å⁻¹ and a weak, albeit discernible, peak at 0.240 Å⁻¹ (marked by solid black arrows), with positions relationship of 1: √2: √8. Using a 2D tetragonally packed lattice model, the scattering data can be fit to reveal two additional hidden peaks (marked by dotted arrows) centered at 0.169 and 0.191 Å⁻¹, respectively. These five peaks are indexed as the q₁₀, q₁₁, q₂₀, q₂₁, and q₂₂ of an in-membrane tetragonal membrane protein lattice with a lattice parameter of ~74 Å. As these peaks are not visible in SAXS spectra which we have obtained on another retinal-binding protein, green proteorhodopsin, produced in a similar membrane environment (Figure 6-7A), we can conclude that they do not originate from background proteins and indicate that ASR is arranged in a tetragonal lattice in the E. coli membrane environment. The
observed scatterings are too weak to extract meaningful estimates of the domain size; however, their weak and diffusive nature suggests a smaller domain size than that observed in PL-ASR.

6.3.2 2D and 3D SSNMR spectra of ASR in cellular membranes

Chemical shifts are exquisitely sensitive to alterations in the chemical and structural environment, and can be used to detect variations in the structure of ASR in the *E. coli* inner membrane. Figure 6-8 compares the 2D NCA spectra measured in PL-ASR, rbUCN EM-ASR, and rbUCN EM-NIC. The 2D NCA spectrum of PL-ASR is generally well resolved (Figure 6-8A), with linewidths of ~0.5 ppm in the direct $^{13}$C and ~1 ppm in the indirect (truncated) $^{15}$N dimensions. The linewidths of resolved ASR peaks in EM-ASR are comparable to those seen in PL-ASR, overall indicating that the protein remains structurally homogeneous even in the heterogeneous environment of the cell membrane. We note that the cross peaks which correspond to other background proteins are similarly narrow and are downshifted in the $^{13}$C dimension and upshifted in the $^{15}$N dimension, confirming the presence of a large amount of proteins with β-strand secondary structure. This suggests that much of the lack of resolution in the center of the spectra is caused by the overlap of many well-resolved peaks, as opposed to the presence of broad, unspecific peaks.

Overall, a total of 16 ASR peaks could be identified in the 2D NCA EM-ASR spectra (Figure 6-8B) using the known resonance assignments obtained from proteoliposome samples, and through a comparison with the spectra collected on EM-NIC (Figure 6-8C). Due to the reduced amount of ASR in the EM-ASR sample, many of the lower intensity resolved peaks in the PL-ASR spectrum which do not overlap with background peaks, most notably in the proline region, are not visible in the EM-ASR spectra. Isolated residues with unchanged (conserved)
chemical shifts are found in both the loop regions (A-B, B-C, C-D, E-F, and F-G) and in the intramembrane regions of helices B, C, D, and F. In agreement with the previous observation of the unchanged chemical shift of the SB $^{15}$N, signals of the residues S47 and T79, which are part of the retinal binding pocket, are also not shifted. Additionally, the chemical shifts of M52 and I56, which are found on the extracellular side of helix B and are involved in the intermonomer contacts of the trimer, are nearly identical to those detected in PL-ASR, indicating that the monomer-monomer interface of the trimer is preserved in the *E.coli* inner membrane.

Due to the relatively large size of ASR, as well as the spectral overlap created by background proteins, 3D spectroscopy is required for more detailed site-specific analysis. In order to extend our assignments, we have performed 3D CANCO and NCACB experiments on PL-ASR, rbUCN EM-ASR, and EM-NIC samples. CANCO experiments were performed to obtain backbone assignments, whereas 3D NCACB experiments provide information on side chain atoms and help validate the identification of shifted residues from the CANCO data in cases of overlapping peaks or where there are multiple assignments possible.

In **Figure 6-8D-H** we show representative 2D planes extracted from the 3D NCACB and CANCO experiments recorded on the PL-ASR, rbUCN EM-ASR, and rbUCN EM-NIC samples. The 2D plane of PL-ASR shown in **Figure 6-8D** contains five cross peaks corresponding to valine correlations. Four of them, although reduced in intensity due to the lower amount of ASR, are also clearly visible in the spectrum of EM-ASR, while V211 is below detection in this spectrum. We additionally show several conserved peaks of interest, such as that of D75, which is situated in the retinal binding pocket (**Figure 6-8F**), and M52, which is involved in the monomer-monomer interface (**Figure 6-8G**), as well as examples of peaks which are shifted in EM-ASR (F213, V169). Additional planes are shown in **Figure 6-9**.
Figure 6-8 NCA SSNMR spectra of A) PL-ASR, B) rbUCN EM-ASR, and C) rbUCN EM-NIC. All peaks are labelled according to their assignments in PL-ASR. All spectra were recorded on a 800 MHz spectrometer and processed identically using 15 Hz of Lorentzian line narrowing and 30 Hz of Gaussian line broadening in the indirect dimension, and 30 Hz of Lorentzian line narrowing and 60 Hz of Gaussian line broadening in the direct dimension. The spectrum in A) was collected with 32 scans, while spectra B) and C) were both collected with 320 scans. In A) the lowest contour is cut at 10xσ, while in B) and C) the lowest contour level is cut at 5xσ. Overlaid representative 2D planes of D-G) 3D NCACB and H) CANCO experiments performed on PL-ASR (red), EM-ASR (black), and EM-NIC (blue). All peaks are labelled according to the assignments. PL-ASR and EM-NIC spectra were recorded on an 800 MHz spectrometer while EM-ASR spectra were recorded on a 900 MHz spectrometer. Additional planes are shown in Figure 6-9.
**Figure 6-9** Additional overlaid representative 2D planes of 3D CANCO (A, B, H) and NCACB (C-G, I) experiments performed on PL-ASR (red), EM-ASR (black), and EM-NIC (blue). All peaks are labelled according to the assignments. PL-ASR and EM-NIC spectra were recorded at 800 MHz, while EM-ASR spectra were recorded on a 900 MHz spectrometer. In many cases, peaks which are observable in the PL-ASR spectra are not observed in the EM-ASR spectra due to the reduced sensitivity seen in these spectra. In other cases, ASR peaks cannot be identified in the EM-ASR spectra due to spectral overlap.

### 6.3.3 Structural evaluation of ASR in the E. coli inner membrane

With combined data from 2D and 3D experiments on PL-ASR, rbUCN EM-ASR, and rbUCN EM-NIC, N, CA, and in some cases CB peaks belonging to 82 of the 205 residues previously assigned in PL-ASR ([123, 124]) have been confidently reassigned in EM-ASR. Additional assignments for CO resonances in 29 residues have also been obtained (summarized in Table 6-1). Although several residues which are visible in PL-ASR are obscured by background peaks in EM-ASR, the majority of unassigned residues are missing due to insufficient sensitivity. Based on the signal to noise ratios of these peaks in the PL-ASR spectra, their absence in the EM-ASR spectra does not indicate that these residues undergo significantly increased motions. To verify that our spectra do not contain resolved peaks belonging to ASR
which we are unable to assign due to large changes in chemical shifts, we have confirmed that
any unassigned, resolved peaks found in the EM-ASR spectra are also found in the EM-NIC
spectra, and thus correspond to the background proteins.

The observable ASR residues are evenly distributed throughout the protein, found in
loops, the membrane-water interface regions, and transmembrane regions alike, thus providing a
site-specific view of conformational perturbations in ASR induced by the environment of the
inner membrane of *E. coli*. In Figure 6-10 we show residue-specific changes in chemical shifts
for N, CO, CA and CB atoms, as well as changes in peak intensities. The largest perturbations
range from up to a ppm in the $^{15}$N dimension to up to half a ppm in the $^{13}$C dimension, with most
of them occurring in helices B and D and in the extracellular B-C loop (Figure 6-11A). Chemical shift perturbations of this magnitude do not correspond to large changes in the
secondary structure of the protein, and we can thus conclude that the structure of ASR is largely
preserved in the *E.coli* inner membrane. As well, noticeable changes in cross peak intensities
ranging from 2 to 3x are observed in the E-F and C-D loops.

Many of the important structural features of rhodopsins previously observed for PL-ASR
remain well conserved in EM-ASR. For example, in line with the aforementioned observations
for unshifted peaks of the $^{15}$N SB and for S47, D75, and T79, located in the retinal binding
pocket, two additional residues interacting with the retinal (W76 and W176) are resolved in the
3D spectra, and display no significant changes in chemical shifts (Figure 6-11B), confirming
that the all-\textit{trans} conformation of retinal and the structure of the retinal binding pocket remain
unchanged.
Figure 6-10 Site-specific chemical shift perturbations and signal-to-noise ratio (SNR) comparisons between EM-ASR and PL-ASR. For N and CA shifts and for the relative SNR EM/PL data is combined from NCA, CANCO, and NCACB experiments. The relative SNR EM/PL is calculated with consideration for the number of scans performed on each sample and the relative amounts of UCN ASR present in each sample.
Three-dimensional spectroscopy also allows for further expansion of our view of the intermonomer interface within the trimer, and confirms that tight protein packing limits structural flexibility. The interface is comprised of interacting residues on helix B1 of one monomer (F42, W46, L49, M52, I56) and helix D2 (S107, T114) and helix E2 (V127, W131, C134) of a second monomer. Although most perturbed residues are found in helices B and D, they are not involved in the intermonomer interface. In addition to M52 and I56, located in extracellular half of helix B and observed in the 2D NCA spectrum, V127, W131, and C134 on the extracellular half of helix E, and an additional residue, L49, in the middle of helix B, do not experience significant chemical shift changes. Thus, the structurally conserved parts of helices B and E involved in the intermonomer interface extend at least to the entire extracellular half (Figure 6-11C).

M41 on the cytoplasmic side and M54 on the extracellular side are amongst the most perturbed residues in helix B. Likewise, there are noticeable (>0.5 ppm) perturbations observed for residues Q109, V112, S115, and D120 of helix D. All of the mentioned residues face either the interior of the trimer, neighboring helices within the same monomer, or lipids, and do not directly affect the monomer-monomer interface (Figure 6-11D). Specifically, M41 whose nitrogen chemical shift is perturbed by ~0.8 ppm, faces the lipid-exposed pocket at the interface formed by helices A and B of one monomer, and helices C and D of the neighboring monomer. It is thus possible that the shift in M41 peak is caused by a change of lipids in this pocket in EM-ASR. Additionally, S115 is tightly packed against G135 of helix E, and Q109 (chemical shift of the backbone nitrogen perturbed by ~0.7 ppm) is closely packed against T80 and P81 of helix C, which are unshifted themselves, indicating that the chemical shift changes observed in backbone atoms do not correspond to large changes in the conformation of amino acid side chains. The
lateral pressure exerted on ASR may differ between the tetragonal lattice of EM-ASR and the hexagonal lattice of PL-ASR, which could explain the changes in chemical shift observed in certain residues of helices B and D. Particularly, Q109 is known to participate in a strong interhelical H-bond (similar to D115 of BR) [263] and may serve as a sensitive indicator of changes in helical packing.

Figure 6-11 A) Structural model of a monomer of trimeric ASR (PDB 2M3G). In all figures unperturbed residues are shown in green, residues experiencing significant changes in chemical shift are shown in yellow, and residues with significantly enhanced signal intensity are shown in orange. B) Close-up of the retinal pocket, showing the side chains of selected unperturbed residues (green) which form the pocket. The retinal molecule is shown in blue. C) Side view and D) view from the extracellular side of the interaction interface (helices B1, E2, and D2) between two monomers and helix C2, which is tightly packed against many of the perturbed residues in helix D2. The side-chains of perturbed residues and residues involved in the interaction interface are shown. In C) unperturbed residues involved in the interaction interface are labelled, while in D) several of the most perturbed residues and residues which pack tightly against these residues are labelled. In A) and C) the cytoplasmic side is on top.
The B-C loop is well structured in many rhodopsins, and appears to be a common structural element which is often conserved in lipidic, detergent, and crystalline environments [32, 110, 114, 116, 264, 265]. In ASR the B-C loop is disordered in the X-ray structure [29], but gains partial β-structure in proteoliposomes [93]. It is composed of two short anti-parallel strands, V61-A63 and Q66-A68, and is subjected to collective motions as indicated by transverse relaxation measurements [125]. While there are some non-negligible chemical shift perturbations detected for the nitrogen of M54 in the immediate vicinity of the loop, for D57 in its flexible part, and for V61, the structure of the loop is largely unaffected by changes in the environment.

The E-F loop, which contains the peripheral helix E’, is another structural element commonly found in several other rhodopsins in the lipid environment [82, 114, 266, 267], but which exhibits different behaviors in detergent micelles and crystals. It is natively helical in BR in purple membrane samples [247], but either not visible or has large B-factors in high-resolution X-ray structures [30–32] and displays multiple conformers in samples prepared in detergent micelles [266] or nanodiscs [246]. While the chemical shifts of residues detected in the E-F loop of ASR (T154, T156, S158, and S159) remain largely unchanged in the E.coli membrane, there is a noticeable increase in the relative cross peak intensities for these residues (Figure 6-10E, Figure 6-11A), indicating partial immobilization of this loop in the E.coli membrane. This increase in intensity, along with the signal increase for K96 in the cytoplasmic C-D loop, may indicate a more symmetric position of ASR in the E. coli inner membrane. Our previous measurements indicated that the cytoplasmic side of ASR is more exposed in the DMPC/DMPA bilayer [123]. This asymmetry could be caused by a mismatch in the hydrophobic dimensions of ASR and the PL environment. While DMPC and DMPA both contain fatty acyl chains with a 14:0 composition, E.coli membrane contains mostly PE C16:0/cyC17:0, and C14:0, and longer
C16:1, C16:0, cyC17:0 and C18:1 constitute almost all of the fatty acyl substituents [268, 269], and will have a comparatively larger hydrophobic thickness.

Alternatively, the stabilization of the E-F loop in EM-ASR could be due to the removal of steric clashes between E-F loops of adjacent monomers in the hexagonal lattice. It has been shown that the E-F loops of adjacent trimers in the PM of BR can create steric conflicts under illumination [270, 271]. ASR, which is somewhat more cytoplasmically open [93, 123, 126] than BR, could produce such a steric clash in the PM-like lattice of PL-ASR, while in more loosely packed tetragonal lattice of EM-ASR, this conflict is removed, resulting in a different dynamic behavior of the E-F loop.

6.4 Summary and Conclusions

We have examined the conformation of a 7TM photoreceptor, ASR, in the environment of the E.coli inner membrane and compared it to that in a bilayer formed by DMPC/DMPA lipids. In both environments ASR forms trimers, as confirmed by visible CD spectroscopy and site-specific comparison of SSNMR chemical shifts of residues forming the intermonomer interface. SAXS measurements show that ASR reconstituted in DMPC/DMPA lipids forms small, 2D crystalline domains of hexagonal symmetry (typical size ~500 Å) with a lattice parameter of ~66.4 Å. While a two-dimensional lattice is also found in the E. coli inner membrane, the domains are likely smaller in size, and packing is of tetragonal symmetry and is looser, with a lattice parameter of ~74 Å.

Although the outer surfaces of ASR trimers in EM-ASR are more loosely packed and therefore more susceptible to the effects of the lipid environment, a comparison of solid-state NMR chemical shifts reveals that the overall structure of the protein in the E. coli inner
membrane remains similar to that in DMPC/DMPA lipids, and is not significantly influenced by different crystal packing or by interactions with lipids and other proteins present in the *E. coli* inner membrane. Many functionally important structural motifs are conserved between the two environments. In particular, the retinal remains in the all-trans conformation in the dark state in the *E. coli* membrane and, in full agreement with this, the conformation of the retinal-binding pocket is unperturbed.

While the intermonomer interface is structurally well conserved between the two environments, there are small but noticeable chemical shift perturbations detected in two of the α-helices, B and D, all occurring on the sides facing either the interior of a monomer, or the lipid bilayer. This may be the consequence of different lateral pressure profile experienced by ASR in the two systems studied. Additionally, the large, threefold increase in the relative cross peak intensities for residues in the cytoplasmic C-D and E-F loops indicate the possibility of a relative shift in the position of a trimer in the *E. coli* membrane bilayer as compared to that in proteoliposomes, or a removal of steric conflict between EF loops of adjacent trimers.

In contrast to small peptides and single-spanning membrane proteins, the inter-protein interactions in and between the ASR α-helical bundles likely minimize the effect of the environment on the protein structure. Data on the adaptability of membrane protein structures to their environment are of fundamental importance for understanding protein function, and for the validation of their high-resolution structures in cell membrane mimetics.
<table>
<thead>
<tr>
<th>Residue</th>
<th>EM-ASR Chemical Shift (ppm)</th>
<th>PL-ASR Chemical Shift (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>CA</td>
</tr>
<tr>
<td>L 6</td>
<td>116.28</td>
<td>57.17</td>
</tr>
<tr>
<td>L 7</td>
<td>117.46</td>
<td>58.85</td>
</tr>
<tr>
<td>V 12</td>
<td>121.20</td>
<td>67.84</td>
</tr>
<tr>
<td>G 14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M 15</td>
<td>120.95</td>
<td>56.78</td>
</tr>
<tr>
<td>T 16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I 17</td>
<td>119.29</td>
<td>65.41</td>
</tr>
<tr>
<td>A 19</td>
<td>122.14</td>
<td>56.59</td>
</tr>
<tr>
<td>N 28</td>
<td>115.41</td>
<td>50.90</td>
</tr>
<tr>
<td>P 29</td>
<td>130.22</td>
<td>65.12</td>
</tr>
<tr>
<td>V 32</td>
<td>125.54</td>
<td>59.63</td>
</tr>
<tr>
<td>E 36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y 37</td>
<td>115.32</td>
<td>63.40</td>
</tr>
<tr>
<td>L 38</td>
<td>121.55</td>
<td>58.93</td>
</tr>
<tr>
<td>A 40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M 41</td>
<td>111.90</td>
<td>59.18</td>
</tr>
<tr>
<td>I 43</td>
<td>114.92</td>
<td>67.73</td>
</tr>
<tr>
<td>P 44</td>
<td>132.00</td>
<td>65.34</td>
</tr>
<tr>
<td>W 46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S 47</td>
<td>106.13</td>
<td>60.01</td>
</tr>
<tr>
<td>G 48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L 49</td>
<td>121.39</td>
<td>58.25</td>
</tr>
<tr>
<td>M 52</td>
<td>126.15</td>
<td>59.02</td>
</tr>
<tr>
<td>A 53</td>
<td>118.45</td>
<td>54.71</td>
</tr>
<tr>
<td>M 54</td>
<td>113.21</td>
<td>58.55</td>
</tr>
<tr>
<td>A 55</td>
<td>126.94</td>
<td>55.53</td>
</tr>
<tr>
<td>I 56</td>
<td>111.09</td>
<td>60.48</td>
</tr>
<tr>
<td>D 57</td>
<td>114.75</td>
<td>56.35</td>
</tr>
<tr>
<td>Q 58</td>
<td>117.66</td>
<td>54.54</td>
</tr>
<tr>
<td>G 59</td>
<td>101.78</td>
<td>45.83</td>
</tr>
<tr>
<td>K 60</td>
<td>114.05</td>
<td>53.77</td>
</tr>
<tr>
<td>V 61</td>
<td>120.15</td>
<td>59.98</td>
</tr>
<tr>
<td>A 63</td>
<td>128.93</td>
<td>52.25</td>
</tr>
<tr>
<td>Q 66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I 67</td>
<td>119.16</td>
<td>58.96</td>
</tr>
<tr>
<td>A 68</td>
<td>134.28</td>
<td>50.80</td>
</tr>
<tr>
<td>A 71</td>
<td>124.83</td>
<td>55.18</td>
</tr>
<tr>
<td>Y 73</td>
<td>112.76</td>
<td>59.87</td>
</tr>
<tr>
<td>I 74</td>
<td>119.97</td>
<td>65.66</td>
</tr>
<tr>
<td>D 75</td>
<td>116.04</td>
<td>56.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>W 76</td>
<td>114.87</td>
<td>56.59</td>
</tr>
<tr>
<td>M 77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V 78</td>
<td>109.88</td>
<td>63.25</td>
</tr>
<tr>
<td>T 79</td>
<td>107.37</td>
<td>66.03</td>
</tr>
<tr>
<td>T 80</td>
<td></td>
<td>113.05</td>
</tr>
<tr>
<td>L 82</td>
<td>113.70</td>
<td>57.52</td>
</tr>
<tr>
<td>A 91</td>
<td>119.91</td>
<td>53.89</td>
</tr>
<tr>
<td>M 92</td>
<td>109.43</td>
<td>53.53</td>
</tr>
<tr>
<td>K 96</td>
<td>119.32</td>
<td>56.59</td>
</tr>
<tr>
<td>T 101</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I 102</td>
<td>117.16</td>
<td>66.53</td>
</tr>
<tr>
<td>M 106</td>
<td>113.70</td>
<td>61.16</td>
</tr>
<tr>
<td>Q 109</td>
<td>121.43</td>
<td>57.17</td>
</tr>
<tr>
<td>V 111</td>
<td>123.07</td>
<td>67.72</td>
</tr>
<tr>
<td>V 112</td>
<td>121.08</td>
<td>67.13</td>
</tr>
<tr>
<td>I 113</td>
<td>113.00</td>
<td>61.51</td>
</tr>
<tr>
<td>S 115</td>
<td>115.57</td>
<td>63.62</td>
</tr>
<tr>
<td>L 117</td>
<td>123.19</td>
<td>57.17</td>
</tr>
<tr>
<td>A 119</td>
<td>122.49</td>
<td>56.12</td>
</tr>
<tr>
<td>D 120</td>
<td>113.00</td>
<td>54.13</td>
</tr>
<tr>
<td>S 122</td>
<td>115.57</td>
<td>61.63</td>
</tr>
<tr>
<td>V 127</td>
<td>122.25</td>
<td>65.84</td>
</tr>
<tr>
<td>W 131</td>
<td>119.56</td>
<td>61.11</td>
</tr>
<tr>
<td>C 134</td>
<td>121.32</td>
<td>65.40</td>
</tr>
<tr>
<td>G 135</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V 136</td>
<td>123.35</td>
<td>67.43</td>
</tr>
<tr>
<td>C 137</td>
<td>118.62</td>
<td>65.49</td>
</tr>
<tr>
<td>G 145</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I 146</td>
<td>114.64</td>
<td>64.26</td>
</tr>
<tr>
<td>W 147</td>
<td>120.89</td>
<td>60.03</td>
</tr>
<tr>
<td>T 154</td>
<td>107.15</td>
<td>63.79</td>
</tr>
<tr>
<td>T 156</td>
<td>104.50</td>
<td>61.51</td>
</tr>
<tr>
<td>S 158</td>
<td>113.00</td>
<td>57.88</td>
</tr>
<tr>
<td>S 159</td>
<td>118.50</td>
<td>61.51</td>
</tr>
<tr>
<td>A 162</td>
<td>122.96</td>
<td>56.00</td>
</tr>
<tr>
<td>L 164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y 165</td>
<td>119.29</td>
<td>63.94</td>
</tr>
<tr>
<td>D 166</td>
<td>119.21</td>
<td>56.94</td>
</tr>
<tr>
<td>V 169</td>
<td>120.38</td>
<td>64.67</td>
</tr>
<tr>
<td>Y 171</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F 172</td>
<td>118.33</td>
<td>60.72</td>
</tr>
<tr>
<td>T 173</td>
<td>113.13</td>
<td>68.94</td>
</tr>
<tr>
<td>L 175</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>W</td>
<td>176</td>
<td>117.71</td>
</tr>
<tr>
<td>G</td>
<td>178</td>
<td>102.39</td>
</tr>
<tr>
<td>I</td>
<td>181</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>182</td>
<td>118.74</td>
</tr>
<tr>
<td>W</td>
<td>183</td>
<td>116.40</td>
</tr>
<tr>
<td>I</td>
<td>184</td>
<td>113.35</td>
</tr>
<tr>
<td>G</td>
<td>186</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>187</td>
<td>133.53</td>
</tr>
<tr>
<td>I</td>
<td>193</td>
<td>107.11</td>
</tr>
<tr>
<td>N</td>
<td>194</td>
<td>118.32</td>
</tr>
<tr>
<td>Q</td>
<td>195</td>
<td>115.81</td>
</tr>
<tr>
<td>F</td>
<td>200</td>
<td>120.38</td>
</tr>
<tr>
<td>L</td>
<td>201</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>202</td>
<td>115.48</td>
</tr>
<tr>
<td>C</td>
<td>203</td>
<td>117.92</td>
</tr>
<tr>
<td>G</td>
<td>212</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>213</td>
<td>119.83</td>
</tr>
</tbody>
</table>

130
Chapter 7
Conclusions and Future Directions
7.1 Concluding Remarks

The research presented in this thesis has centered around the development and application of novel SSNMR approaches to study large, seven transmembrane (7TM) α-helical proteins – a class of proteins of which microbial rhodopsins are representatives-- with a particular focus on the effect of the lipid environment on membrane protein structure. We have determined that under ultra-fast MAS conditions both covalently attached and exogenously added CuEDTA can uniformly distribute paramagnetic relaxation enhancement throughout these proteins. With the implementation of PACC, increased sensitivity per unit sample can be achieved, with the overall sensitivity available in these systems rivalling that which can be obtained from much larger rotors when proton-detection is implemented. These experiments are therefore an attractive option when limited sample amounts are available. Secondly, we have demonstrated that sensitivity enhancements of ~10x are available when direct proton detection is applied to the study of mobile resonances in fully protonated proteins under standard MAS conditions. These experiments provide a new method with which to study sufficiently mobile loop regions of membrane proteins, as well as the mobile regions of lipid and carbohydrate moieties which are tightly bound to, and may significantly affect the structure and function of, membrane proteins.

Finally, methods for the study of membrane proteins in the native E. coli membrane have been developed and applied to ASR. Through sample preparation procedures which strategically incorporate isotopes into ASR and the physical separation of regions of the cell which do not contain ASR, structural perturbations of ASR between the proteoliposome and E. coli membrane environments could be site-specifically analyzed using SSNMR. 40 % of residues, which were distributed evenly throughout the protein, were analyzed for changes in intensity or chemical shift, the results of which indicated that many functionally important structural motifs are
conserved in this environment, though moderate adaptations to the structure were observed which may result from changes in the bilayer thickness or protein crowding.

### 7.2 Continued Investigations of Membrane Proteins in Native Environments

The results presented in Chapter 6 represent the most extensive characterization of a membrane protein in the *E. coli* membrane environment reported to date. Though 40% of residues could be resolved, many more could not be analyzed due to insufficient sensitivity or overlap with background proteins. Several methods could be implemented in order to resolve these issues. To begin, the signals resulting from background proteins could be further eliminated from the spectra through the implementation of several recently developed methods. Specifically, the use of a mutant *E. coli* BL21Star(DE3) strain which lacks the two major OM proteins OmpF and OmpA [100], the repression of endogenous protein expression through the addition of rifampicin [272], and/or lower incubation temperatures during expression [101] could all lead to a significantly reduced presence of background signal, potentially leading to a significant increase in resolvable ASR residues in the spectra. In addition, the removal of OmpF and OmpA from the sample may allow for an increase in ASR content of the sample, and therefore increased sensitivity. The sensitivity of these experiments could also potentially be increased through the combination of these methods with Dynamic Nuclear Polarization (DNP), as has been demonstrated on similar samples [102–105]. However, the decrease in resolution which accompanies the DNP enhancement may lead to increased spectral overlap and preclude the obtaining of additional assignments. The ability to resolve additional resonances could facilitate the acquisition of a high-resolution protein structure, solved within the context of the *E. coli* membrane.
7.3 Detailed investigation on the effect of lipid acyl chain length on membrane protein structure

Results from our studies of EM-ASR have indicated that differences in the bilayer thickness of the *E. coli* membrane and the proteoliposomes more commonly used in NMR studies of ASR and PR could be the cause of several of the structural changes observed in EM-ASR. Studies which specifically investigate this hypothesis are currently underway in our group. As we have discovered in the course of these studies that ASR forms 2D crystals in lipid membranes, making it less ideal for the investigation of the effects of hydrophobic mismatch, we chose to investigate the effect of bilayer thickness on PR, which does not form 2D crystals.

Initially, three samples of PR reconstituted in lipids which differ only in the length of their acyl chains, DMoPC (14:1, 1,2-dimyristoleoyl-sn-glycero-3-phosphocholine), DOPC (18:1, 1,2-dioleoyl-sn-glycero-3-phosphocholine) and DEPC (22:1, 1,2-dierucoyl-sn-glycero-3-phosphocholine), were prepared at a relatively low protein:lipid ratio (~1:80 molar ratio) (Figure 7-1). These lipids were chosen to not only vary only in their acyl chain length, but also to have relatively low transition temperatures so that experiments could be performed on bilayers in the fluid phase. Although the low protein:lipid ratio allows for better isolation of PR monomers, thereby enhancing the effect of the hydrophobic thickness on protein structure, the increased lipid content of the samples leads to a decrease in the overall amount of protein in the sample, resulting in reduced sensitivity.

Initial SSNMR experiments were performed on the samples described above at a temperature of 25 °C. Unfortunately, significant sample degradation and protein-lipid phase separation were observed in these samples after the application of MAS SSNMR experiments. Therefore, two new samples were created in which PR was reconstituted at a 1:80 molar ratio.
into liposomes composed of DMPC/DMPA or DSPC/DSPA (1,2-distearoyl-sn-glycero-3-phosphocholine, 1,2-distearoyl-sn-glycero-3-phosphate) which were premixed at a 9:1 w/w ratio. Initial characterization of these samples using FTIR and one- and two-dimensional SSNMR indicate that the protein is well-folded and uniformly incorporated in the proteoliposomes. Site-specific analysis of changes in chemical shift and sensitivity are currently underway and three-dimensional experiments will be performed in order to extend this data.

Figure 7-1 Structural models of the lipids used in this study and a 3D model of PR based on homology modeling using the 3D-JIGSAW[273] program and the BR structure (PDB ID: 1C3W) as a template. Lipids and protein are not to scale.
References

137


52. Weingarth, M., Van Der Cruijsen, E.A.W., Ostmeyer, J., Lievestro, S., Roux, B., Baldus,


Agarwal, V., Diehl, A., Skrynnikov, N., Reif, B.: High resolution 1H detected 1H,13C


