uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Yanqiu Wu

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Chemistry)

GRADE / DEGREE

Department of Chemistry

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Development of Methods to Evaluate the Utility of Detergents for Membrane Protein Study by Solution NMR Spectroscopy

TITRE DE LA THÈSE / TITLE OF THESIS

Dr. N. Goto

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Dr. D. Bryce

Dr. J. Giorgi

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Development of methods to evaluate the utility of detergents for membrane protein study by solution NMR Spectroscopy

Yanqiu Wu

Thesis submitted to the
Faculty of Graduate & Postdoctoral Studies
University of Ottawa
In partial fulfillment of the requirements for the M.Sc. degree in the
Ottawa-Carleton Chemistry Institute

Thèse soumise à
Faculté des études supérieures et postdoctorales
Université d'Ottawa
En vue de l'obtention de la maîtrise es science à
L'Institut de chimie d'Ottawa-Carleton

uOttawa

L'Université canadienne
Canada's university

Candidate	Supervisor
Yanqiu Wu	Natalie Goto

Your file Votre référence
ISBN: 978-0-494-18480-6
Our file Notre référence
ISBN: 978-0-494-18480-6

Abstract

Structure determination of membrane proteins by solution nuclear magnetic resonance spectroscopy requires that they can be incorporated into detergent micelles in order to ensure that they are maintained in a folded, water-soluble state. However, some detergents appear to disrupt structurally important interactions between transmembrane (TM) helices, while others produce micelle-protein complexes that may be too large for solution NMR. As a result there are a limited number of detergents suitable for solution NMR studies, making the search for better solution conditions one of the remaining challenges for structure elucidation of membrane proteins by this technique. This thesis presents the development and application of methods to address this problem by characterizing detergents for suitability for membrane protein solution NMR. Since the overall size of the membrane protein-detergent complex is one of the major factors that affects the quality of NMR spectra, modified versions of pulsed field gradient (PFG) NMR pulse programs for translational diffusion measurement were set up to estimate the size of these complexes. These methods were applied to a model membrane protein-detergent system allowing an approximate complex size and composition to be determined. ^{15}N relaxation parameters were also measured for this complex and provided evidence that extensive motion of the protein within the micelle can occur to improve the quality of the resulting NMR spectrum. In addition to these experiments that probe

complex molecular weight, a model system was also established to help identify detergent characteristics that can promote helix-helix interactions. This system is based on the properties of the single transmembrane segment from the M13 bacteriophage major coat protein (MCP_{TM}) which forms homodimers in both lipid membranes and detergent micelles. Using solution NMR in combination with paramagnetic probes, oligomerization of this peptide under a range of detergent conditions was investigated. The results show that both denaturing and non-denaturing detergents are capable of preserving the TM-helix homodimer interaction. In accord with previous studies on this dimer, results from this study indicate that this peptide exists as a parallel dimeric structure at low concentrations of detergents. Together the methods described in this thesis have the potential to be used in the characterization of a wider range of detergent-protein complexes to assist in the development of new solvents that can ultimately advance the utility of solution NMR of membrane proteins.

Table of Contents

List of Figures

Abbreviations

BCA	Bicinchoninic acid
BPP	Bipolar gradient pulse pair
BSA	Bovine serum albumin
CD	Circular dichroism
CMC	Critical micelle concentration
DAGK	Diacylglycerol kinase
DHPC	L-α-1,2-dihexanoylphosphatidylcholine
DIEA	*N,N*-diisopropylethylamine
DPC	Dodecylphosphocholine
D_s	Translational diffusion coefficient
FID	Free induction decay
GpA	Glycophorin A
HATU	*N*-[(dimethylamino)-1*H*-1,2,3-triazolo[4,5-b]pyridine-1-ylmethylene] -*N*-methylmethanaminium hexafluorophosphate *N*-oxide
HSQC	Heteronuclear single quantum coherence
kDa	kilodalton
LED	Longitudinal eddy current delay
LOPG	(1-oleoyl-2-hydroxy-sn-glycero-3-[phosphor-RAC-(1-glycerol)])
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
MCP	Major coat protein
MTSL	(1-Oxyl-2,2,5,5-tetramethyl-Δ^3-pyrrolin-3-yl) methyl methanethiosulfonate
OD	Optical density
OG	β-octyl glucoside
PFG	Pulsed field gradient

PFGSE	Pulsed field gradient spin echo
ppm	Parts per million
RAW	Randomization approach to water suppression
rf	Radiofrequency
rpm	Rotations per minute
SDS	Sodium dodecyl sulphate
SDSL	Site-directed spin labeling
STE-PFG	Stimulated echo–PFG
ssNOE	Steady-state nuclear overhauser effect
T_1	Longitudinal relaxation time
T_2	Transverse relaxation time
TFA	Trifluoroacetic acid
TM	Transmembrane
UV	Ultraviolet
WT	Wild type
τ_c	Overall rotational correlation time

Acknowledgements

I would like to express my sincere gratitude to Dr. Natalie Goto, who sparked my interest in the present studies. Natalie gave me a good start and the freedom to follow my interests. Her wide knowledge and her logical way of thinking have been of great value for me.

I wish to thank Dr. Charles M. Deber from University of Toronto for providing the protein sample at the beginning of my research. I also extend my sincere thanks to Michael Barnes from NRC (Ottawa) for helping with the peptide synthesis.

I would like to thank all the members from the Goto lab for their generous assistance during the progression of this research. It was always appreciated.

Finally, I would like to dedicate this thesis to my parents whose example and inspirations make me work hard. Especially, I would like to give my special thanks to my husband Jie for his patience and support. Throughout it all, you were there.

Chapter 1

Introduction

1.1 Membrane protein structures

Integral membrane proteins (MPs) are embedded in the membranes of cells or organelles, forming an important part of the interface between the cell and the surrounding environment. The parts of these proteins that are associated with the membrane are composed of amino acid residues with hydrophobic side chains that can be immersed in the hydrophobic core of the phospholipids bilayer without incurring a large energetic penalty. However, since the amide bonds in the polypeptide backbone have polar, hydrogen-bonding character, membrane-embedded portions of these proteins tend to form structures that reduce the exposure of the backbone to the lipid environment. This is most commonly achieved by adopting an α-helical structure (Figure 1). When comprised of approximately 20 or more hydrophobic amino acids, the energy for the α-helix to traverse the membrane is typically large and negative. For this reason, membrane-spanning domains of membrane proteins are usually comprised of one or more hydrophobic α-helices. Less commonly, a transmembrane structure can also be achieved with multiple β strands so long as they form "barrels" that completely satisfy the internal hydrogen bonding capacity of each β-strand that crosses the membrane (Popot & Engelman, 2000; White, 1999; White, 2001; Tamm, 2004).

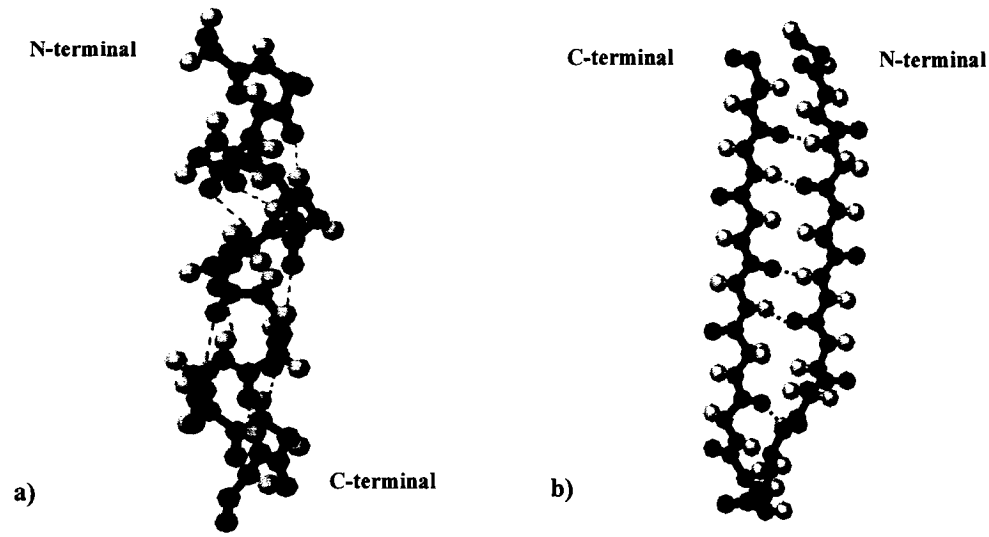


Figure 1: Ball and stick model of representative a) α-helical and **b)** β-sheet backbone secondary structures generated by MOLMOL (Karadi, 1996). The black balls represent backbone carbon atoms, red the carbonyl oxygen atom, and the blue and white the nitrogen and proton of the amide group respectively. In the α-helix hydrogen bonds between the oxygen atom of the carbonyl group from residue *i* and the amide proton from residue *i*+4 are highlighted by the green line. In the β-sheet structure hydrogen bonds are formed between backbone carbonyl oxygen atoms and amide protons between the two β strands, also shown with green lines. Side chains are not shown.

Since membrane proteins play an important role in key biological functions such as energy transduction, ion transport and regulation, molecular recognition and response. Given the functional importance of this class of proteins, knowledge of the three-dimensional (3D) structure of membrane proteins would be of great utility in the elucidation of their biological functions at the molecular level. However, our understanding of the function of this broad range of proteins has been hindered by a lack of structural information. In fact, this group is severely under-represented for this group in the database of high resolution protein structures, with less than 100 integral membrane proteins that have solved structures (White, 2006). The reason behind this striking deficit of structures is largely due to limitations normally associated with the

production of large amounts of these proteins and the difficulty in obtaining the crystals required for x-ray diffraction. In addition, the extreme hydrophobicity of membrane-embedded segments drives them to unfold and precipitate when excluded from the native lipid environment. However, this can be prevented by including detergents or lipids that are capable of maintaining a membrane protein in a folded, water-soluble form (Figure 2).

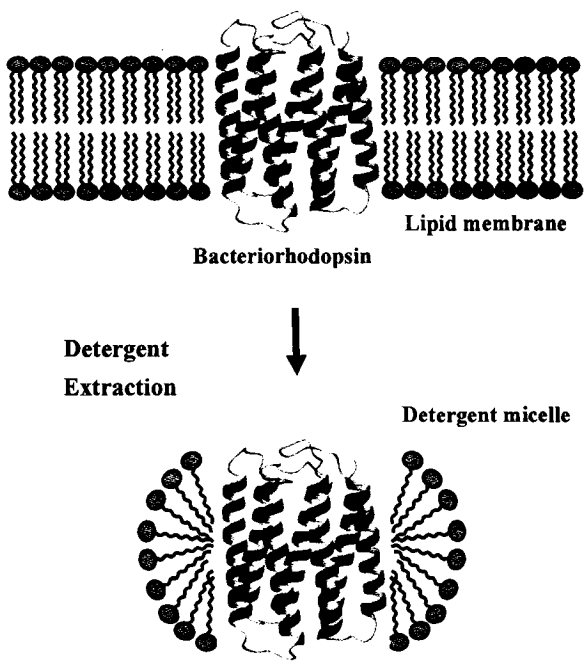


Figure 2: Schematic representation of a multi-spanning helical membrane protein in a lipid bilayer (top) and a micelle (bottom). To isolate membrane proteins such as bacteriorhodopsin (PDB ID 2BRD, shown as a ribbon diagram) from its native lipid bilayer environment (top), solubilization agents (detergents) are required. Micelles formed by detergents can maintain membrane proteins in a folded, water soluble form for structural studies by solution NMR.

Although most available membrane protein structures have been determined by x-ray crystallography, it can be difficult to find appropriate detergents that are compatible with crystal formation (Arora and Tamm, 2001). For this reason there is increasing interest in the application of another technique that has been widely applied to the determination of protein structures: nuclear magnetic resonance (NMR) spectroscopy. Our laboratory focuses on the use of high-resolution solution NMR spectroscopy for the study of membrane protein structure, a goal which has underlined the work described in this thesis.

1.2 Basic principles of NMR spectroscopy

NMR spectroscopy is a powerful technique that has been widely applied to the study of protein structure. This form of spectroscopy takes advantage of nuclei that have spin angular momentum, an intrinsic property which makes these nuclei behave like bar magnets (reviewed in Cavanagh, 1996). According to quantum mechanics, under the influence of an external magnetic field the magnetic dipoles associated with these nuclei will adopt a discreet number of orientations that depend on the spin quantum number of the nuclei involved. Typical nuclei used for NMR spectroscopy of biomolecules have a spin quantum number (I) of ½, including ^{1}H, ^{15}N and ^{13}C. These nuclei can adopt only two possible orientations relative to the external magnetic field; a low energy orientation called the α-state (also known as spin-up) and a slightly higher energy orientation called the β-state (spin-down).

As shown in Figure 3, the difference in energy between these two states

depends on the strength of the external magnetic field and the magnitude of the gyromagnetic ratio, an intrinsic property of each nucleus type. Consequently, if a population of spins in a strong magnetic field is subjected to a radiofrequency pulse with an energy that matches the energy difference between α- and β-states then this energy will be absorbed, giving rise to transitions between these states. NMR spectroscopy detects the frequency at which this absorption occurs, and typically reports this absorbance frequency in terms of a parts per million (ppm) shift from a reference frequency. Since this energy separation also depends on the local chemical environment of the spin, giving rise to a shift in the absorption frequency from reference values, this absorption frequency is also known as chemical shift. This is the basic signal that is detected when NMR spectroscopy is applied to samples isotropically tumbling in the solution state.

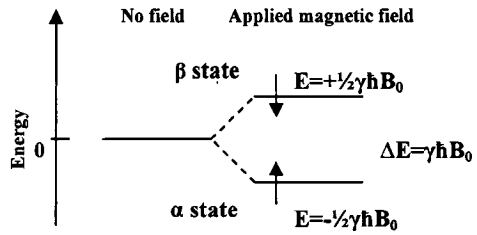


Figure 3: Energy level diagram for isolated nuclei with spin quantum number ½ in the absence (left side) and presence (right side) of an external magnetic field B_0. Under the influence of an external magnetic field the spins adopt two orientations that are slightly different in energy. The α-state spins (magnetic quantum number m_I = +1/2) have a low energy orientation (spin-up, denoted by the up arrow) of $E=-1/2\gamma\hbar B_0$ where ħ is Planck's constant and γ is the gyromagnetic ratio of the nucleus. The β-state spins (m_I = -1/2) have a slightly higher energy orientation (spin-down, shown by the down arrow) with a corresponding energy of $E=1/2\gamma\hbar B_0$. The energy required for a spin to undergo a transition from the α-state to β-state is $\Delta E= \gamma\hbar B_0$.

Our laboratory focuses on the use of NMR spectroscopy in the solution state for the study of membrane protein structure, a goal which has underlined the work described in this thesis.

1.3 Solution NMR of membrane proteins

Solution NMR has been widely applied to structure determination of a range of water soluble proteins, following procedures similar to those first demonstrated with a 6 kDa protein (Williamson & Wüthrich, 1985). Since then major improvements in spectrometer hardware, pulse sequence methodology, molecular biology and biochemical techniques have dramatically increased the range and utility of NMR for the characterization of protein structure and dynamics (Schmidt, 1984; Clore, 1989; Sattler, 1996; Tjandra & Bax, 1997). For example, the advent of uniform ^{15}N, ^{13}C isotope labeling raised the molecular weight limit of proteins for which protein structures could be determined from ~10 kDa to ~20 kDa (Bax, 1991). More recently, transverse relaxation-optimized spectroscopy (TROSY) (Pervushin, et al., 1997) has allowed the acquisition of high-resolution spectra of proteins with molecular weights exceeding 100 kDa (Salzmann et al., 1998; Wider, 1999; Pervushin, 2000; Salzmann, 2000). However, in spite of impressive advances that have been made in the application of solution NMR to water-soluble proteins, much less progress has been made in its application to membrane proteins.

Most integral membrane proteins contain hydrophobic sequences approximately 20 amino acid residues in length that cross the lipid bilayer in an

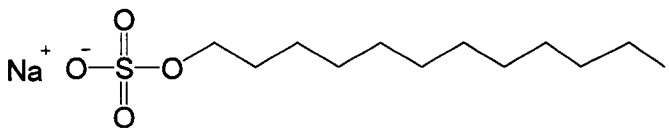
α-helical conformation (Popot & Engelman, 2000). While small membrane proteins or membrane protein fragments containing a single transmembrane (TM) helix have proven amenable to structure determination by solution NMR in detergent micelles (MacKenzie, 1997; Papavoine, 1998), its application to proteins with more than one TM helix has been extremely limited. To avoid the extra mass of the detergent micelle, a mixture of organic solvents and water has been used to study the structures of a small number of multi-spanning α-helical membrane proteins (*e.g.* the multidrug transporter EmrE, subunit c of the E. *coli* F_0F_1 ATP, and bacteriorhodopsin fragments, Schwaiger, 1998; Girvin, 1998.). However, organic solvent-water mixtures are generally poor mimics of the lipid bilayer environment and appear to disturb tertiary contacts between TM helices (Buck, 1995; Lou, 1997). In contrast, the amphiphilic character of a lipid or detergent micelle provides both hydrophilic and hydrophobic domains that are capable of preserving the active form of a membrane protein. This has been recently demonstrated by solution NMR structures determined for multi-spanning β-barrel types of membrane proteins in detergent micelles (Fermández, 2001; Arora, 2001; Hwang, 2002). However, the more commonly occurring multi-spanning helical type of membrane protein is considerably more hydrophobic than the β-barrel and is also generally less stable, making it a much greater challenge for structure determination. To date there are no examples of solution NMR structures for typical multi-spanning helical membrane proteins in a more native-like micelle environment. This reflects the fact that structure determination of multi-spanning α-helical membrane proteins comprises one of the last frontiers in biomolecular

NMR.

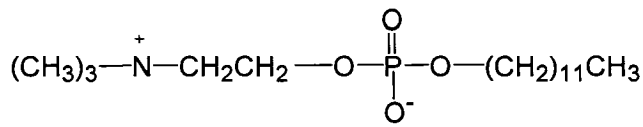
1.4 The need for new detergents for solution NMR studies

The use of detergents in solution NMR studies of membrane proteins can give rise to two significant problems. Firstly, the membrane protein-detergent complex formed can become very large, surpassing the molecular weight limit of solution NMR. A second and equally important issue is the possibility that protein activity, structure and/or dynamics may be perturbed by strong interactions between amphiphilic detergent molecules and the protein of interest. Owing to these issues there is only a very small range of detergents that have been used in membrane protein NMR, (e.g. sodium dodecyl sulfate, dodecyl phosphocholine). Although these detergents form small micelles they also tend to destabilize interactions between TM helices (Choma, 2000; Zhou, 2000; Gratkowski, 2001; Therien, 2002) generally limiting their use to the study of single-spanning TM helices. The utilization of milder non-denaturing detergents has also been problematic for solution NMR since they often form larger protein-detergent complexes that give rise to NMR spectra with broad peaks and poor resolution (e.g. Triton X-100, Krueger-Koplin, 2003). Consequently, the successful application of solution NMR to membrane protein structure determination could benefit significantly from the development of a wider range of NMR-compatible mild detergent systems.

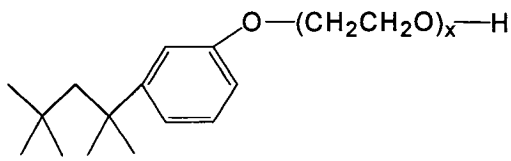
SDS



DPC



TRITON X-100



DHPC

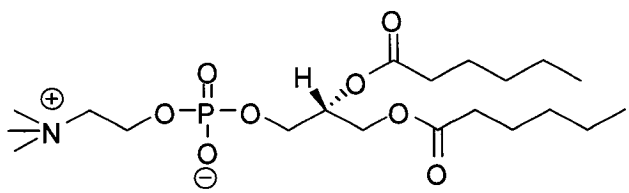


Figure 4: Chemical structures of common detergents.

An additional complication in the study of membrane protein structure by solution NMR has been the difficulty in predicting how a membrane protein will behave in different detergents. In fact membrane protein folding and stability will depend on the balance of protein-protein and protein-detergent interactions required to

maintain the complex in the solution phase while preserving native interactions between TM helices. In addition, the nature of this balance will depend on the membrane protein under study, as well as the detergent being used. Even in the relatively simple case of a single TM helix interaction formed within a homodimer, the affinity of the TM helix interaction is dependent on the nature of the detergent micelle into which it inserts (Fisher, 1999; Therien, 2002). In addition, very small detergent micelles, such as those formed by L-α-1,2-dihexanoylphosphatidylcholine (DHPC), can induce non-native curvature in micelle-associated peptides (Chou, 2002), highlighting the influence that detergents can have on membrane protein structure. These issues further demonstrate the need for the development of a wider range of detergent systems that can provide a more natural environment for a membrane protein in an NMR-compatible micelle.

1.5 Thesis objectives

Given the need for a wider range of detergents that could be used for solution NMR of membrane proteins it would be useful to first develop tools that can characterize the suitability of various detergents. Therefore the aim of this thesis has been to develop methods that can probe membrane protein tertiary structure in the detergent micelle while evaluating the size of the detergent-micelle complex. Towards the latter goal, I developed and implemented NMR experiments for the measurement of translational diffusion coefficients as described in Chapter 2. These experiments were tested on a model membrane protein composed of a single TM helix in SDS micelles.

^{15}N NMR relaxation studies were also performed on this system to evaluate the internal motion of the membrane protein within the detergent micelle since this motion can lead to improvements in NMR spectral quality. Chapter 3 describes the development of a technique to study the interaction between TM helices by combining spin-labeling with NMR spectroscopy. Successful implementation of this method helped to establish that the system under investigation forms parallel dimers in the non-ionic detergent β-octyl glucoside (OG). Overall these techniques described in this thesis open the door to future investigations on the suitability of different detergents for structural studies by solution NMR.

Chapter 2

Evaluation of the size and dynamics of membrane proteins in detergent micelles using solution NMR

2.1 Introduction

Detergents that are utilized to solubilize membrane proteins may form large protein-detergent complexes which give rise to broad peaks that would not be compatible with solution NMR studies. In order to determine whether this will be a problem for a particular protein-detergent complex, it is necessary to implement a method that can be easily used to estimate the size of these complexes. This method, in combination with spin relaxation data, could then be used to screen detergents for suitability in solution NMR studies. Therefore the primary goal in this chapter is the development and application of NMR experiments that can be used for this purpose.

2.1.1 NMR spectroscopy of large molecular weight systems

The quality of a solution NMR spectrum is a function of sample sensitivity and resolution, where intense, well-separated narrow peaks would comprise the ideal spectrum. Among a number of variables that can affect spectral quality, the size of the molecule being analyzed is one of the most important factors (Claridge, 1999). The linewidth of an NMR signal depends on the exponential decay of the NMR signal over time, which is related the transverse relaxation time (T_2). This parameter is

modulated by the molecular tumbling rate in solution which is represented by the rotational correlation time (τ_c), the average time it takes for the molecule to rotate through one radian. For a spherical molecule rotating in a liquid, τ_c can be described using the Debye equation (Debye, 1945; Bloembergen, 1948; Hennel, 1993)

$$\tau_c = \frac{4\pi\eta R_s^3}{3k_B T} = \frac{V\eta}{k_B T} \qquad [1]$$

where k_B is the Boltzmann constant, T the temperature, η the viscosity of the medium, and R_s and V are the hydrodynamic radius and the volume of the molecule respectively. Assuming that the volume is proportional to molecular weight, the rotational correlation time can therefore be related to the molecular weight.

As indicated by Equation 1, larger molecules will tumble more slowly in solution. Since T_2 is a function of τ_c, these relationships indicate that molecules with larger τ_c values will have shorter T_2 times. Therefore, larger molecules tend to suffer from rapid transverse relaxation rates that reduce both signal intensity and resolution. This is a particular issue for solution NMR studies of membrane proteins, since detergents must be used to maintain the hydrophobic protein in solution. The result is a membrane protein-detergent micelle complex, usually of large molecular mass with long rotational correlation times. Consequently, solution NMR of membrane proteins will have a greater tendency to suffer from the detrimental effects of large, slowly tumbling molecules. For this reason, detergents must be chosen that can maintain the membrane protein in solution while minimizing the size of the protein-detergent complex. However, in order to find the appropriate conditions for a particular

membrane protein it would be useful to have methods available to determine the size of the membrane protein-detergent complex in the same conditions being used to record the NMR experiment.

2.1.2 Diffusion coefficients as an estimate of complex size

The size of the protein-detergent complex can be determined by measuring the self-diffusion coefficient (D_s), a parameter which describes the translational motion of the assembly through the sample medium and is indicative of hydrodynamic properties and molecular size.

If a spherical shape is assumed for the complex, the diffusion coefficient D_s can be described by the Stokes-Einstein equation:

$$D_s = \frac{k_B T}{6\pi\eta R_s} \qquad [2]$$

where the variables are the same as those defined in Equation 1. As predicted by this equation, large molecules move more slowly than small molecules, and therefore exhibit a small degree of self-diffusion compared to small molecules. Consequently larger molecules have smaller diffusion coefficients that can be more difficult to measure.

The size of a protein-detergent complex can be estimated from the translational diffusion coefficient as measured by pulsed field gradient (PFG) NMR. The use of NMR for quantitative measurement of translational diffusion coefficients has become widespread in the past decade with the introduction of reasonably linear, self-shielded gradient coils in high resolution NMR probes (Altieri et al., 1995;

Dingley, 1995; Kuchel & Chapman, 1993; Andrec & Prestegard, 1996; Gounarides, 1999). Although there are other approaches available to determine the size of a molecule, *e.g.* light scattering, molecular sieving, ultracentrifugation (Cavanagh, 1996), the advantage of PFG-NMR methods for this study is that the measurements can be carried out under experimental conditions (protein and detergent concentration, temperature, viscosity, etc.) identical to those that are used to collect structure NMR data used for structural determination.

2.1.3 Pulsed field gradients (PFG) in NMR

There are a number of NMR methods that exist for the measurement of self-diffusion, all of which employ pulsed field gradients (PFGs). During application of a PFG, the external magnetic field B_0 is made deliberately inhomogeneous by the application of a magnetic field gradient. For example, when a PFG is applied along the direction of the external magnetic field vector (the z-axis), the magnetic field becomes inhomogeneous in a position-dependant manner. Consequently the Larmor frequency varies with position along this axis causing nuclei to precess at frequencies that depend upon their longitudinal position as shown in Figure 5 (Stejskal & Tanner, 1965). Under these conditions the Larmor frequency of a nucleus at position z along the z-axis, ω_Z (rad s^{-1}), can be described by Equation 3.

$$\omega_z = \gamma B_0 + \gamma G_z z \qquad [3]$$

where γ is the gyromagnetic ratio, B_0 (in units of Tesla, T) is the external magnetic field strength and G_z (in units of $T \cdot m^{-1}$) is the gradient applied along the z-axis. If we

consider the effect of the applied gradient pulse in the rotating frame (where the coordinate system rotates at Larmor frequency, ω_0, then the gradient pulse applied for a time δ imposes a position-dependent phase angle, φ_z (rad), given by:

$$\varphi_z = \delta\gamma G_z z \qquad [4]$$

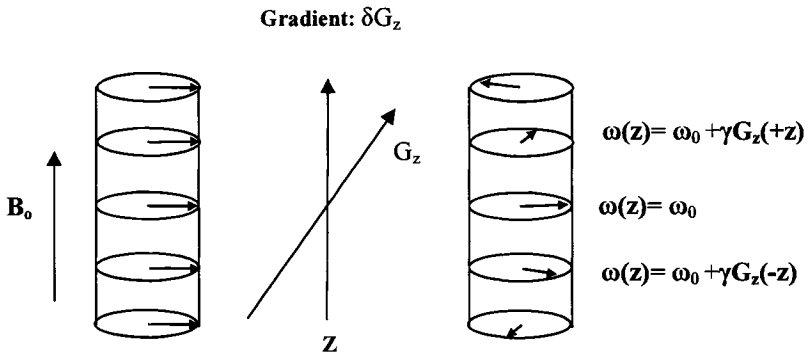


Figure 5: Schematic diagram showing the effect of an applied gradient on the bulk magnetization vector at different positions along the z-axis. In this example the magnetic field changes linearly with z. During application of this gradient pulse spins near the bottom of the NMR sample will precess more slowly than spins at the top of the sample. The resulting position-dependent phase differences are shown on the right.

As shown in Figure 5, during application of a linear z-gradient, the spins at the centre of the sample would not experience a change in magnetic field strength, while those above and below the centre would have higher and lower frequencies respectively due to the different magnetic fields at those positions. The resulting position-dependent de-phasing can be reversed if another gradient equal in magnitude

but opposite in intensity to the initial gradient is subsequently applied. This ability to use PFGs to de-phase and re-phase magnetization in a position-dependent manner is the general principle that allows the measurement of diffusion coefficients using solution NMR.

2.1.4 Use of PFG approaches for measurement of diffusion coefficients

It is possible to measure the translational diffusion coefficient D_s by PFG-NMR approaches using the simple PFG spin-echo (PFGSE) experiment (Stejskal and Tanner, 1965) shown in Figure 6.

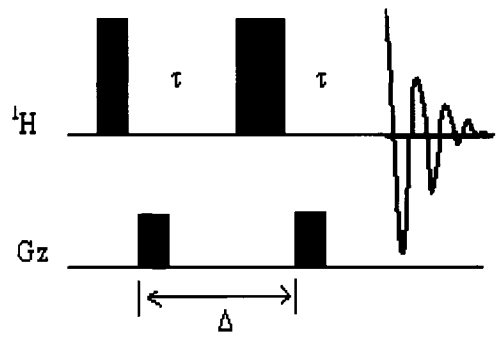


Figure 6: Pulsed field gradient spin echo (PFGSE) sequence. The narrow black rectangle represents a 90° radiofrequency pulse on protons applied along the x-axis, and the wide rectangle is a 180° pulse. Shaded grey bars are the PFGs of magnitude G_z and duration δ. The NMR signal is recorded after the second delay τ.

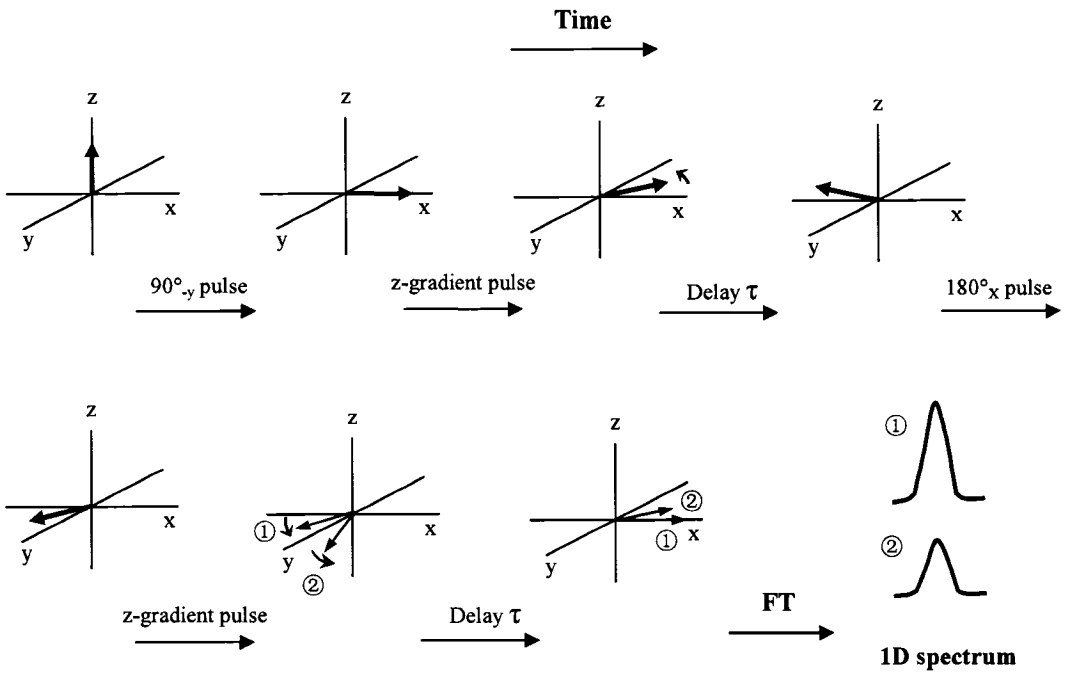


Figure 7: Vector diagram representation of the evolution of the bulk magnetization vector associated with two isolated spins that start at same z-position in the NMR sample. Assuming that both spins evolve under the influence of chemical shift during delay time τ at the same rate, (*i.e.* both have the same resonant frequency) then the magnetization vector arising from the two spins should not be distinguished until the second gradient pulse. The first gradient pulse imposes the same phase on these two spins. However, if spin ① does not diffuse during the time of the pulse sequence, its magnetization will be refocused after the second gradient pulse, while if spin ② does diffuse the second PFG will fail to perfectly reverse the phase imposed by application at the first PFG. The result on the NMR spectrum for each spin is shown in the last panel.

The evolution of magnetization through this pulse sequence can be described using classical vector diagrams (shown in Figure 7). In this experiment a non-selective 90° ^{1}H pulse rotates proton magnetization into the transverse plane, where upon a PFG is applied to 'label' or 'encode' spatially dependent phase angles

(φ_z). During the subsequent diffusion period (Δ), the spins are allowed to evolve under the influence of chemical shift while diffusion also occurs. The 180° pulse followed by a second delay τ allows complete refocusing of the magnetization vector back to the x-axis, so long as the strength of the gradient pulse before the refocusing pulse is the same magnitude as that after the refocusing pulse. This will be the case if the spins (e.g., spin 1 in Figure 7) do not undergo translational diffusion during the time of the pulse sequence. However, if the spins do diffuse to a different spatial position along z (e.g., spin 2 in Figure 7), then the second PFG will be different in magnitude from the first PFG. In this case refocusing will not be complete and the resulting signal intensity will be attenuated. Therefore, the signal intensity after the decoding gradient will depend on the degree to which the magnetization was re-phased by the second refocusing PFG, which in turn depends on the extent of diffusion that occurred during time Δ.

Although the PFGSE sequence is a simple method that can be used to measure diffusion constants of small molecules, macromolecules with small translational diffusion coefficient ($D_s \sim 10^{-9}\ m^2/s$) require the use of strong gradients and long diffusion delays between gradient pulses in order to observe the effect of diffusion on the signal. However, signal decay due to rapid transverse relaxation prevents the application of this method to larger molecules and hence a variation on this basic sequence must be used to enable the measurement of smaller diffusion coefficients.

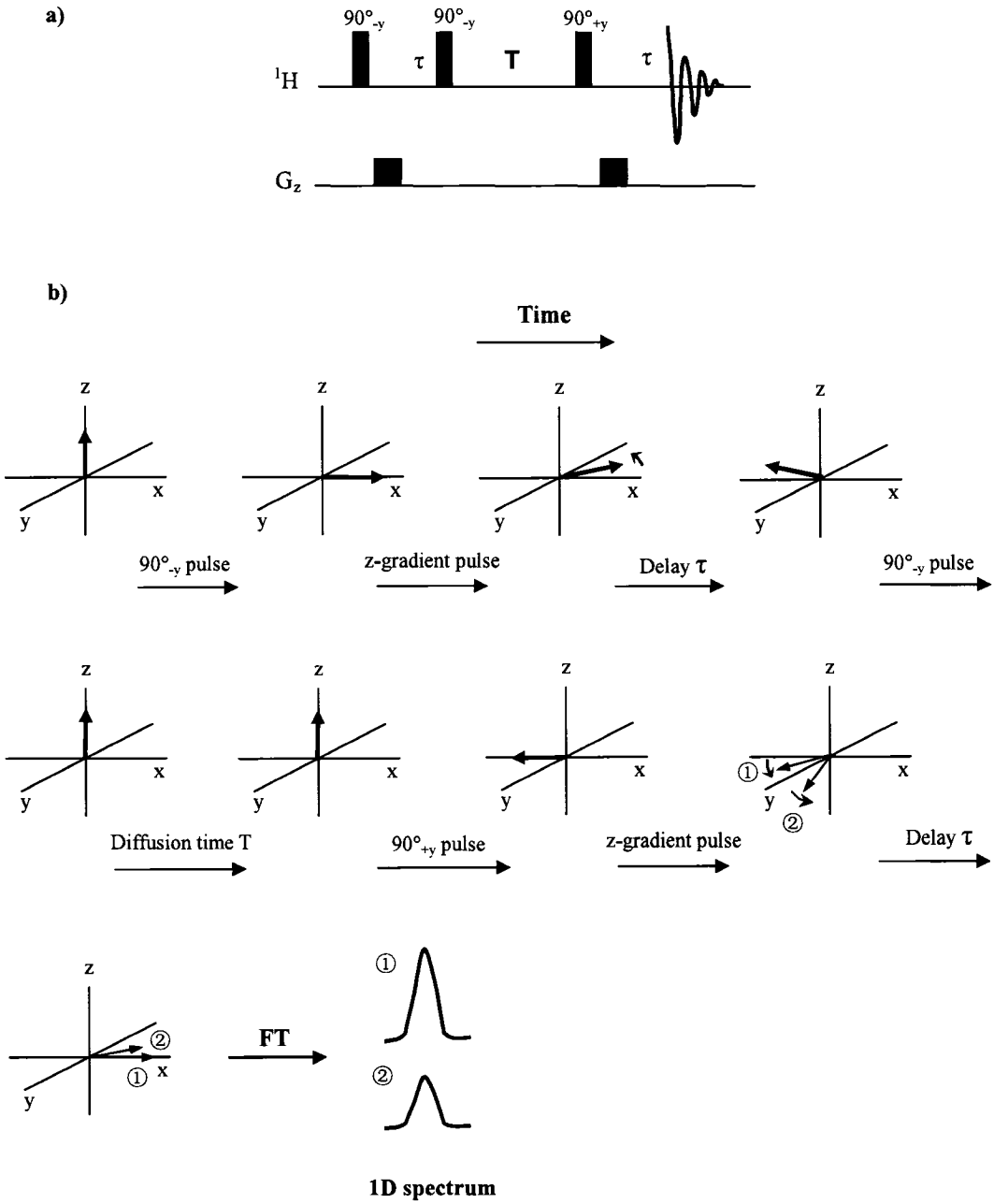


Figure 8: a) Basic pulse sequence for STE-PFG experiment. b) Vector diagram representation of the evolution of magnetization through the STE-PFG pulse sequence for two isolated spins originally located at the same z position in the NMR sample. Spin 1 does not diffuse allowing its magnetization to be refocused at the end of the sequence. Conversely, if a spin diffuses during T as shown for spin 2, then refocusing will be imperfect leading to a reduction in signal intensity shown on the right.

A popular way to measure diffusion coefficients for larger molecules takes advantage of the fact that large molecules have long longitudinal relaxation times. For example, amide protons in lysozyme (a 14.3 kDa protein) exhibit an average T_1 of 1.0 s while the average T_2 is 26.8 ms (Altieri et al., 1995). Therefore, if the transverse coherence is stored along the z-axis prior to the diffusion delay, much longer diffusion times can be utilized than are possible in the basic PFG experiment. This can be achieved with the sequence shown in Figure 8 a), called the stimulated echo–PFG (STE-PFG) experiment.

As shown in Figure 8 b) the PFG following the first 90° ^{1}H pulse is used to 'encode' the position of the spins prior to the diffusion delay. After chemical shift evolution during delay τ, another 90° proton pulse will move the transverse ^{1}H magnetization along the z-axis. Diffusion is then allowed to occur for time T during which the stored signal decays slowly at a rate that is primarily determined by the longitudinal relaxation time. The magnetization is then transferred back to the transverse plane by another 90° ^{1}H pulse and another gradient pulse applied. Again, as was seen in the PFG experiment, the magnitude of the gradient experienced by each spin will depend on its position along z and will not be the same as that experienced during the first PFG if significant diffusion has occurred during time T (e.g. spin 2 in Figure 8 b). In this case the magnetization will not be completely refocused after the second chemical shift evolution period, leading to a reduction in signal intensity. This basic STE-PFG scheme therefore achieves the same result as the PFGSE experiment with the advantage that it can be applied to larger molecular weight systems,

including proteins in complex with detergent micelles.

2.1.5 Eddy current effects

The use of strong gradient pulses such as those used for diffusion measurement also raises potential problems with the occurrence of eddy currents which are induced in the sample by the gradient. Since eddy currents can cause inhomogeneity in the magnetic field, they can give rise to severe signal distortions in NMR spectra. Therefore it is necessary to use pulse sequences that can minimize these effects prior to data acquisition. These can be achieved by two approaches. The first approach is longitudinal-eddy-current-delay (LED) sequence (Gibbs & Johnson, 1991), in which an eddy current delay time, τ_e, is incorporated after the gradient echo and before acquisition of the NMR spectrum. As a result signal distortion is reduced since a settling period for eddy currents allows the field to equilibrate before signal acquisition. However, one drawback is that the eddy current delay time τ_e leads to signal attenuation through T_1 relaxation. Even more serious is the problem that the time constant for eddy-current decay can be on the order of hundreds of milliseconds, and consequently, some lineshape distortion could remain even with an eddy current delay time greater than 100 ms.

Another approach that can be used in addition to LED to suppress the effect of eddy currents is the introduction of bipolar gradient pulse pairs (BPP) (Wider et al., 1994; Wu et al., 1995). In this method a single gradient of duration δ is replaced by a pair of gradients of duration δ /2 pair that are opposite in polarity but equal in

magnitude PFGs, separated by a 180° radiofrequency pulse. Since the 180° pulse negates the sign of the phase change generated by the first gradient, and the second gradient adds to the phase change of the first to generate a total phase change that is equivalent to a single PFG of duration δ. Since the BPP approach uses shorter gradients, less eddy current is generated. Moreover, the opposite polarity of the second gradient effectively compensates for the eddy current created by the first gradient to dramatically reduce the eddy-current effects, thereby improving the overall quality of the spectrum.

2.1.6 Water suppression

An important consideration in the study of any aqueous protein sample by solution NMR is that the concentration of water protons approaches ~110 M whereas the concentration of a typical protein sample is only on the order of ~1 mM. This large excess of water protons gives rise to a huge signal from water that affects the ability to detect protein signals (Figure 9). This difficulty arises from limitations imposed by the dynamic range of the electronic components of the spectrometer. Specifically, the analog-to-digital converters cannot adequately digitize the protein signal when this large signal from water is present. Moreover, observation of protein signal can also be obscured by the broad water signal (Cavanagh, 1996). Although it is possible to eliminate this problem by making the sample in 100% D_2O, this would eliminate signals from the amide protons since they are exchangeable with the solvent water protons. Since amide protons are of great value in protein structure

determination it is usually necessary to employ pulse techniques that can selectively suppress the NMR signal from water protons.

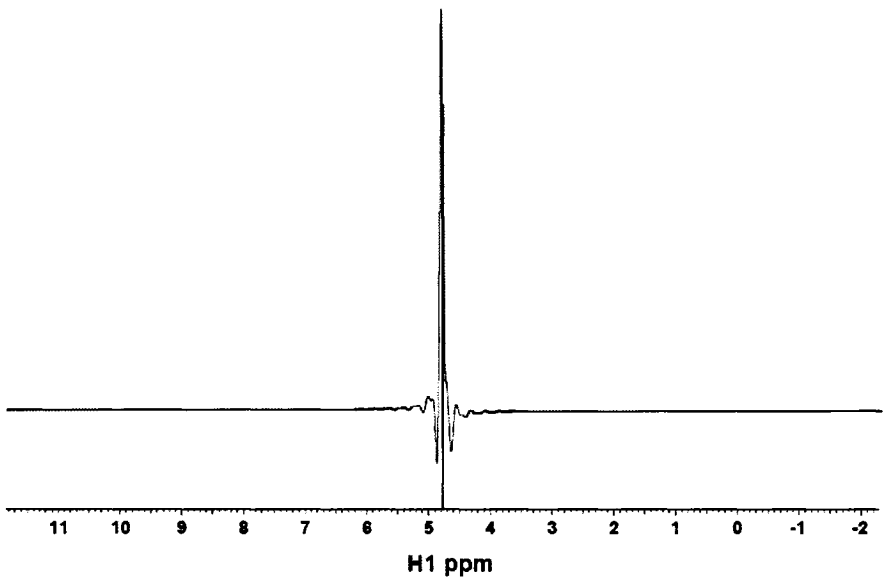


Figure 9: 1D spectrum of lysozyme obtained without water suppression. Sample contains 2 mM lysozyme, 50 mM sodium phosphate, pH 6.0, 10% D_2O, at 25 °C.

In order to suppress the large water signal, it is possible to use the randomization approach to water suppression (RAW) (Altieri & Byrd, 1995) at the beginning of a pulse sequence. The RAW sequence is comprised of a water-selective Gaussian-shaped pulse which is used to selectively rotate the magnetization from the water protons into the transverse plane, followed by a z-gradient pulse which de-phases the transverse water magnetization. This eliminates the component of the bulk magnetization vector originating from water and prevents it from being stimulated by the subsequent non-selective 90° pulse. Utilization of a

Gaussian-shaped pulse with a longer duration and lower power gives rise to high selectivity in the excitation profile since an increase in pulse width decreases the range of frequencies that the radiofrequency (rf) field will be effective in exciting (Figure 10). If this pulse is carefully calibrated such that it is centered on the water proton resonance frequency with a power level adjusted to achieve a 90° tip angle, high selectivity for water protons should be achieved, allowing selective elimination of the water signal (Kyogoku, 2003).

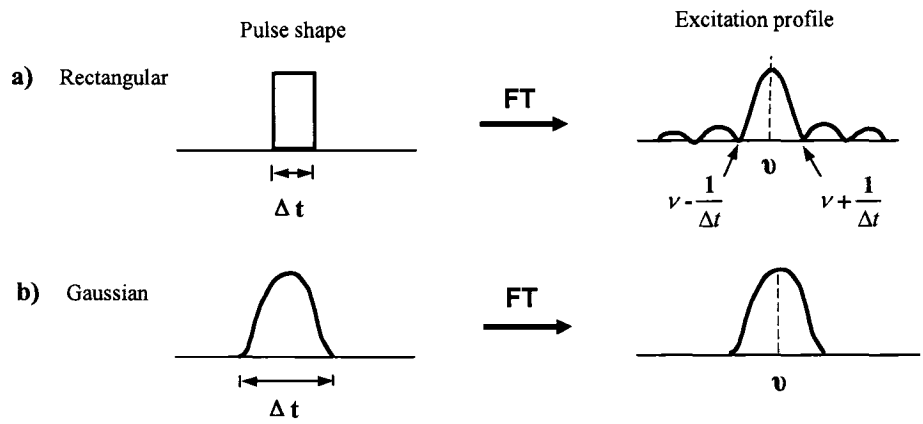


Figure 10: a) A rectangular pulse of frequency υ and of duration Δt will excite the range of frequencies given by its Fourier transform (FT). The result is a |sinc| profile, where the bandwidth of each lobe is approximately $\frac{2}{\Delta t}$. Therefore the longer the pulse the narrower the excitation profile. b) Elimination of the additional bands beyond $\nu \pm \frac{1}{\Delta t}$ in the excitation profile is achieved through the use of a Gaussian shaped pulse, since Fourier transform of a Gaussian shaped pulse is also Gaussian.

Since water magnetization has different relaxation characteristics relative to protein magnetization, at the end of a diffusion delay in a RAW-suppressed sequence

longitudinal relaxation may have restored a significant amount of water magnetization back to the z-axis. To suppress signal from this restored magnetization, a final water suppression element called WATERGATE (Figure 11) (Piotto, et. al., 1992) can be used at the end of the sequence. In this approach a series of selective 90° pulses with a nonselective 180° pulse to selectively refocus protein signals while dephasing solvent signals. As a result, most of the signal from water protons is effectively eliminated, allowing observation of protein peaks.

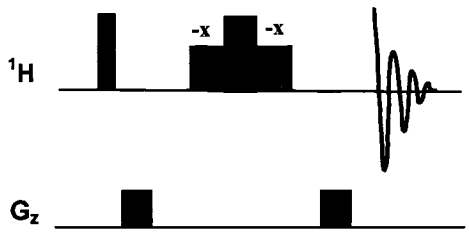


Figure 11: WATERGATE pulse sequence. A nonselective pulse rotates the magnetization from both protein and solvent water into the transverse plane and the following strong gradient pulse dephases this magnetization. A water-selective proton pulse (grey) that has no effect on solute magnetization is then applied, rotating the solvent magnetization back to the z-axis. The following nonselective pulse rotates all magnetization vectors by 180° while the second water selective pulse (grey) returns the solvent magnetization back to transverse plane. As a result, application of another gradient pulse equal in duration and intensity to the previous one will rephase the solute magnetization, while the solvent magnetization will continuous to be dephased. In other words, the combination of the selective 90° pulses and the nonselective 180° pulse is equivalent to a 360° pulse for solvent protons and a 180° pulse for protein protons.

2.1.7 Spin relaxation and protein dynamics

Although a large membrane protein-detergent complex will have a long

overall rotational correlation time (τ_c), internal rotational motion of a membrane protein within its detergent micelle can decrease τ_c for the protein itself, thereby improving the quality of the protein NMR spectrum. Therefore, in addition to the diffusion measurements that utilized elements outlined in the previous section, we also studied the ability of the detergent to permit motion of the membrane protein within the confines of the micelle by measuring longitudinal and transverse ^{15}N spin relaxation times. In addition, the steady state nuclear Overhauser effect (ssNOE) was also measured to monitor relative mobility of residues within the protein. Since less mobile spin pairs give rise to more efficient NOE transfer the magnitude of this effect depends on the rotational correlation time of the spin pair being examined. Assuming the absence of internal motions on an intermediate time scale (ns - ms), these relaxation data can be used to extract an estimate for the apparent size of the complex via Equation 1 (Kay et al., 1989; Cavanagh et al., 1996). Since this relationship only holds if the molecular complex is spherical with rotational dynamics that can be characterized by a single correlation time, differences in size estimates obtained from ^{15}N spin relaxation and translational diffusion coefficient measurements can be used to evaluate whether there is rotational diffusion of the membrane protein within a particular detergent micelle.

2.1.8 Model system for diffusion measurement - Pf1 coat protein

To test the suitability of the pulse sequences for the measurement of diffusion coefficients of micelle-embedded membrane proteins, we used a simple model system

that could easily be produced in ^{15}N-labeled form: the major coat protein from the filamentous bacteriophage Pf1. As shown in Figure 12, this is a small protein of 46 amino acids. The structure and dynamics of this coat protein in both lipid bilayers and detergent micelles have been extensively characterized by solid-state (Nambudripad, 1991; Shon & Opella, 1991) and solution NMR (Schiksnis, 1987; Lee, 2003). These studies showed that the Pf1 coat protein contains a hydrophobic helix (residue 19 to 42) that spans the membrane (or micelle) and a short amphipathic helix (residue 6 to 13) located on the membrane surface. Since this protein-detergent complex is already well characterized and can be straightforwardly produced using established protocols, it is an ideal test system that can be used the development and implementation of the NMR experiments described above.

```
1         10        20        30        40        46
GVIDTSAVQSAITDGQGDMKAIGGYIVGALVILAVAGLIYSMLRKA
```

Figure 12: Amino acid sequence of Pf1 coat protein (Nakashima et al., 1975). The single underline denotes the N-terminal amphipathic helix. The double underline indicates the α-helical transmembrane domain.

2.1.9 Objectives

In this work, pulse sequences were developed and optimized to enable the measurement of diffusion coefficients on the 500 MHz Varian INOVA NMR spectrometer at University of Ottawa. The Pf1 coat protein dissolved in SDS micelles was utilized as a model system to test the suitability of these sequences to measure

translational diffusion coefficients of detergent-solubilized membrane proteins. The size estimate obtained from these experiments was compared to the rotational correlation time derived from ^{15}N relaxation experiments. The results from these studies indicate that water suppression and ^{15}N editing is necessary to obtain more sensitive, accurate measurements of translational diffusion coefficients for membrane protein-detergent complexes.

2.2 Material and methods:

2.2.1 Pf1 bacteriophage propagation and purification

The Pf1 filamentous bacteriophage was propagated by growing the phage host, namely *Pseudomonas aeruginosa,* on a non-selective agar plane and using a single colony to inoculate 50 mL of M9 minimal media (48 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.5 mM NaCl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 16 mM $^{15}NH_4Cl$, 30% (w/v) glucose, 1% (v/v) Gibco MEM Vitamin Solution) in a 250 mL Erlenmeyer flask, which was incubated overnight at 37 °C with shaking at 200 rotations per minute (rpm). The bacteria culture was allowed to grow until the optical density (OD) at 600 nm reached ~0.7-0.8. The bacteria culture was then infected with Pf1 phage at a multiplicity of infection exceeding ~50 (Cross, 1983) and allowed to sit for 10 minutes. The infected culture was then transferred into 950 mL of M9 minimal media in 2500 mL Fernbach flasks and incubated overnight at 220 rpm, 37 °C. The culture was centrifuged at 4500 g for 1 hour and 30 minutes (Beckman rotor JLA-10.500) at 4°C and sodium chloride and polyethylene glycol (PEG, MW 8000 g/mol, EMD Chemical Inc.) were added to the supernatant to a final concentration of 20 and 60 g/L, respectively. The pellet was recovered by centrifugation at 4500 g for 1 hour at 4°C and then resuspended by vortexing in ~10 mL of ultra-pure water. Remaining cell debris was removed by centrifugation at 23500 g for 45 minutes and the supernatant was then dialyzed against 1 mM EDTA, 10 mM Tris (pH 8.0) overnight. The final Pf1 phage solution was stored in the fridge or freezer (-20 °C) for future use.

2.2.2 Purification of Pf1 coat protein using phenol extraction

Phenol extraction is a basic and effective method for protein purification, since phenol can dissolve proteins and lipids leaving water soluble matter (carbohydrates, nucleic acids, etc.) in the aqueous layer (Cohn & Conant, 1926). Borate-buffered phenol (pH 8.2) was added to an equivalent volume of Pf1 phage solution prepared as described in 2.2.1 and agitated using a vortex for 1 min. The mixture was spun in a centrifuge at 3,000 g for 10 minutes and allowed to sit until phase separation had occurred. The phenol phase containing the coat protein was collected, while the aqueous layer was carefully removed by using pipette. To maximize yields the aqueous layer could be re-extracted with borate-buffered phenol if desired. The phenol phase was diluted with methanol to a ratio of 1:1 (v/v) and then dialyzed at 4°C against 4 L mixtures of 10 mM borate and methanol, first in a 3:1 (v/v) ratio, followed by a 2:1 (v/v) ratio before a final exchange against distilled water. The Pf1 coat protein was lyophilized and stored in the freezer (-20°C) for NMR sample preparation.

2.2.3 Diffusion coefficient measurement and calculation

The method implemented in this thesis to measure diffusion coefficients (D_s) is based on the bipolar gradient pulse pair longitudinal-eddy-current-delay or BPP-LED pulse scheme (shown in Figure 13 a) in which bipolar gradient pulse pairs (Wider et al., 1994) were incorporated into LED experiment (Wu et al, 1995) to suppress the effect of eddy currents that can arise from the rapid application of an

applied gradient pulse. This original BPP-LED pulse sequence was applied on the lysozyme sample (2 mM in 50 mM phosphate buffer, 10% D_2O) to get reference spectra that were used to evaluate the water suppression effect of RAW-BPP-LED pulse sequence.

Diffusion coefficient experiments were all performed at 25°C, using the RAW-BPP-LED (Figure 13 b) for the lysozyme sample (2 mM lysozyme (Sigma) in 50 mM phosphate buffer) and the ^{15}N-edited BPP-LED (Figure 13 c) for Pf1 coat protein (~ 1.5 mM) in 60 mM SDS detergent, 40 mM phosphate buffer, 90% H_2O, 10% D_2O, pH 5.5. The gradient duration δ was fixed at 2.4 ms, delay τ at 0.2 ms and the diffusion delay T at 100 ms. Transient number was set to 512 for lysozyme and 1024 for Pf1 coat protein in SDS. Diffusion measurements started with a low relative gradient strength (2% maximum gradient strength) followed by 2% increases until the signal decayed into the noise at around 80% of the maximum gradient strength.

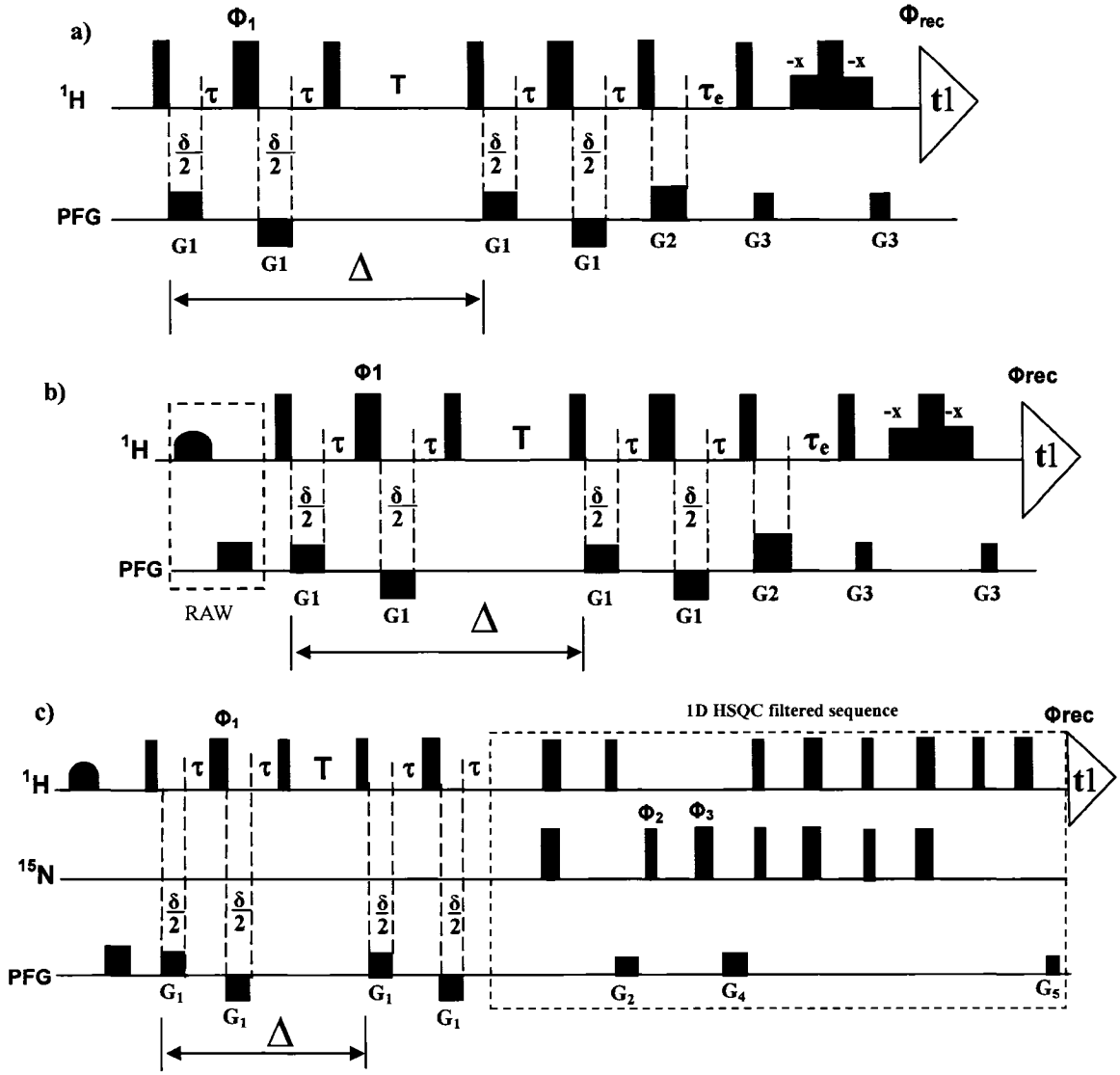


Figure 13: Pulse sequences used for diffusion coefficient measurements. a) BPP-LED (Chou, 2004). b) BPP-LED with the RAW water suppression element (RAW-BPP-LED). c) ^{15}N-edited BPP-LED pulse sequence with Rance-Kay (RK) coherence selection as the readout element. In all three pulse schemes used for diffusion measurement, the bipolar gradient pulse pairs (G_1, shown in blue) were applied along the z direction at variable strength with δ=2.0ms, τ=0.2ms and T=100ms. G_2 is a rectangular gradient with a peak amplitude of 24 G/cm applied for 1.2ms. The WATERGATE gradient G_3 in a) and b) has a peak amplitude of 36.3 G/cm and duration of 2 ms and the water selective proton pulses have a amplitude of ~ 11.64 dB and duration of 1 ms. The rectangular phase coherence selective gradient G_4 in c) has duration of 27.8 G/cm, with corresponding gradient G_5 in the Rance-Kay read-out element being 10-fold shorter. The phase cycling was:

a) Φ_1=x,y,-x,-y, Φ_{rec}=x,-x,x,-x. b) Φ_1=x,y,-x,-y, Φ_{rec}=x,-x,x,-x.

c) Φ_1=x,x,x,x,x,x,x,x,y,y,y,y,y,y,y,y,-x,-x,-x,-x,-x,-x,-x,-x,-y,-y,-y,-y,-y,-y,-y,-y;Φ_2 =x,-x; Φ_3=x,x,y,y,-x,-x,-y,-y; Φrec =x,-x,-x,x,

All data analysis was performed using Varian VNMR, Version VNMRJ 1.1. For all analyses, peaks were selected that were far from the water signal to minimize potential effects of baseline distortion. The ratio of peak intensities (I) obtained in the presence of stronger gradient strength (G) to a reference signal (I_o) at relatively low gradient strength (G_o) was calculated and plotted with respect to gradient strength (Stejskal and Tanner, 1965). A mono-exponential decay was obtained that could be fit to:

$$I = I_0 \exp[-(\gamma\delta)^2 (G-G_0)^2 (\Delta-\delta/3) D_s] \qquad [5]$$

using a non-linear curve-fitting protocol provided by xmgr (http://plasma-gate.weizmann.ac.il/Xmgr/) to extract the translational self-diffusion coefficient D_s.

The Stokes-Einstein equation (Equation 2) was used to relate D_s to the hydrodynamic radius of a spherical molecule, R_s, which is proportional to the apparent molecular weight of the molecule, M_{app}. For a solid sphere, R_s is proportional to $M_{app}^{1/3}$ (Cantor & Schimmel, 1980). Using the ratio of our experimentally determined diffusion coefficient to that of another molecule with known molecular weight and hydrated volume (e.g. lysozyme with M_{app} of 14.3 kDa and hydrated volume 17 nm^3), the apparent molecular weight and hydrated volume of the membrane protein-detergent complex were estimated using:

$$\frac{D_{s1}}{D_{s2}} = \left(\frac{M_{app2}}{M_{app1}}\right)^{\frac{1}{3}} = \left(\frac{V_2}{V_1}\right)^{\frac{1}{3}} \qquad [6]$$

2.2.4 Gradient Calibration

Since every gradient probe will has its own characteristic gradient strength for a given current value, the gradient strength needs to be carefully calibrated to obtain accurate diffusion coefficient (D_s) values. For this purpose, the BPP-LED experiment was performed at 25°C on a sample of 0.9388 mole fraction D_2O (468 μL of 99.99% D_2O and 32 μL H_2O) since its D_s is known to be 1.922 x 10^{-5} cm^2/s (Longsworth, 1960). The gradient strength of the spectrometer was calculated using the same procedure indicated in 2.2.3 with the signal intensity from water at 2% maximum gradient strength as the reference intensity (I_0). A series of experiments were run at various gradient strengths and the corresponding signal intensity ratio $\left(\frac{I}{I_o}\right)$ versus gradient strength was plotted and fit to Equation 5 using non-linear curve-fitting provided by software xmgr (Figure 14). The maximum gradient strength on z-axis was found to be 69.9 G/cm which compares well with the manufacturer's specifications of 68.9 G/cm.

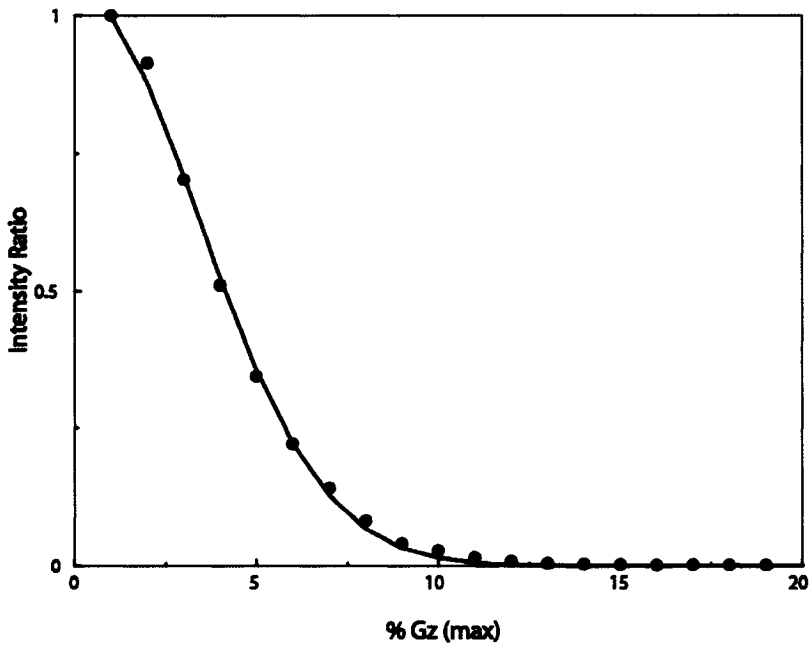


Figure 14: Plot of intensity ratio (I/I_0) vs. percent gradient strength for the water proton signal in a D_2O:H_2O sample with a D_2O mole fraction of 0.938. The black dots are the experimental points and the red curve shows the result of fitting of the data to Equation 5.

2.2.5 ^{15}N relaxation experiments

The inversion recovery version of the 2D gradient selectivity enhanced ^{1}H-^{15}N HSQC (Kay et al., 1992) was used for ^{15}N relaxation measurement and the steady-state ^{1}H - ^{15}N heteronuclear NOEs measurement (Farrow, 1994) both of which are included in the Varian Biopack suite of pulse sequences. The T_1 longitudinal relaxation delays used were 10, 300, 650, 920, 1500 and 1800 ms and the T_2 transverse relaxation delays were 10, 50, 110, 170, 230, 270, 310 ms. $^{15}N\{^{1}H\}NOE$ values were obtained from spectra with and without the 3 s relaxation time for the ^{15}N NOE to build up. The T_1, T_2 and NOE data sets were processed with NMRPipe (Delaglio *et al.*, 1995) and NMRView (Johnson & Blevins, 1994) with resolution

enhancement (90° shifted sine-bell squared function in both dimensions). Intensities of cross-peaks were obtained from peak-picking routines provided in nmrDraw (Delaglio et al., 1995). T_1 and T_2 values were extracted in a straightforward fashion by measuring the volumes of cross peaks in 2D spectra as a function of relaxation delay, T, and fitting the volumes to y=exp(-T/T_i), {i=1,2} using xmgr.

The T_1 and T_2 relaxation rates and the NOE enhancement of an amide ^{15}N spin are largely determined by dipolar coupling interactions with its directly bonded proton and by chemical shift anisotropy, giving rise to the following relationships (Abragam, 1961)

$$\frac{1}{T_1} = d^2\{J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)\} + c^2 J(\omega_N) \qquad [7]$$

$$\frac{1}{T_2} = 0.5d^2\{4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)\} + \frac{1}{6}c^2\{3J(\omega_N) + 4J(0)\} \qquad [8]$$

and

$$NOE = 1 + \left[\left(\frac{\gamma_H}{\gamma_N}\right) d^2\{6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)\}T_1\right] \qquad [9]$$

in which $d^2 = \frac{0.1\gamma_H{}^2\gamma_N{}^2h^2}{4\pi^2\left(R_{HN}\right)^6}$ and $c^2 = \frac{2\gamma_N{}^2B_0{}^2(\sigma_{\|} - \sigma_{\perp})^2}{15}$. γ_H is the proton gyromagnetic ratio, (26.7519 x 10^7 rad $T^{-1}s^{-1}$), γ_N is the ^{15}N gyromagnetic ratio, (-2.712 x 10^7 rad $T^{-1}s^{-1}$), h is Planck's constant, R_{HN} is the internuclear ^1H-^{15}N distance (1.02Å), B_0 is the magnetic field strength, (11.74 T) and $\sigma_{||}$ and $\sigma_{\perp}$ are the parallel and perpendicular components of the axially symmetric ^{15}N chemical shift tensor. The assumption of an axially symmetric chemical shift tensor has been shown to be valid

for peptide bond with $\sigma_{\parallel}$-$\sigma_{\perp}$ = -160 ppm (Hiyama et al., 1988). $J(\omega_i)$ is the spectral density function which, for a protein, depends on both the overall motion of the macromolecule as a whole and on the internal motion of the ^{1}H-^{15}N bond vector and can be expressed by using the formalism of the model free approach (Lipari & Szabo, 1982a, b).

$$J(\omega_i) = \frac{S^2\tau_c}{1+(\omega_i\tau_c)^2 + \frac{(1-S^2)\tau}{1+(\omega_i\tau)^2}} \qquad [10]$$

where S is the order parameter measuring the degree of spatial restriction of the motion and its value is between 0 and 1, τ_c is the correlation time for the overall motion, and $1/\tau=1/\tau_c+1/\tau_e$, where τ_e is an effective correlation time describing the rapid internal motions. The parameters S and τ_e described above are defined in a model-independent way that assumes that the overall tumbling of the macromolecule is isotropic.

Assuming the absence of internal motions on an intermediate time scale (μs – ms), the ^{15}N T_1 and T_2 relaxation rates are essentially independent of τ_e and hence the overall correlation time τ_c can be extracted on a residue basis according to the Equation 11 (Kay et al. 1989; Cavanagh et al., 1996)

$$\frac{T_2}{T_1} \sim \frac{d^2\{J'(\omega_H-\omega_N)+3J'(\omega_N)+6J'(\omega_H+\omega_N)\}+c^2J'(\omega_N)}{0.5d^2\{4J'(0)+J'(\omega_H-\omega_N)+3J'(\omega_N)+6J'(\omega_H)+6J'(\omega_H+\omega_N)\}+\frac{1}{6}c^2\{3J'(\omega_N)+4J'(0)\}} \qquad [11]$$

where T_1 and T_2 are the measured longitudinal and transverse relaxation time for each ^{15}N spin. $J'(\omega_i)$ is a simplified spectral density function of the form:

$$J'(\omega_i) = \frac{S^2 \tau_c}{1 + (\omega_i \tau_c)^2} \qquad [12]$$

Since Equation 11 is independent of S^2, the overall correlation time τ_c was obtained by a computer minimization of the difference between the left- and right-hand side of equation 11 for each T_2/T_1 ratio. In order to identify the residues giving rise to the relaxation data obtained, backbone amide proton and nitrogen assignments were obtained from standard triple-resonance backbone assignment spectra run under similar conditions (unpublished results).

2.3 Results and discussion

2.3.1 Water suppression in the BPP-LED

In order to evaluate the utility of the various pulse sequences for the measurement of protein diffusion coefficients, we utilized the well-characterized and commercially available protein called lysozyme as a test sample. The first sequence that was implemented was the BPP-LED sequence (Figure 13 a) described by Chou & Bax (2004) that was designed to analyze macromolecules in aqueous systems. Since the BPP-LED sequence is based on the STE-PFG approach, it is ideal for the measurement of diffusion coefficients for large macromolecules. In addition, it contains elements to suppress eddy currents (BPP and LED) and the WATERGATE water suppression sequence. However, in spite of the inclusion of the WATERGATE element, a significant water signal was obtained at low gradient strength (Figure 15 b), resulting in base-line distortion. This significantly affected the quality of the fit of the experimental data to Equation 5, making it impossible to obtain accurate diffusion coefficients on this system. To circumvent this problem, I modified the BPP-LED pulse sequence to include the randomization approach to water suppression (RAW) (Altieri & Byrd, 1995) at the beginning of the sequence (RAW-BPP-LED Figure 13 b). As shown in Figure 15, excellent water signal suppression was obtained, allowing the protein spectrum to be obtained with high sensitivity. Even at the every low PFG strengths, very little detectable signal from water was observed and consequently baseline distortion was no longer a problem.

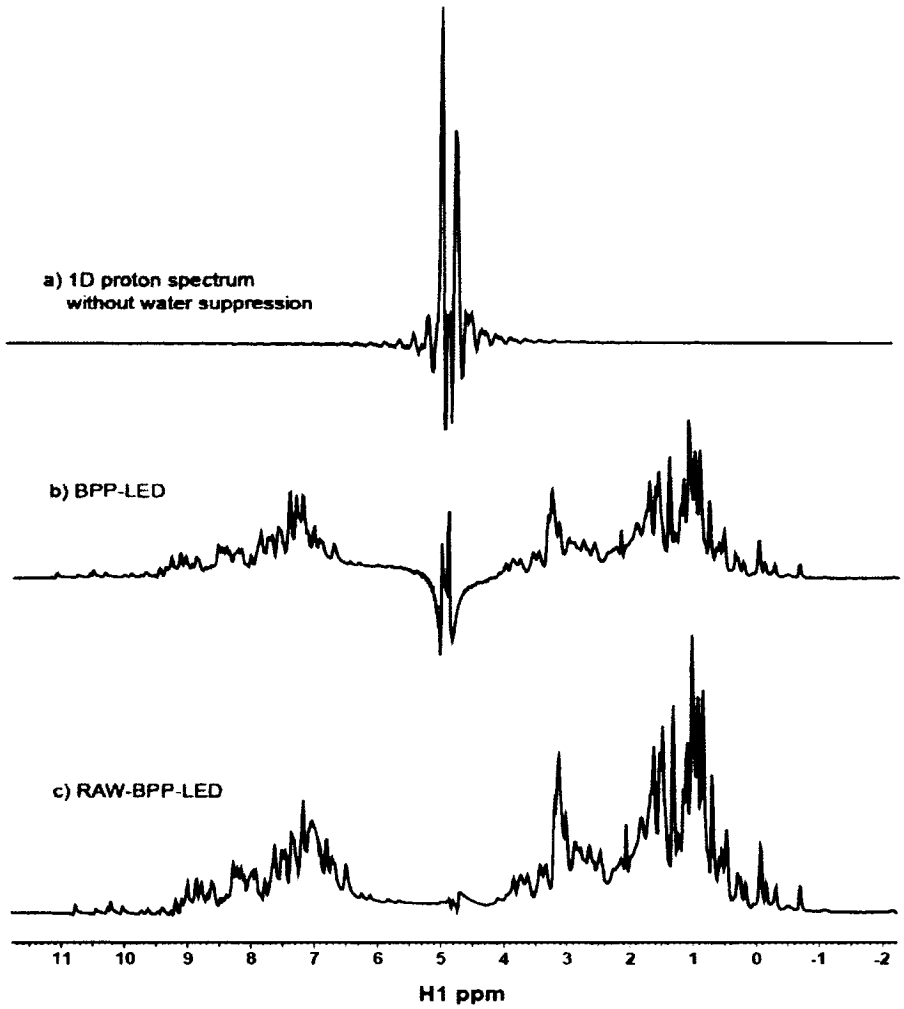


Figure 15: 1D ^{1}H NMR spectra for lysozyme acquired using either a) a single 90° proton pulse with no water suppression, b) the BPP-LED pulse sequence from Figure 13 a or c) the RAW-BPP-LED sequence from Figure 13 b. All spectra were obtained at 25 °C on 2 mM lysozyme, 50 mM sodium phosphate, pH 6.0, 10% D_2O. The scale of top the spectrum has been reduced 30 times relative to the scale of the other two spectra. A digital filter of bandwidth 600 was applied to all of the original FIDs in VNMRJ to suppress the solvent peak.

The series of spectra highlighting both aliphatic and amide regions of the lysozyme spectrum that were obtained at different gradient strengths using the RAW-BPP-LED sequence is shown in Figure 16. As expected, the signal intensity was observed to decrease with increasing gradient strength, illustrating the impact of

translational diffusion on the recovery of the gradient-dephased signal.

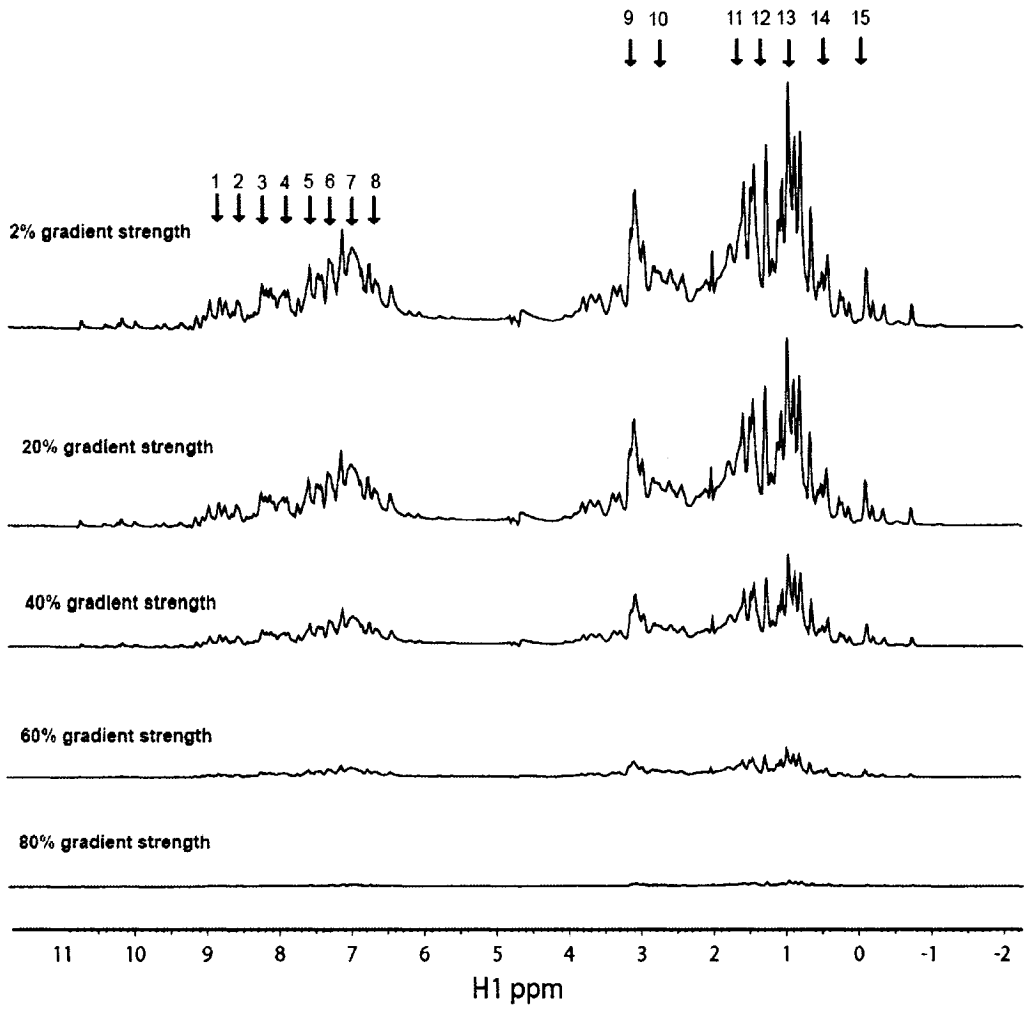


Figure 16: Selected spectra from the RAW-BPP-LED experiment run on 2 mM lysozyme in 50 mM phosphate buffer with 512 transients, pH 6.0, 10% D_2O, at 25°C. 15 peaks selected from both the amide and aliphatic regions of the spectrum for diffusion coefficient measurement were indicated by blue arrows.

For this series of spectra, 15 peaks were selected from both the amide and aliphatic regions of the spectrum for diffusion coefficient measurement. The intensities of these peaks were normalized by their respective intensities in the 2%

gradient strength reference spectrum and then fit to Equation 5. Figure 17 shows representative plots of I/I_0 versus gradient strength fit to this equation for peak 5 from the amide region (Figure 17 a) and peak 13 from the aliphatic region (Figure 17 b) of the 1D spectrum.

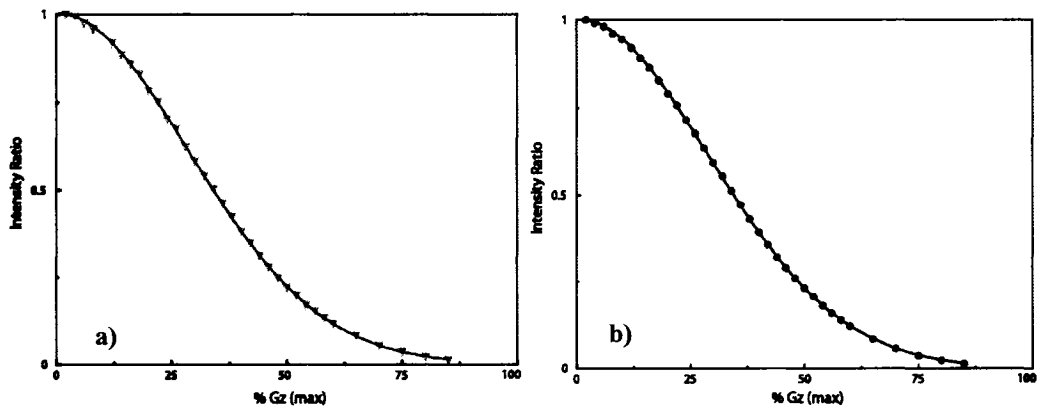


Figure 17: Plot of normalized peak intensity ratios (I/I_0) versus gradient strength (% G_{Zmax}) for a peak from a) the amide region, or b) the aliphatic region of the spectra of 2 mM lysozyme in 50 mM sodium phosphate pH 6.0, 10% D_2O. Brown and violet symbols denote experimental data points; the line shows the results of the curve fitting of the experimental data to Equation 5. The fitting values of D_s obtained from the fitting procedure in a) and b) are 1.04 (±0.02) x 10^{-6} cm^2/s and 1.01(±0.03)x 10^{-6} cm^2/s respectively.

The average diffusion coefficient determined with all 15 peaks in the lysozyme spectrum was found to be 1.03 (± 0.03) x 10^{-6} cm^2/s. As show in Table 1, this value is in excellent agreement with previous measurements for lysozyme obtained by using the LED pulse sequence (Altieri et al., 1995) and Rayleigh interferometry (Longsworth, 1960). This value is also consistent with the known molecular weight for lysozyme of 14.3 kDa and hydrated volume of 1.7 nm^3.

Our results with the BPP-LED pulse sequence indicate that measurement of accurate diffusion constant values on our spectrometer requires the addition of the RAW water suppression element. This may reflect the fact that there is only a z-axis gradient system available on this spectrometer, leading to a reduction in water suppression efficiency relative to that obtained with a triple axis gradient system. Since our results showed that the RAW-BPP-LED sequence could be used to accurately measure the diffusion coefficient for water soluble proteins (e.g., lysozyme), the next step was to evaluate its suitability for measurements on membrane protein systems.

Table 1: Diffusion coefficient values for lysozyme at 25°C.

Method	D_s (cm^2/s)	Reference
Rayleigh Interferometry	1.06×10^{-6}	Longsworth, 1960
LED	1.08×10^{-6}	Byrd et al, 1995
RAW-BPP-LED	$1.03 \pm 0.03 \times 10^{-6}$	Current study

2.3.2 ^{15}N-edited BPP-LED for diffusion coefficient measurements of membrane proteins

In order to test whether the RAW-BPP-LED sequence could be used to measure the diffusion coefficient of detergent-solubilized membrane proteins, we ran this experiment on the Pf1 coat protein dissolved in SDS micelles. However, as shown in Figure 18, we found that the signals from detergent protons were much more

intense than those from proteins due to the large excess of detergent. While this was not surprising since the concentration of detergent utilized for solubilizing a membrane protein is usually much higher than that of the protein (40 fold excess in this sample), this made it difficult to observe peptide amide peaks in the spectrum. Instead, the dominant peaks in this spectrum originated mainly from methyl and methylene protons from free SDS molecules that appear to be in rapid exchange with micelle-bound species, while the proton signal from protein was more than 1000-fold lower in intensity. For this reason it was not possible to use the modified BPP-LED experiment to measure the diffusion coefficient for the coat protein-micelle complex.

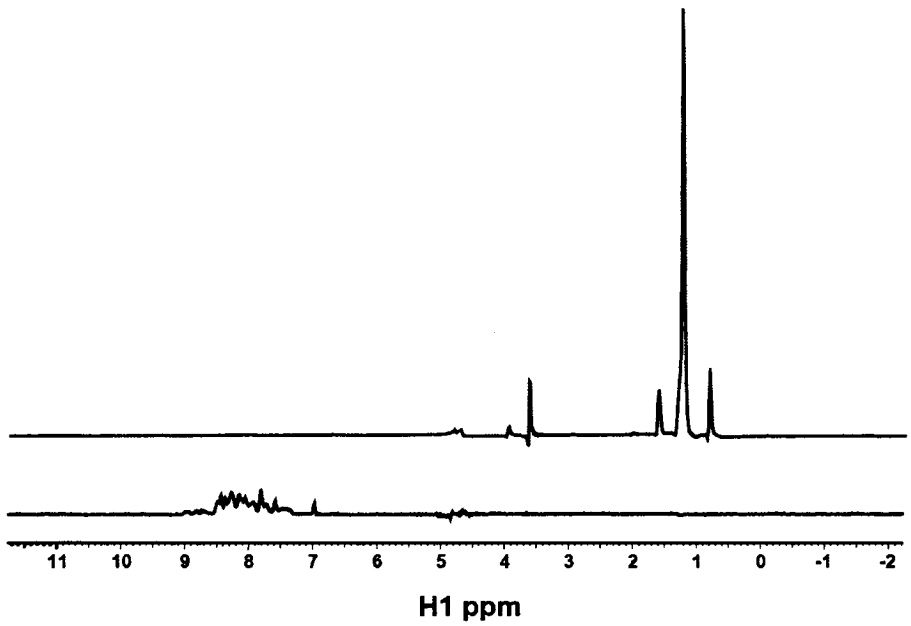


Figure 18: Comparison of 1D ^{1}H NMR spectra for 1.5 mM Pf1 coat protein in 60 mM SDS, 40mM phosphate buffer pH 5.5 at 25°C obtained using RAW-BPP-LED (top) and the ^{15}N-edited BPP-LED sequence (bottom). The scale on the top spectrum has been reduced 100 times relative to the scale of the bottom spectrum. Both spectra were recorded with the following parameters: a sweep width of 7018 Hz with 898 complex points and 256 transients (top spectrum) or 512 (bottom spectrum).

In order to suppress the huge detergent signals and to measure the size of protein-detergent complex, I decided to edit the signal obtained after RAW-BPP-LED sequence to select for magnetization originating from protons attached to ^{15}N. This was done by appending a 1D ^{1}H - ^{15}N HSQC read-out element to the RAW-BPP-LED (Figure 13 c). This ^{15}N-edited BPP-LED diffusion experiment was applied to a sample of uniformly ^{15}N-labeled Pf1 coat protein (~ 1.5 mM) in 60mM SDS micelle and 40 mM phosphate buffer at 25°C. As shown in Figure 18 b, the primary signals shown in the spectrum originate from amide protons, with no detectable signal from the SDS, making it possible to obtain an undistorted spectrum of the protein alone.

To test the utility of this sequence in the measurement of diffusion coefficients for protein-detergent complexes, 1D spectra of ^{15}N-labeled Pf1 coat protein were acquired at a range of gradient strengths (Figure 19). As shown in Figure 20, normalized peak intensities were plotted as a function of gradient strength for 6 peaks. This gave rise to a monoexponential decay that gave rise to an excellent fit to Equation 5. Diffusion coefficients were extracted from these fits, giving rise to an average D_s of $6.44 \pm 0.11 \times 10^{-7}$ cm^2/s.

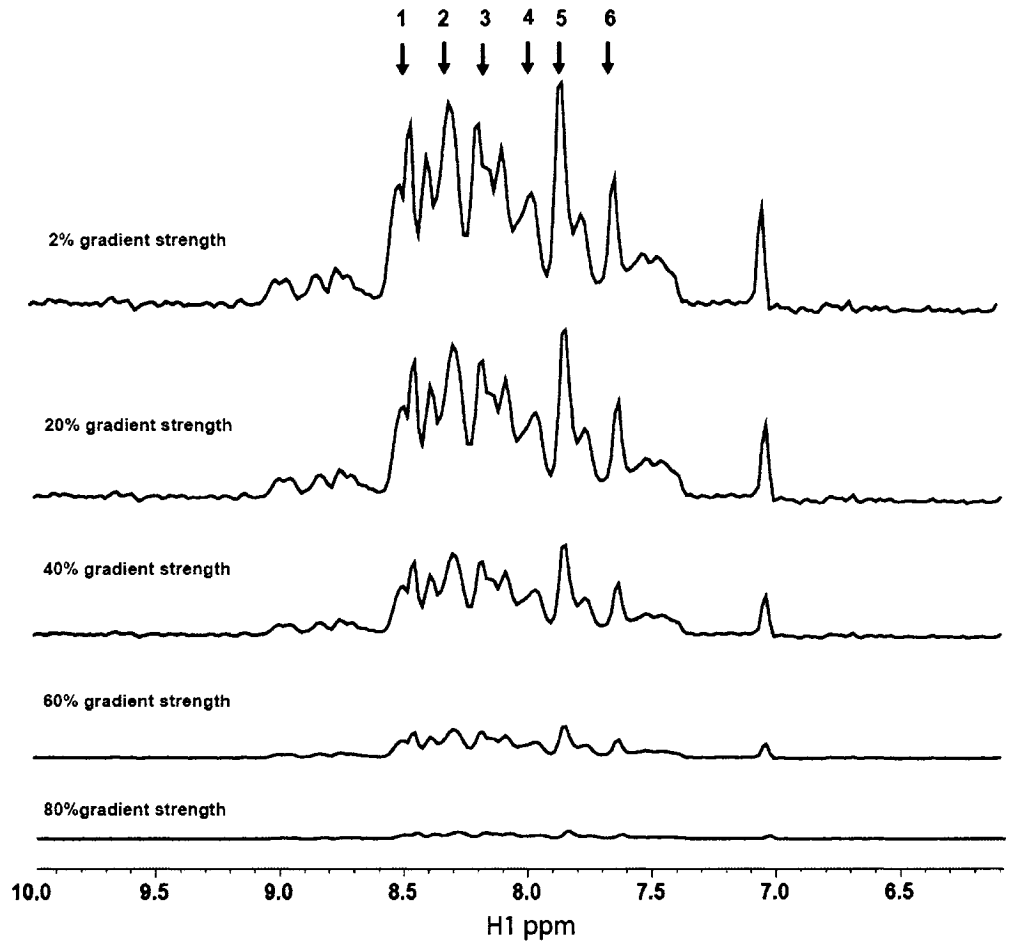


Figure 19: Series of spectra from the ^{15}N-edited BPP-LED experiment run on 1.5 mM ^{15}N-labeled Pf1 coat protein in 60 mM SDS, 40 mM sodium phosphate pH 5.5 at 25°C. Spectra were recorded by using the ^{15}N-edited BPP-LED sequence (Figure 13 c), with a sweep width of 7018 Hz and 898 points in the proton dimension with 1024 transients. Peaks highlighted with blue arrows were used to extract D_s for the protein-detergent complex using the procedure described in section 2.2.3.

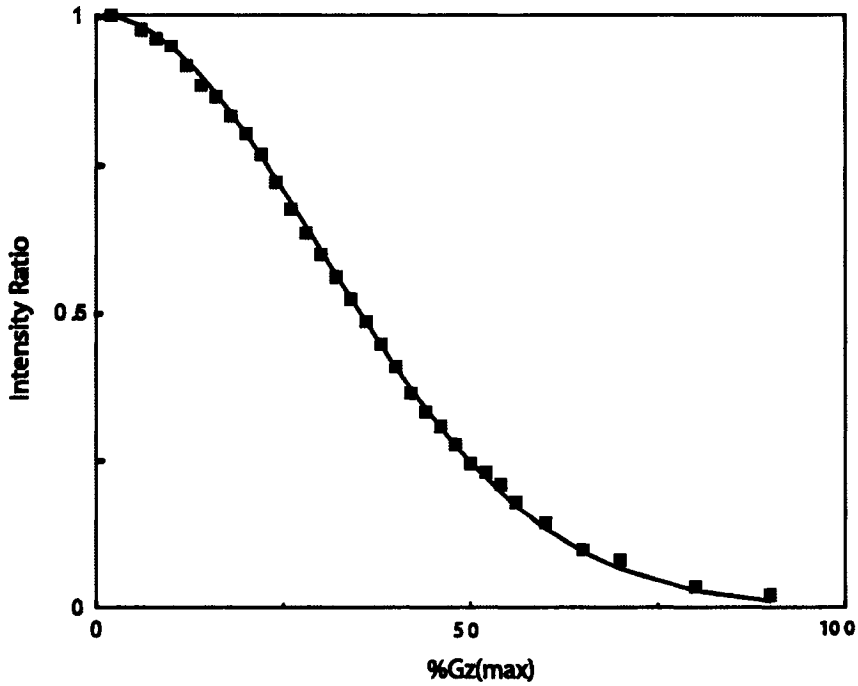


Figure 20: Normalized signal intensity (I/I_o) vs. gradient strength (% G_{zmax}) obtained from peak 5 in the series of spectra shown in Figure 19. The line shows the best fit of the data to Equation 5 to yield $D_s = 6.35 \times 10^{-7}$ cm^2/s.

2.3.3 Estimation of Pf1-micelle complex size

Using Equation 6, the size of the protein-detergent complex was estimated from our diffusion coefficient measurement and found to be approximately 58 kDa. For the purpose of comparison we also ran the RAW-BPP-LED experiment on a protein-free SDS sample under the same conditions. The molecular weight estimate obtained from the resulting diffusion coefficient (8.79×10^{-7} cm^2/s) was ~ 22 kDa which corresponds to a micelle composed of approximately 80 detergent molecules. This value is higher than that measured for an SDS micelle in absence of salt (N=68, Quina, 1995) due to the presence of the 40 mM sodium phosphate buffer in the

sample. However, SDS is an ionic surfactant, and the addition of salt is known to increase the micelle size since this can shield of repulsive electrostatic interactions between SDS head groups (Bezzobotnov, 1988; Missel, 1989; Quina, 1995). This increase in micelle aggregation number has also been observed in experiments using light scattering (Hayashi, 1980), neutron scattering (Bezzobotnov, 1988) and fluorescence (Bales, 1995) which generally showed an increase in aggregation number from ~60 to ~90 when the counterion concentration increased from 20 to 140 mM. Therefore, our results from the diffusion coefficient measurements seem to provide a reasonable estimate for the size of the SDS micelle in the absence of Pf1 coat protein.

Given that the protein:detergent molar ratio in this study was ~1:40 and that one SDS micelle contains at least 80 detergent molecules, there must have been an average of at least 2 coat protein molecules associated with each micelle. However, since the coat protein is 6 kDa, this model would suggest a molecular weight of 34 kDa, a result that is 24 kDa smaller than estimates obtained from our diffusion measurements. This implies that the incorporation of coat protein has a non-linear additive effect on the apparent molecular weight of the coat protein-detergent complex. The origin of this non-linear effect is not yet known, although one possibility is that incorporation of Pf1 coat protein into the SDS micelle increases the number of detergent molecules that could be associated with it. Since the protein:detergent ratio is ~1:40, the increase in micelle size means that a proportional number of Pf1 protein molecules would have to be associated with each micelle.

Based on a 58 kDa estimate for the Pf1 coat protein-SDS micelle, each complex likely contains approximately 110 detergent molecules and ~4 coat proteins. This increase in SDS aggregation number could reflect the ability of the N-terminal poly-Lys tail to reduce repulsive interactions between SDS sulfate groups, a factor which would be predicted to increase the size of the micelle. As a test of this hypothesis it should be possible to repeat these measurements in the future for a series of Pf1 samples at a range of concentrations. If incorporation of Pf1 increases the aggregation number of the SDS micelle then the size of the complex should decrease as with decreasing Pf1 concentration. However, it should also be noted that the incorporation of the Pf1 coat protein could affect the hydrodynamic properties of the SDS micelle to induce a non-spherical shape, or a change in the amount of water associated with the complex. Also, since the measured diffusion coefficient D_s decreases linearly with the concentration of detergent used at low volume fraction (Tokuyama & Oppenheim, 1994; Chou & Bax, 2004), it may be useful to do additional D_s measurements at various concentrations in the future in order to extrapolate the single particle self-diffusion coefficients for SDS micelles with and without protein at infinite dilution, and investigate how the incorporation of protein affects the effective hydration volume and apparent molecule weight of detergent micelle.

A non-linear additive effect of membrane protein incorporation on the molecular weight of the membrane protein-detergent complex has been observed in other systems and appears to depend on the detergent being used. For example, the diffusion coefficients of the transmembrane protein diacylglycerol kinase (DAGK)

(13 kDa) were measured in SDS, β-octyl glucoside (OG) and DHPC micelles and found to be about 53%, 27% and 64% smaller than that of corresponding protein-free micelles respectively (Vinogradova, 1998). These observations, in combination with our results for SDS showing a 28% decrease in D_s upon incorporation of Pf1 coat protein –detergent micelle size suggest that the size of protein-detergent complex will depend on the detergent being used, the nature of the membrane protein interaction with the micelle and the solution conditions (e.g. salt, buffer concentrations, etc.). The difficulty in predicting the molecular weight of a protein-detergent micelle complex a priori therefore illustrates the importance of applying methods such as the ^{15}N-edited BPP-LED sequence to estimate complex size.

2.3.4 Investigation of Pf1 dynamics in the SDS micelle

Our measurements on the translational diffusion coefficient of the Pf1 coat protein-SDS complex indicate that the micelle-protein complex in this case should be large enough to give rise to significant loss of the NMR signal from rapid transverse relaxation. Assuming that the protein-detergent complex is spherical and the viscosity of sample solution is the same as water at 25°C, its rotational correlation time can be estimated from D_s using Equation 1, giving rise to an estimate of ~32 ns. This suggests that if the coat protein were rigidly fixed to the micelle with no rotational motion within the micelle allowed, then its NMR spectrum should be very broad. Nonetheless the 2D ^{1}H-^{15}N HSQC spectrum of the Pf1 coat protein in SDS exhibits reasonable resolution and sensitivity (Figure 21), indicating that the actual rotational

correlation time may be lower than that suggested by translational diffusion measurements due to mobility of the bound protein relative to SDS micelle.

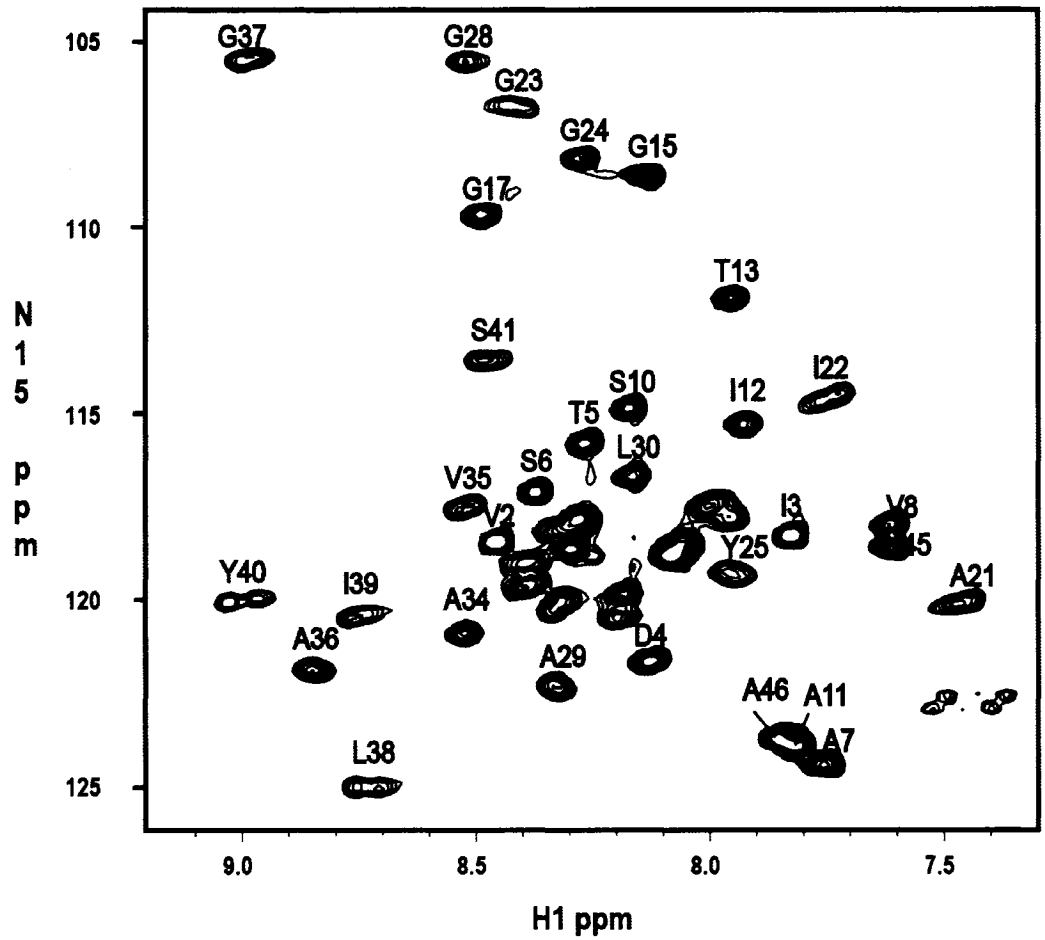


Figure 21: 2D ^{1}H-^{15}N HSQC spectrum for 1.5 mM Pf1 coat protein in 60mM SDS micelle and 40 mM phosphate buffer, pH 5.5 at 25°C. The spectrum was collected with 32 transients, and 898 and 100 real points over a spectral width of 7017.54 Hz and 1520.65 Hz in t_2 and t_1 respectively. Data were processed with a shifted squared sine-bell apodization function and zero-filled to 1024 (^{1}H) and 256 (^{15}N) points. Note: Some of peaks were not assigned because of peak overlap.

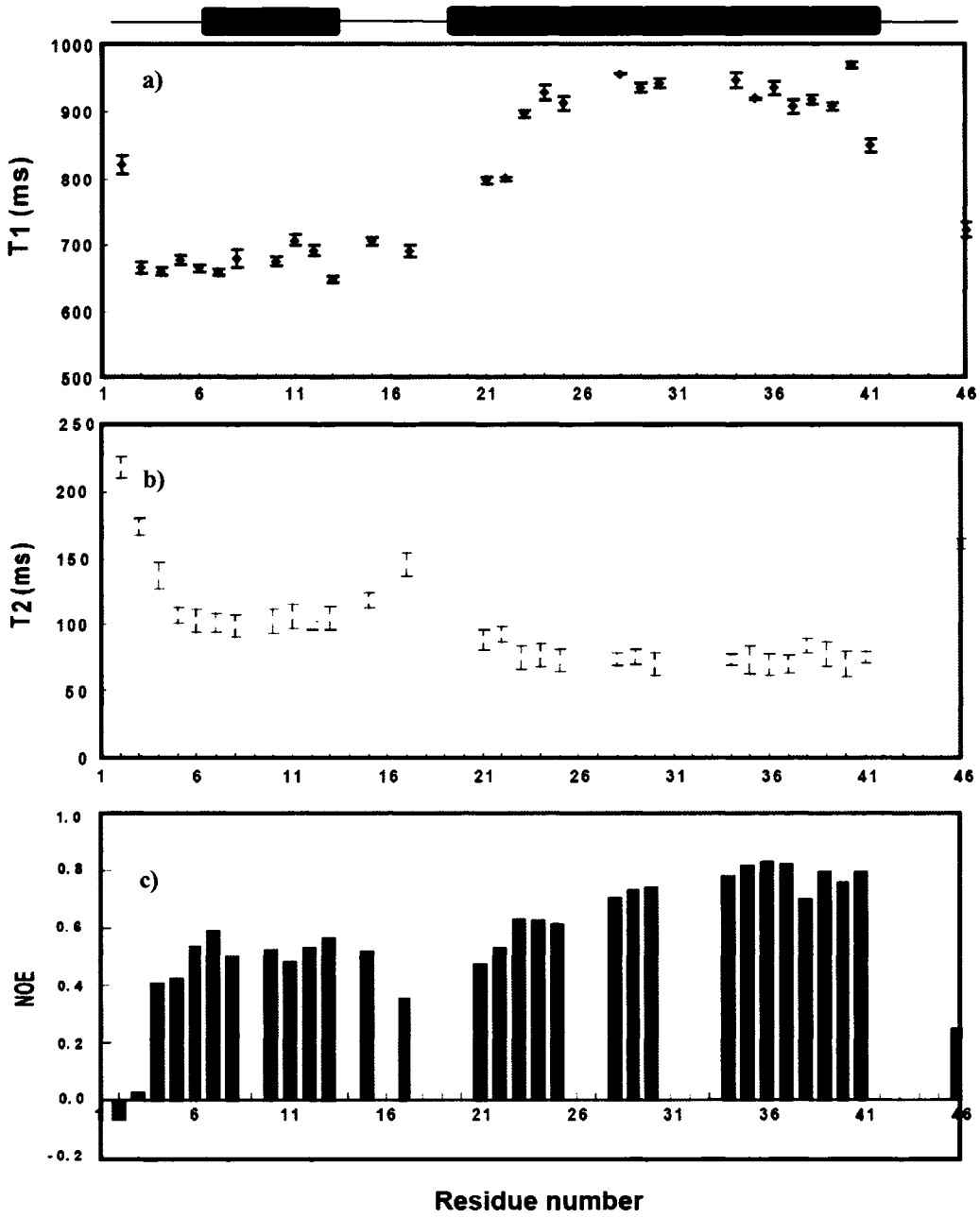


Figure 22: ^{15}N relaxation data for Pf1 coat protein in 60 mM SDS micelles, 40 mM phosphate buffer, 10% D_2O, pH 5.5, at 25 °C plotted with respect to Pf1 residue number. In some cases data is not shown since residues with a high degree of overlap in the HSQC were difficult to accurately assign and extract accurate intensities in the spectra. a) Longitudinal relaxation time T_1. b) Transverse relaxation time T_2. c) Heteronuclear NOE with an average estimate error of ± 0.05. The location of secondary structure elements previously determined for Pf1 coat protein in SDS micelles (Schiksnis, 1987; Lee, 2003) is shown schematically at the top.

In order to determine the effective rotational correlation time for Pf1 coat protein and investigate the dynamic properties of this sample, ^{15}N T_1 and T_2, and $^{15}N\{^1H\}$ NOE relaxation measurements were performed. The relaxation data shown in Figure 22 indicates that the residues located on the helices had similar T_1 and T_2 values within the helices. In addition, residues in the helical regions exhibited larger positive heteronuclear NOEs, while residues within both terminal and loop regions had smaller NOE values. Overall this data shows that residues in the N- and C-terminal regions, as well as in the loop connecting the helices, are more mobile than residues in the helices. Moreover, the sequence corresponding to the amphiphilic N-terminal helix had a lower average ssNOE (0.53 ± 0.04) than that for the than the transmembrane helix at the C-terminal end (0.71 ± 0.11). This amphipathic helix also had a longer average T_2 (102.76 ms ± 2.72) and shorter average T_1 (675.41 ms ± 20.11) than the transmembrane helix (76.46 ± 6.62 ms T_2 and 908.88 ± 30.75 ms T_1). Overall, these results suggest that the amphipathic helix is much more mobile than the transmembrane helix. This is consistent with previous relaxation studies that suggested that the amphipathic helix was only transiently associated with the SDS micelle surface while the TM helix was an integral part of the micelle (Lee & Opella 2003).

From the relaxation data the average correlation time of the C-terminal helix was found to be approximately 14 ± 1 ns. This would correspond to a molecular weight for the Pf1 coat protein-detergent complex on the order of ~20 kDa. The fact that this correlation time is much shorter than the one obtained from the diffusion

coefficient measurements (32 ns), indicates that the SDS micelle must fluid enough to permit extensive rotation of Pf1 coat protein within the confines of the micelle. This suggests that it will be possible to study larger molecular weight membrane protein-detergent complexes than would normally be possible for globular water-soluble proteins as long as there is sufficient mobility of the membrane protein within the detergent micelle. This has also been observed for other membrane proteins, such as subunit c of the *E. coli* F_1F_0 ATP in LOPG detergent micelles (1-oleoyl-2-hydroxy-sn-glycero-3-[phosphor-RAC-(1-glycerol)]) (Krueger-Koplin & Girvin, 2004). However, it still remains to be determined whether this extent of mobility will be observed for membrane proteins with a larger number of TM segments.

2.4 Conclusions

The incorporation of the RAW water suppression element into the BPP-LED diffusion coefficient measurement experiment was found to effectively suppress the solvent water signal to obtain accurate measurement of diffusion coefficients of water soluble proteins. However, to measure the diffusion coefficient for membrane protein-detergent complexes it was necessary to implement the ^{15}N-edited BPP-LED pulse sequence. This experiment provided a fast and convenient method to determine the size of a membrane protein-detergent complex under NMR conditions. This method was applied to the Pf1 coat protein-SDS micelle (60 mM) complex and yielded a diffusion coefficient that suggested a size for the detergent-protein complex

on the order of 58 kDa. This information, in combination with the relaxation data, was used to evaluate the impact of Pf1 mobility in the SDS micelle on spectral resolution, which was generally found to be significant in this detergent. In the future, the experiments developed and implemented here will help to investigate a range of detergents for suitability in structural studies of membrane proteins by solution NMR.

Chapter 3

Transmembrane helix interactions in detergent studied by solution NMR

3.1 Introduction

One of the important requirements for solubilization reagents utilized for membrane protein structure determination by solution NMR is that they should allow the protein to fold into its native form. In order to investigate the ability of a detergent to promote membrane protein tertiary structure, we need to develop a system that can be used to monitor the impact of a particular detergent on the interaction between TM helices. Therefore, the main objective in this part of the thesis is to establish and characterize a model system that can be used to test the ability of different detergents to promote the interactions between TM helices.

3.1.1 The M13 bacteriophage major coat protein (MCP) TM segment as a model system

An ideal model system for the study of detergent effects on TM helix interactions should be easy to produce and be amenable to rapid screening by solution NMR. For this purpose we have chosen to study the single TM segment of major coat protein (MCP) from bacteriophage M13. This protein is composed of 50 amino acids (~5.2 kDa, sequence shown in Figure 23), and its 3D structures in SDS and DPC

detergent micelles have been determined by using solution NMR (Papavoine *et al.*, 1998). These studies showed that the structure is composed of a single TM helix (Y25-F45) that is embedded in the hydrophobic region of the micelle, while a second, amphipathic, helix (P6-E20) interacts with the surface of the micelle.

1 11 21 31 41 50

AEGDDPAKAAFRSLQASATEYIGYAWAMVVVIVGATIGIKLFKKFTSKAS

Figure 23: The primary amino acid sequence of the M13 bacteriophage major coat protein (MCP). The single underlined sequence denotes the N-terminal amphipathic helix. The double underline indicates the α-helical transmembrane domain. The residues highlighted in red are those found to be important for homodimerization as identified by site-directed mutagenesis (Melnyk et al., 2002).

It has been shown using SDS-PAGE analyses and fluorescence resonance energy transfer experiments, excimer formation studies and analytical ultracentrifugation experiments that MCP can self-associate into dimers via specific interactions between the TM segments in detergent micelles (Melnyk et al., 2002), a finding that was confirmed in lipid bilayers in vivo (Haigh et al., 1998). In addition, mutagenesis studies demonstrated that the self-association of MCP in membrane environments is sensitive to mutations located on one face of the TM α-helix (Deber, 1993; Li, 1993; Williams, 1995). These results led to the identification of a putative dimeric interface within the core of the TM domain (Figure 23). Based on this information, it was proposed that the bulky amino acid residues (V, I) form a ridge along the length of the TM helix that fits into a complementary groove formed by the smaller amino acid residues (G, A). This type of hydrophobic interaction has been

termed "knobs into holes" and is a common motif underlining the interaction of TM helices within multi-spanning helical membrane proteins.

The importance of the TM domain for the dimerization of MCP has also been illustrated with synthetic peptides containing only the TM segment from MCP (MCP_{TM}) (Melnyk et al., 2002). SDS-PAGE analysis, a useful method to characterize the molecular weight of proteins, as well as membrane protein complexes, showed a mobility for MCP_{TM} that was consistent with a dimeric complex, as confirmed by fluorescence-based studies (Melnyk et al., 2002). Since these studies indicated that MCP_{TM} can form homodimers in detergents at concentrations that should be amenable for solution NMR studies, we decided to use MCP_{TM} as a model system to investigate the effect of various detergents on the monomer-dimer equilibrium state.

Since an over-riding interest in the Goto lab is the development of new detergents for solution NMR studies, we are interested in designing an assay that could rapidly identify the oligomerization state of MCP_{TM} under solution NMR conditions. For this reason we chose to use a basic 2D NMR experiment that is widely considered to provide a 'fingerprint' spectrum for proteins, called the ^{1}H-^{15}N heteronuclear single quantum correlation spectroscopy (HSQC) experiment.

3.1.2 Features of the ^{1}H-^{15}N HSQC

In a ^{1}H-^{15}N HSQC experiment each amino acid residue (except proline) gives rise to one peak at the chemical shift of the amide proton (x-axis) and amide ^{15}N atom (y-axis). Since the amide proton chemical shift is very sensitive to its local chemical

environment, any change in the conformational and/or oligomerization state in a protein will give rise to changes in peak positions (chemical shifts) in the HSQC spectrum. Therefore, we anticipated that this experiment would be a useful tool for the characterization of MCP_{TM} oligomerization state in different detergent environments.

The HSQC experiment also has the benefit that it can be used to assess whether the MCP_{TM} peptide is folded, since folded proteins display a broad distribution of amide proton chemical shifts (> 1 ppm), while 'unstructured' random coil configuration exhibit a narrow range of amide proton shifts (~ 0.5 – 1.0 ppm). Therefore, the 2D ^{1}H-^{15}N HSQC spectra acquired in this work will also provide some insight into the confirmation of MCP_{TM} in different detergent environments. Moreover, since the HSQC is one of the basic building blocks in a number of triple resonance multidimensional NMR experiments routinely applied in structure determination, the quality of the HSQC will also provide useful insight into the suitability of various detergent systems for solution NMR studies of membrane proteins. In order to acquire a ^{1}H-^{15}N HSQC NMR spectrum it will be necessary to synthesize the MCP_{TM} peptide with ^{15}N-label incorporated into the peptide backbone. With the ^{15}N label it should be possible to use this HSQC experiment as a tool for the rapid evaluation of MCP_{TM} oligomerization state, folding and spectral quality.

3.1.3 Site-directed spin labelling (SDSL) as a probe for TM helix-helix interactions

Although 2D HSQC spectroscopy can be used to monitor changes in the local chemical environment of the protein backbone, any changes in chemical shifts could arise either from a change in the conformational or oligomerization state of the protein. Therefore, in order to determine whether a difference in amide chemical shifts is a result of a change in oligomerization, I developed an assay for MCP_{TM} that combines site-directed spin labelling (SDSL) with HSQC spectroscopy. This technique takes advantage of the fact that the unpaired electron of a spin label produces local fluctuations in the magnetic field that can influence the corresponding spin-lattice (T_1) and spin-spin (T_2) relaxation times of protons in a distance-dependent manner. For commonly used spin labels, such as (1-Oxyl-2,2,5,5-tetramethyl-Δ^3-pyrrolin-3-yl) methyl methanethiosulfonate (MTSL; Toronto Research Chemicals), the result is a substantial broadening of the resonance for protons within ca 15 Å of spin label (Schmidt, 1984; Kosen et al., 1986). We decided to use this phenomenon to look for evidence of *inter*molecular interactions between MCP_{TM} peptides across the dimer which should be detected as paramagnetic broadening of peaks in the HSQC spectrum.

This strategy in the assay developed here required the use of an MCP_{TM} with no ^{15}N-label that instead contained a non-native cysteine residue that could be derivatized with a sulfhydril-reactive spin label. When mixed with ^{15}N-labeled MCP_{TM}, an ^{1}H-^{15}N HSQC spectrum can be recorded to ensure that only peaks from the ^{15}N-labeled (non-spin-labeled) peptide will be monitored. Paramagnetic broadening of these peaks can only occur when a spin-labeled peptide is close enough

to an ^{15}N-labeled peptide to effect the relaxation properties of the ^{15}N-labeled peptide (Figure 24). In order to observe a residue-specific paramagnetic effect on the HSQC spectrum as illustrated in Figure 24, not only must the ^{15}N-labeled peptide occupy the same micelle as the spin-labeled peptide, but they must also interact with each other in a specific configuration. If there is no interaction between helices, then broadening will either not occur, or affect all ^{15}N-labeled residues to a similar extent. Therefore, the extent of broadening in the HSQC spectrum should provide a straightforward measure of the ability of detergent micelles to promote TM helix-helix interactions in the MCP_{TM} peptide.

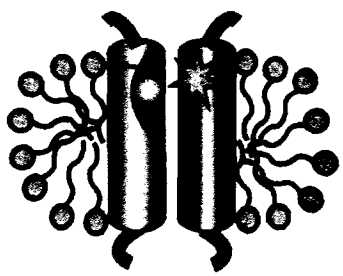


Figure 24: Schematic diagram of ^{15}N- (red) and spin-labeled MCP_{TM} (grey) heterodimer embedded in a detergent micelle. The spin label will cause paramagnetic broadening of protons (grey patch) within ca 15 Å of the spin label.

In addition to providing an indication of dimer formation by MCP_{TM}, the extent of paramagnetic broadening can also be used to quantitatively measure the distance between the spin label and the affected proton resonances. This can be done through application of the Solomon-Bloembergen equations (Solomon &

Bloembergen, 1956; Kosen, 1989), which describe the enhancement of proton longitudinal and transverse relaxation rates ($1/T_1$ and $1/T_2$, respectively) according to:

$$\Delta(1/T_1) = \Delta R_1 = \frac{2K}{r^6\left(\frac{3\tau_c}{1+\omega_H^2\tau_c^2}\right)} \qquad [13]$$

$$\Delta(1/T_2) = \Delta R_2 = \frac{2K}{r^6\left(4\tau_c + \frac{3\tau_c}{1+\omega_H^2\tau_c^2}\right)} \qquad [14]$$

where K is a nitroxide radical specific constant, (1.23×10^{-32} cm^6s^{-2}), r is the distance between the electron and the proton, τ_c is the correlation time for the electron-proton vector, and ω_H is the Larmar frequency of the proton. As described by these two equations, the paramagnetic enhancement of relaxation is inversely proportion to the sixth-power of the distance between the spin label free electron and the proton. Since the linewidth of peaks in the NMR spectrum is directly proportional to the transverse relaxation rate ($1/T_2$), this paramagnetic broadening effect can be most straightforwardly estimated by comparing peak intensities in a 2D ^1H-^{15}N HSQC spectrum of the spin-labeled/^{15}N-labeled peptide mix before and after reduction of the spin label. Therefore, using this approach it should be possible to gain new structural insights into the nature of the MCP_{TM} dimer as well as obtain additional information on the nature of the detergent-peptide interaction.

3.1.4 Circular dichroism (CD) spectroscopy – a probe of protein secondary structure

In order to investigate whether there are any changes in the secondary structure of MCP_{TM} peptide in different detergent environments, circular dichroism (CD) spectroscopy was also used in this study. CD spectroscopy is an excellent biophysical method to evaluate the secondary structure content of peptides and proteins. This form of spectroscopy relies on the interactions of right- and left- handed circularly polarized light and asymmetrical chromophores. For proteins and peptides the peptide bond is the major CD chromophore in the far-UV region of the spectrum. When arranged in a specific conformation these chromophores give rise to a characteristic CD spectrum that can be used to identify the secondary structures contained in this structure.

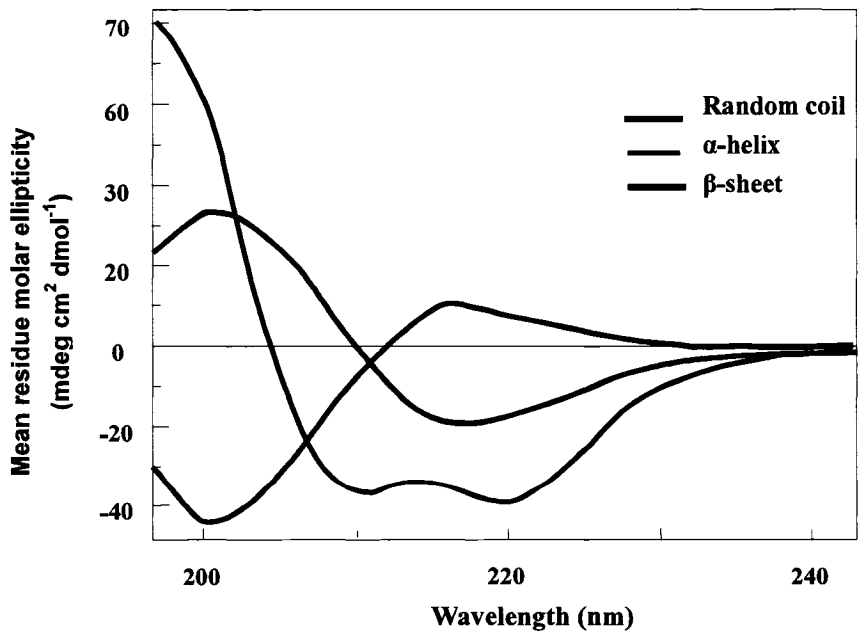


Figure 25: Characteristic CD spectra for random coil (black), α-helix (red) and β-sheet (blue) structures.

As shown in Figure 25, the CD spectrum for an α-helix (red) has two characteristic minima found at approximately 222 nm and 208 nm, while a typical β-sheet conformation (blue) gives rise to a spectrum with a minimum at ~215 nm and a maximum at ~198 nm (Sreerama, 2003; Sreerama, 2004). A typical random coil conformation also has a characteristic CD spectrum (black), showing a maximum at ~215 nm and a minimum at ~202 nm. These different and readily distinguishable CD spectra for different types of protein secondary structures make this an ideal tool for the study of the effect of different detergents on the structure of MCP_{TM}.

3.1.5 Objectives

In this work, a spin labelling strategy was used in combination with 2D ^{1}H-^{15}N HSQC spectroscopy to investigate the oligomerization state of the MCP_{TM} model system in different detergent environments. The results from this study showed that both denaturing detergents (sodium dodecyl sulfate: SDS) and non-denaturing detergents (β-octyl glucoside: OG) are able to support the dimerization of MCP_{TM} under NMR conditions. The pattern of resonance broadening observed in OG is consistent with a parallel arrangement for MCP_{TM} in a dimeric configuration. Overall, the method developed here should be useful for the investigation of interactions between TM helices in different detergent micelles in the future.

3.2 Material and methods

3.2.1 MCP_{TM} synthesis

MCP_{TM} and its mutants were synthesized using standard Fmoc chemistry on a Perseptive Biosystems Pioneer peptide synthesizer as previously described (Melnyk et al., 2002). The syntheses employed the Pioneer's standard (45 min) cycle. The *N*-[(dimethylamino)-1*H*-1,2,3-triazolo[4,5-b]pyridine-1-ylmethylene]-*N*-methylmetha naminium hexafluorophosphate *N*-oxide (HATU, PerSeptive Biosystems) / *N,N*-diisopropylethylamine (DIEA, Sigma-Aldrich) activator pair was used with a 4-fold excess of Fmoc-protected amino acid. A low load (0.8 - 0.22 mmol/g) Fmoc-PAL-PEG-PS resin (0.17 mmol/g, PerSeptive Biosystems) was used to produce an amidated C-terminus. After synthesis, peptides were cleaved from the resin with a cocktail of 88% trifluoroacetic acid (TFA), 5% phenol, 5% ultrapure water and 2% triisopropylsilane (TIPS). The cleaved peptides were precipitated using ice-cold diethyl ether and then filtered though a Buchner filter funnel. The filtered crude peptide was re-suspended in 10% acetonitrile, 0.01% TFA in HPLC-grade H_2O for purification by the HPLC.

To synthesize selectively-^{15}N-labeled MCP_{TM} the same procedure was used except Fmoc-protected ^{15}N-labeled amino acids (Cambridge Isotope Laboratories) were used instead of the unlabeled amino acids at specific sites in the amino acid sequence.

3.2.2 MCP_{TM} purification

Peptide purification was carried out on a reverse-phase C4 semi-preparative column (Waters, 15μm spherical silica, 300Å pore sizes with C4 bonded phase) using a 2%/min acetonitrile gradient in 0.01% (v/v) aqueous trifluoroacetic acid (TFA) and elution monitored by absorbance at 215 nm. The desired peptide product corresponded to the major peak and eluted at ~70% acetonitrile. The purified product was lyophilized and ESI mass spectrometry performed by the University of Ottawa Mass Spectrometry Centre to confirm the molecular weight of the purified sample. Analytical HPLC chromatography on a C4 column (Waters, 5μm spherical silica, 100Å pore sizes with C4 bonded phase) was used to confirm the purity of the peptide which was generally found to be > 98%.

3.2.3 Determination of MCP_{TM} concentration

3.2.3.1 BCA protein assays

The MCP_{TM} concentration was determined using the Bicinchoninic acid (BCA) Protein Assay kit provided by Pierce Biotech. This is a detergent-compatible formulation that allows the colorimetric quantitation of total protein or peptide concentrations. In this method, protein concentrations were determined and reported with reference to bovine serum albumin (BSA) protein standards. BSA standards were prepared by serial dilution of a BSA stock solution (Pierce Biotech) of known concentration (2.0 mg/mL). The protocols provided in the kit were followed for this assay, which usually required 0.1 mL of each standard or unknown sample be mixed

with 2.0 ml of BCA working solution (Pierce Biotech). After incubation at 37°C for 30 minutes (or at room temperature for 2 hours), all samples were allowed to cool to room temperature. The absorbance at 562 nm was measured within 10 minutes and a standard curve was obtained by plotting a blank-corrected 562 nm measurement for each BSA standard vs. its concentration. A least squares linear fit of the data was done to find the relationship between absorbance and concentration. The peptide concentration of each unknown sample was then determined from this equation. Protein concentration determination was always done in duplicate with a standard curve generated in parallel with the MCP_{TM} peptide analysis.

3.2.3.2 Ultraviolet absorption

The MCP_{TM} concentration can also be determined by ultraviolet (UV) absorption since MCP_{TM} peptide contains aromatic amino acids - tyrosine (Y) and tryptophan (W) which absorb at approximately 280 nm. UV absorption provides a quick way to estimate the concentration of protein with <10% experimental error when compared to BCA protein assay. Therefore, we also used this method to determine the MCP_{TM} concentrations in this study. 3 μL MCP_{TM} or its mutant was mixed with 350 mL guanidinium chloride (6M, 0.02 M phosphate buffer, pH 6.5) and the absorbance of the mixture was measured using UV/vis spectrophotometer (Ultrospect 2100, Biochrom) at 280 nm. The extinction coefficients (ε_{280}) for 1 mg/mL MCP_{TM} and its mutants were determined to be 2.367 mg/mL/a.u. with the assistance of ProtParam at ExPASy which allows the computation of various physical

and chemical parameters from a user entered sequence. After the UV spectrophotometer was zeroed with a buffer blank, the absorbance of MCP_{TM} peptide or its mutants at 280 nm (A_{280}) were determined and the corresponding concentration (mg/mL) was equal to $\varepsilon_{280} \times A_{280}$.

3.2.4 Preparation of spin-labeled MCP_{TM}

MCP_{TM} mutants containing a cysteine residue at a specific site were synthesized and purified as described above. To derivatize the peptide with a sulfhydril-reactive nitroxide spin label, a tenfold molar excess of (1-oxyl-2,2,5,5-tetramethyl-Δ^{3}-pyrrolin-3-yl) methyl methanethiosulfonate (MTSL; Toronto Research Chemicals) was added to an aqueous peptide MCP_{TM} sample and the pH of the mixture was then adjusted by addition of sodium hydroxide until the pH meter reading reached ~8. In this weakly basic solution, spin label was specifically attached to the peptide though disulfide bond formation with the cysteine sulfhydrl group (Berliner et al., 1982; Kosen et al., 1986; Todd et al., 1989, Figure 26). The reaction was allowed to proceed overnight at room temperature to ensure complete derivatization of the peptide. Excess spin label was removed by a subsequent purification step using HPLC as described in section 3.2.2. The identity of the final spin-labeled MCP_{TM} was confirmed by ESI mass spectroscopy.

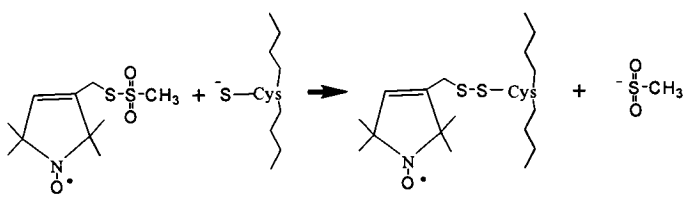


Figure 26: Alkylation of a cysteine thiolate by the nitroxide spin label.

When required, reduction of the nitroxide spin label was achieved by addition of a 2-3 fold molar excess of ascorbic acid. Typically the sample was allowed to incubate at 37 °C for 1 hour to ensure the complete reduction of the nitroxide spin label prior to NMR spectroscopy of the reduced sample.

3.2.5 NMR spectroscopy

~0.3-0.5 mM of lyophilized MCP_{TM} peptide was dissolved in 10% D_2O with various concentrations of either SDS or β-octyl glucoside (OG) detergents while maintaining the pH at 5.5. The sensitivity enhanced version of the 2D $^{1}H - ^{15}N$ HSQC experiments (Kay et al., 1992) were recorded at 37°C on the 500-MHz Varian INOVA or at 41°C on the 500-MHz Bruker Avance spectrometer in the University of Ottawa Faculty of Science NMR Facility. Spectra were processed and analyzed using NMRPipe (Delaglio *et al.*, 1995) and NMRView (Johnson & Blevins, 1994) with resolution enhancement (90° shifted sine-bell squared function in both dimensions). Peak volumes were measured using the nlinLS function in NMRPipe.

Backbone amide proton and nitrogen assignments of the MCP_{TM} peptide in detergent micelles were obtained from previously published assignments when possible (Henry & Sykes, 1992). ^{15}N-HSQC-TOCSY spectra (Kay, 1992) were also run to confirm the identity of the amino acid.

3.2.6 Circular dichroism (CD) spectroscopy

Far-UV circular dichroism (CD) spectroscopy was used to characterize the

secondary structure content for peptide samples. Spectra were recorded on a Jasco-810 instrument kindly made available in the Scaiano lab with 50 and 100 uM MCP_{TM} samples using solvent conditions identical to parallel NMR experiments. Spectral scans were performed from 250-200 nm with a step-resolution of 0.2 nm, a speed of 20 nm/minute and a bandwidth of 1.0 nm using a 0.1 mm path-length quartz cuvette. Ellipticity measurements were converted from millidegrees to mean residue ellipticity in Microsoft Excel using:

$$\text{Mean residue ellipticity} = \frac{100\theta}{cnl} \qquad [15]$$

where θ is the measured ellipticity in mdeg, c the peptide concentration in μM determined using the BCA protein assay or UV absorbance, n the number of amino acid residues in the peptide sequence, and l the pathlength in cm.

3.3 Results and discussion

3.3.1 Peptide design

The MCP_{TM} utilized in this study was synthesized by solid-phase methods using Fmoc-protected amino acids. It spans residues 21 to 48 of wild-type MCP, except that Leu was substituted for Met 28 and Val 31 in order to enhance the stability of the TM helix dimer (Melnyk et al., 2002). Moreover, three non-native lysine residues were added to the N-terminus to improve its solubility without affecting the ability of peptide to dimerize (Melnyk *et al.*, 2002). Twelve ^{15}N-labeled amino acids were also incorporated into this peptide (Gly, Ala, Leu and Val) (underlined in Figure 27) in order to facilitate characterization of oligomerization state by NMR.

$$H_2N\text{-}\mathbf{KKKYI\underline{G}Y\underline{A}W\underline{A}\overset{28}{\underline{L}}\underline{V}\overset{31}{\underline{V}}\underline{L}IV\underline{GA}TI\underline{G}IK\underline{L}FKKFTSK}\text{-}COONH_2$$

Figure 27: The sequence of the MCP_{TM} peptide used in this study. Underlined regions of the sequence identify sites where an ^{15}N-label was incorporated in the backbone. Mutations Met28Leu and Val31Leu (red) were introduced in order to promote TM helix association. The C-terminus was amidated to increase the net charge of the peptide to improve solubility.

3.3.2 NMR analysis of MCP_{TM} in SDS

In order to evaluate the utility of MCP_{TM} as a model system it was necessary to characterize its NMR spectrum under a range of solution conditions. Since MCP_{TM} used in the present work is soluble in water, 2D ^1H-^{15}N HSQC spectra were first

recorded for the selectively ^{15}N-labeled MCP_{TM} peptide in water with 10% D_2O (Figure 28). As shown in Figure 28, 12 peaks were observed as would be expected from the selective incorporation of ^{15}N into the backbone amide nitrogen positions for 12 amino acids. These peaks exhibited a small degree of amide proton chemical shift dispersion, suggesting that in the absence of detergents this peptide is largely unfolded and flexible as was previously suggested by CD studies (Melnyk et al., 2002).

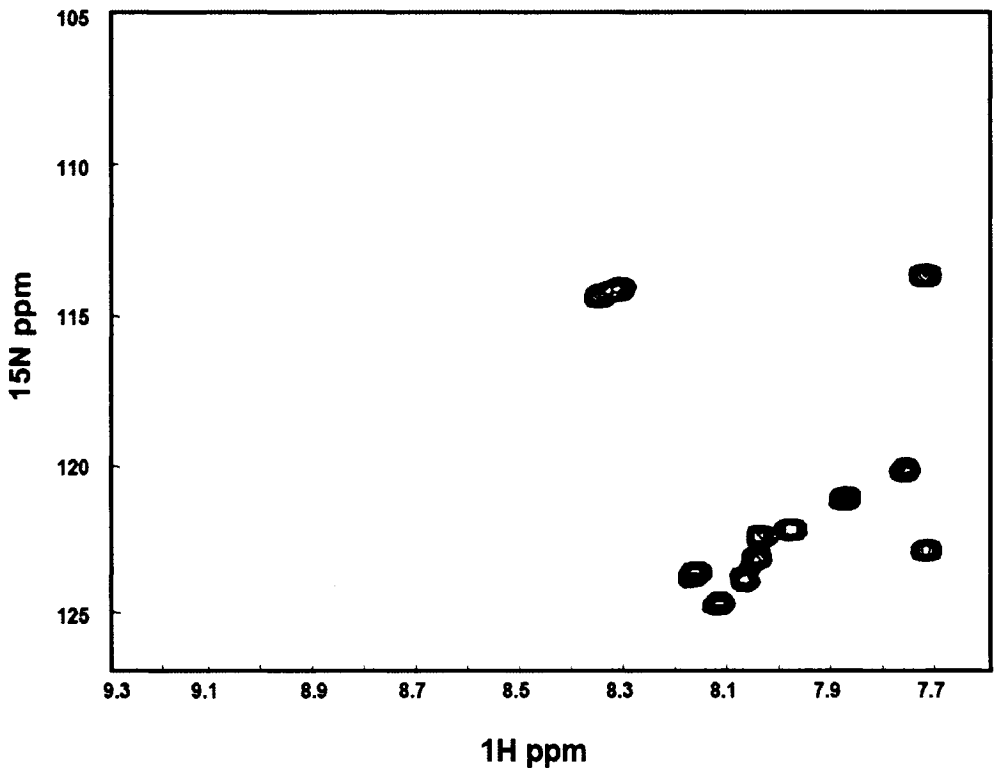


Figure 28: 2D ^{1}H-^{15}N HSQC spectrum of selectively ^{15}N-labeled MCP_{TM} (~0.2 mM) in water with 10% D_2O and pH 5.5 at 37°C.

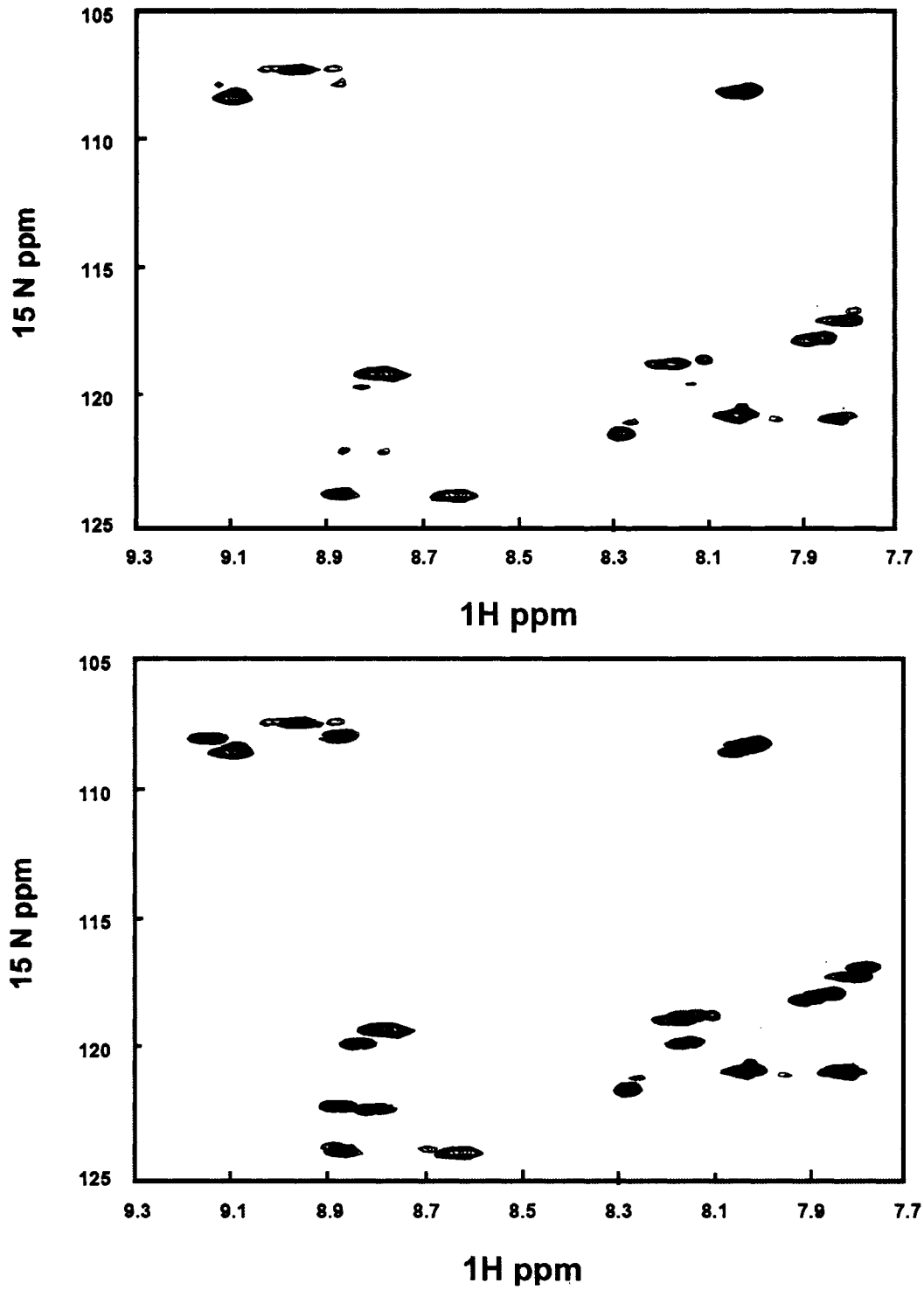


Figure 29: 2D ^{1}H-^{15}N HSQC spectra of 0.2 mM MCP_{TM} in (top spectrum) 50 mM SDS, and (bottom spectrum) 250 mM SDS (red) and 500 mM SDS (green) superimposed on the 50 mM SDS spectrum (black). All spectra were acquired at 41 °C, pH 5.5 on a 500-MHz Bruker Avance NMR spectrometer.

During addition of submicellar concentrations of SDS detergent (less than 10 mM) to this MCP_{TM} solution, the HSQC spectra showed no change at all, with superimposable peaks in the presence and absence of SDS. However, above ~15-20 mM, which is also above the critical micelle concentration (CMC) for SDS (~ 8 mM), the appearance of the spectrum was altered dramatically with large changes in chemical shifts and significant peak broadening. The increase in amide proton chemical shift dispersion suggests that the peptide adopted a more structured conformation, while the increase in peak linewidth likely arises from an increase in molecular mass due to association of the peptide with the micelle.

CD spectroscopy was used to evaluate the secondary structure of the MCP_{TM} at a similar concentration SDS. As shown in Figure 30, a typical α-helical spectrum with two minima at 208 nm and 222 nm was obtained. Moreover, the mean residue molar ellipticity at 222 nm approaches maximum values observed for completely helical standards (e.g. polyLys with mean residue molar ellipticity of ~ -30×10^3 deg cm^2 $dmol^{-1}$, Greenfield & Fasman, 1969), indicating that MCP_{TM} is almost completely helical under these conditions. Therefore, similar to previous observations with a related peptide (Deber, 1993; Melnyk et al., 2002), MCP_{TM} binds to SDS micelles and adopts an α-helical structure.

A particularly interesting feature of the NMR spectrum for MCP_{TM} in low SDS concentrations was the presence of small peaks corresponding to a minor component which indicated the presence of more than one species of peptide in these conditions. For simplicity, we will refer to the minor component of this spectrum as

the species A and the major component species B. However, as the concentration of SDS was increased, the peaks from the species A decreased in magnitude, while peak intensities from species B increased. At 500 mM SDS only peaks from species B were present in the spectrum, with no further changes occurring at higher concentrations of SDS. The appearance of these two distinct species in the NMR spectra whose relative population is SDS concentration-dependent indicates that the structure and/or the oligomerization state of this peptide changes as a function of detergent concentration.

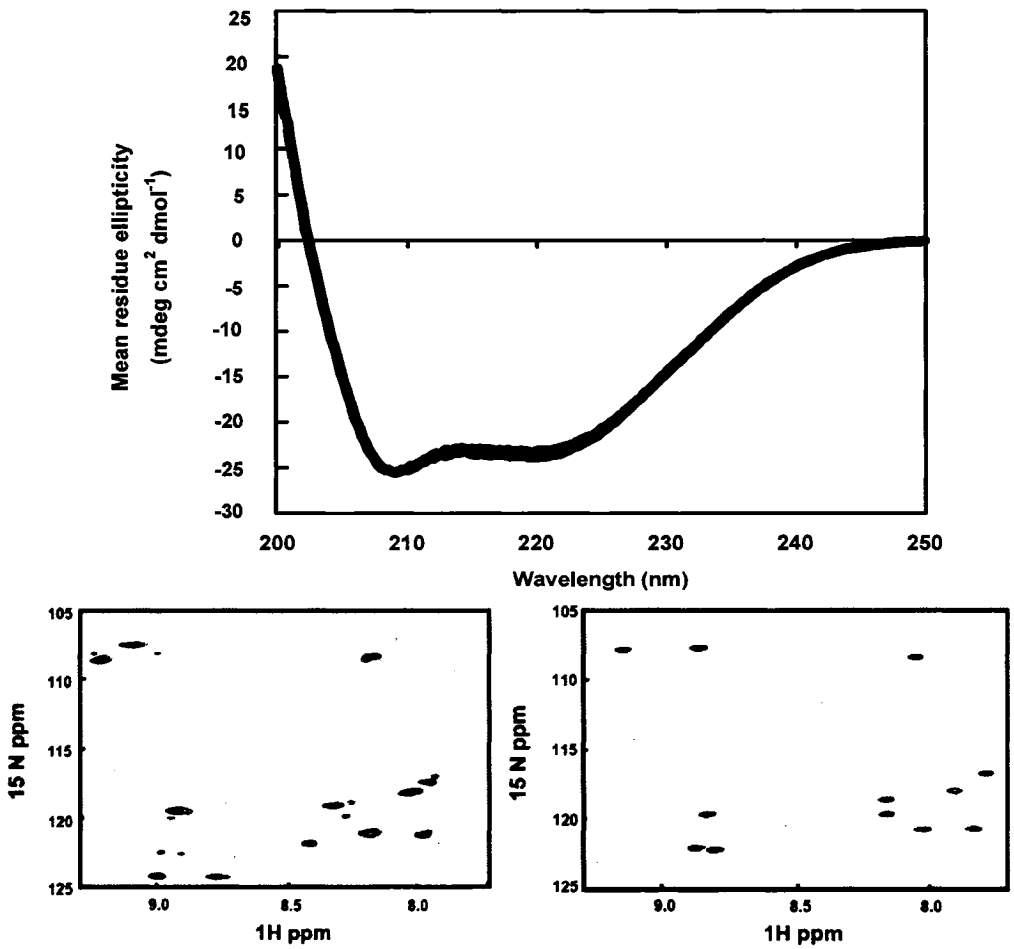


Figure 30: Secondary structure of 0.1 mM MCP_{TM} (top) determined by circular dichroism spectroscopy at 37°C in 15 mM (black) and 100 mM (green) SDS, pH 5.5. 2D ^{1}H-^{15}N HSQC spectra (bottom) acquired for these samples under the same conditions confirms that either species A (black) or species B (green) was the main component.

In order to determine whether the two species observed in the NMR spectra reflect a difference in MCP_{TM} structure, a CD spectrum was recorded under conditions designed to favor species A and compared to the CD spectrum of this peptide in a high concentration of SDS that should give rise to species B. Since it was not possible to record the CD spectra at the same concentrations that were used to acquire the spectra shown in Figure 29, it was necessary to adjust the concentration of SDS to obtain the desired species. NMR spectra were recorded for these CD samples to confirm the presence of A or B species. As shown in Figure 30, for a 100 μM MCP_{TM} sample, species A was dominant at 15 mM SDS while species B dominated at a concentration of 100 mM SDS.

CD spectra were acquired for both samples and found to be almost superimposable, indicating MCP_{TM} adopts an α-helix structure at both low and high concentrations of SDS. Since there was no significant change in secondary structure content, it is likely that the changes observed in the NMR spectrum do not arise from a change in the secondary structure of this peptide. Instead these detergent-dependent spectral changes probably arise from a change in the oligomerization state of MCP_{TM} in SDS. This would be consistent with the wide range of studies showing that M13 TM peptides form dimers in SDS (Melnyk et al., 2002). Moreover, studies of another TM homodimer called glycophorin A (GpA), have demonstrated that the monomer-dimer equilibrium can be similarly affected by changes in the detergent concentration (Fleming, 2002; Fisher, 2003), with higher concentrations giving rise to a larger dissociation constants. This is thought to occur since raising the detergent

concentration dilutes the peptide in a micellar phase, thereby driving dissociation entropically. This is consistent with our results here, indicating that MCP_{TM} peptide adopts a dimeric form at lower concentrations of SDS and a monomeric form at higher concentrations of SDS. Overall, these results show that the ^{1}H-^{15}N HSQC spectrum can be used to identify conditions that favor TM helix interaction in MCP_{TM}.

3.3.3 NMR analysis of MCP_{TM} in β-octyl glucoside (OG)

In order to test the generality of this approach, we examined the effect of detergent concentration on MCP_{TM} NMR spectra in a different type of detergent. For this study we chose the non-ionic detergent β-octyl glucoside (OG), a mild detergent that has been widely utilized for solubilizing membrane proteins in their native state. Given the mild nature of this detergent and the relatively small size of the OG micelle (~30 kDa) it should be possible to preserve a homodimeric interaction between MCP_{TM} peptides in this solvent.

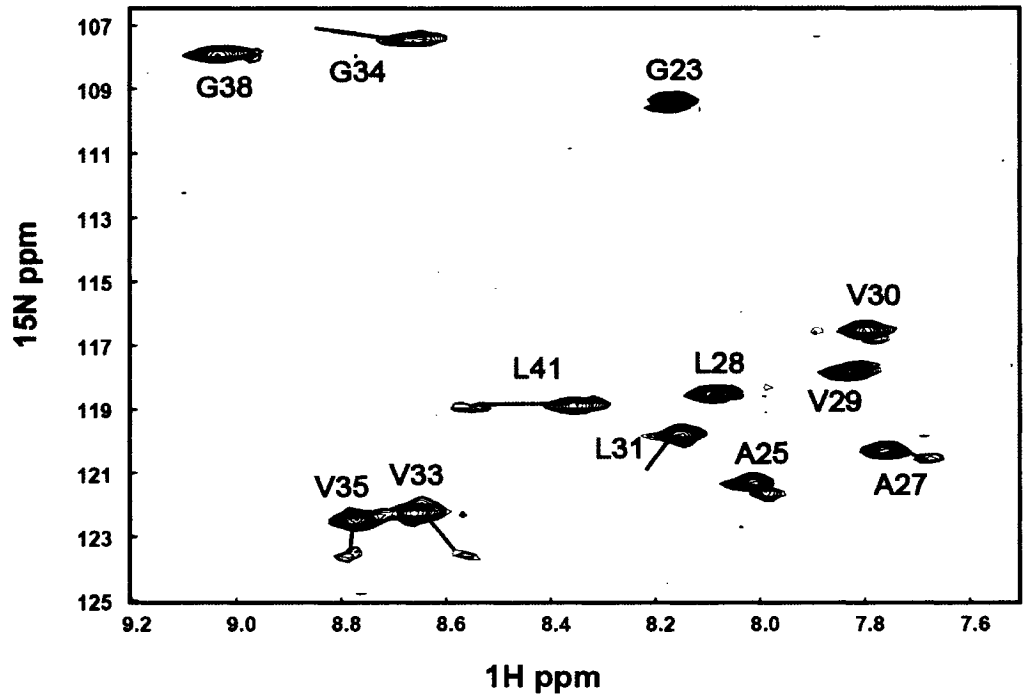


Figure 31: 2D ^{1}H-^{15}N HSQC spectra of 0.3 mM ^{15}N-labeled MCP_{TM} in 30 mM OG (red) or 100 mM OG (green), pH 5.5, at 37°C.

To test this possibility, selectively ^{15}N-labeled MCP_{TM} was prepared in 30 mM OG which is just above its critical micelle concentration and a 2D ^{1}H-^{15}N HSQC spectrum was recorded. As shown in Figure 31 (red spectrum) 12 peaks were observed with different chemical shifts from those seen in the SDS spectra. Given the different chemical environments presented by the two different detergents, these differences in chemical shifts are expected. However, we were interested in seeing whether OG concentration-dependent changes in the NMR spectrum would arise as was observed for SDS. For this purpose OG was added to a final concentration of 100 mM yielding the spectrum in green (Figure 31). Significant shifts in peak positions relative to the 30 mM OG spectrum were observed, suggested an OG concentration-dependent change in the peptide conformational or oligomerization

state. No further changes in the spectrum were observed when more OG was added to the sample.

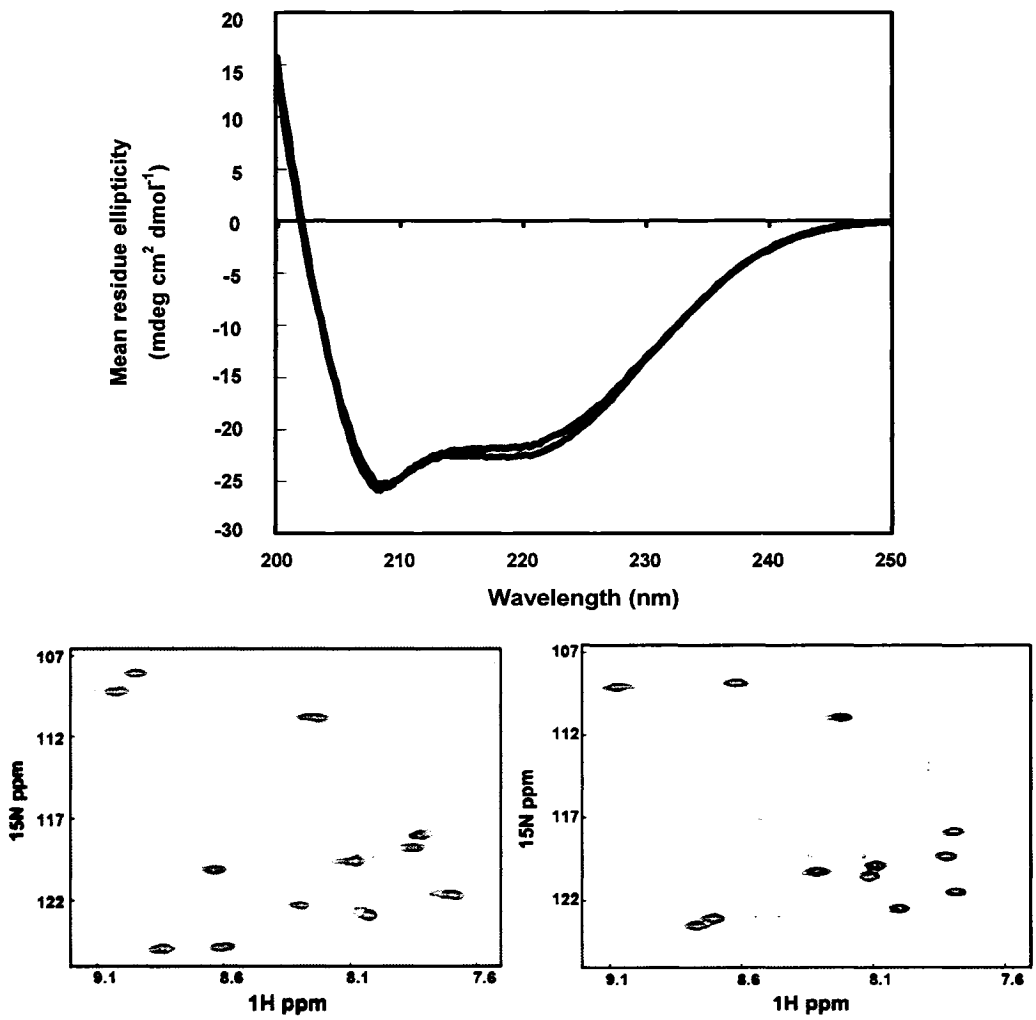


Figure 32: Top: CD spectra of MCP_{TM} in 30 mM OG (black) and 100 mM OG (green), pH 5.5. Bottom: ^{1}H-^{15}N HSQC spectra acquired for the same samples (same color scheme as CD spectra), demonstrating that the CD spectra are representative for predominantly one species of MCP_{TM}.

To evaluate the role of peptide structure in this change in the HSQC spectrum, CD spectroscopy was performed on samples giving rise to the two different types of

OG spectra as was done for the SDS samples. The CD spectra of MCP_{TM} in low and high OG concentrations in Figure 32 showed typical features of an α-helical spectrum. More importantly there was only a very small change in the helical content for these two samples, indicating that there was only a slight conformational change in MCP_{TM} upon increase of OG concentration. Although the CD data suggested that some chemical shift changes may be attributed to a slight increase in α-helical content at higher OG concentrations, this does not rule out the possibility that some of these changes arise due to a change in the oligomerization state of MCP_{TM}. However, unlike the SDS studies, no information is available regarding the oligomerization behavior of the M13 TM segment in this type of detergent. Therefore in order to confirm that the spectral changes that we observed in OG actually reflect a change in the oligomerization state of MCP_{TM}, it was necessary to perform additional experiments that could provide information about any *inter*molecular interactions that may be occurring. For this purpose we developed a system that utilizes the paramagnetic broadening effect of spin labels.

3.3.4 Investigation of MCP_{TM} oligomerization using spin label

In order to introduce a spin label into MCP_{TM}, a cysteine residue was introduced into the sequence to allow derivatization by the sulfhydril-reactive spin label MTSL. To maximize the possibility that an *inter*molecular paramagnetic broadening effect could be observed we introduced this cysteine residue in a central region of the peptide. In order to reduce the possibility that this non-native cysteine

could cause a structural perturbation that could affect the ability of MCP_{TM} to homodimerize, we substituted Val 33 for Cys since extensive mutagenesis studies suggest that the side chain from this residue will face the lipid membrane and not the dimeric interface (Deber, 1993). This peptide was synthesized with no ^{15}N-isotope label and was called $MCP_{TM\text{-}Cys33}$. To confirm that this mutant did not affect the structure of MCP_{TM}, CD spectroscopy was used. As shown in Figure 33, CD spectrum of $MCP_{TM\text{-}Cys33}$ in 30 mM OG was almost identical to that of the original MCP_{TM} peptide.

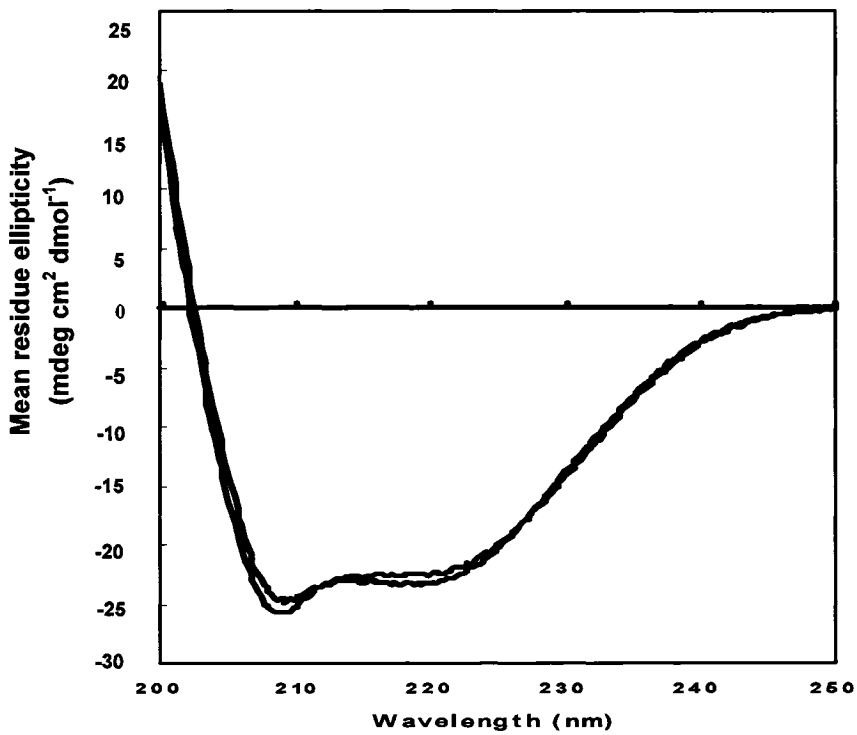


Figure 33: Secondary structure comparison of ~0.1 mM MCP_{TM} (black) and its mutant $MCP_{TM\text{-}Cys33}$ (red) in 30 mM OG micelles, as determined by circular dichroism spectroscopy.

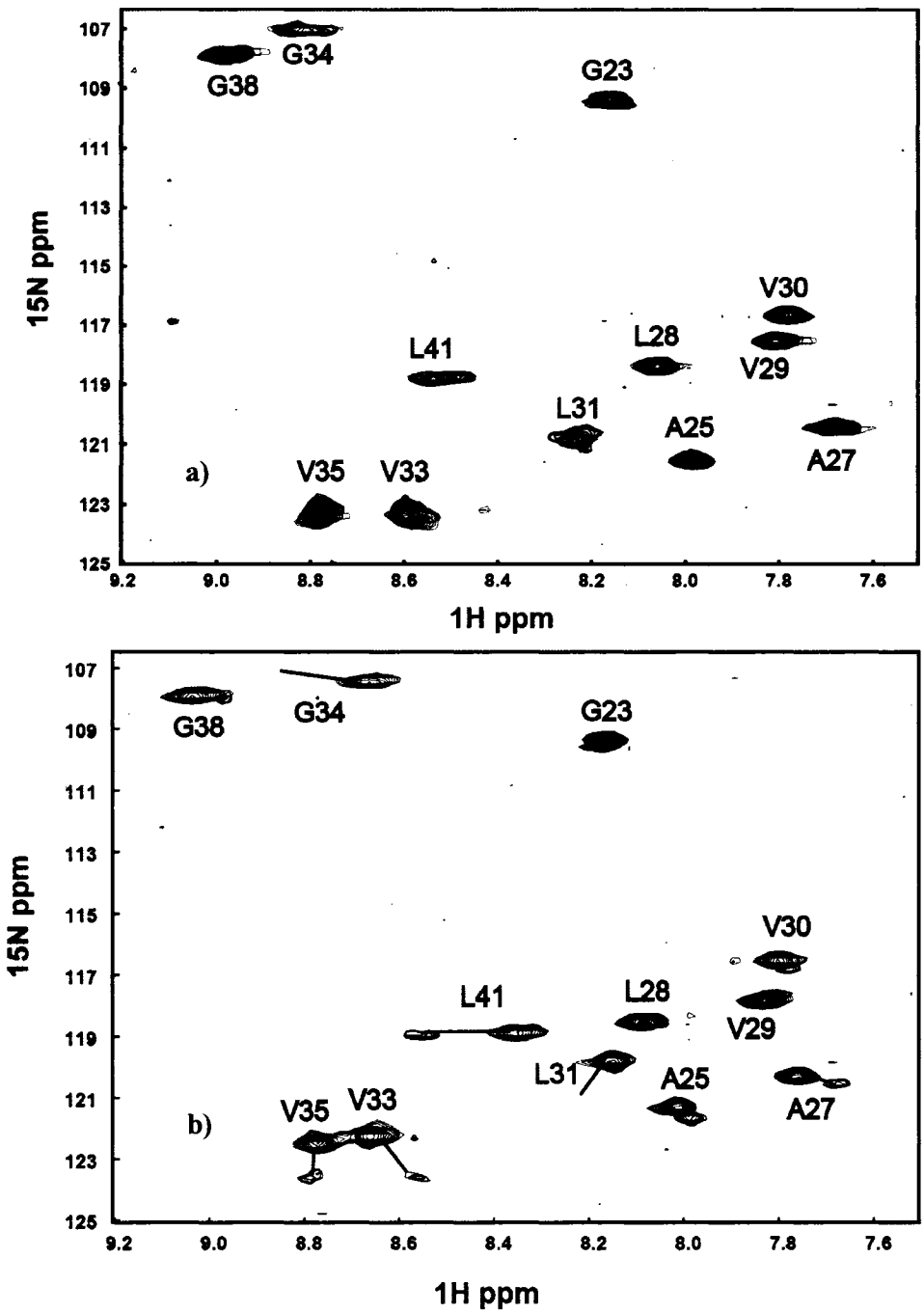


Figure 34: a) 2D ^{1}H-^{15}N HSQC spectra of 0.3 mM selectively ^{15}N-labeled MCP_{TM} at 37°C either alone (black spectrum), with 0.15 mM spin-labeled $MCP_{TM\text{-}Cys33}$ (molar ratio ~1:0.5, blue spectrum) or 0.25 mM spin-labeled $MCP_{TM\text{-}Cys33}$ (~1:0.8, red spectrum). b) 2D ^{1}H-^{15}N HSQC spectra of the 1:0.8 ratio sample from a) in the presence of in 30 mM OG (red) and 100 mM OG (green).

To see if the spin-labeled peptide would affect the spectrum of the ^{15}N-labeled peptide that had no spin label, $MCP_{TM\text{-}Cys33}$ was derivatized with the MTSL nitroxide spin label and titrated into a 0.3 mM solution of ^{15}N-labeled MCP_{TM} in 30 mM OG. As shown in Figure 34, as an increasing amount of spin-labeled peptide was added to ^{15}N-labeled peptide, the HSQC spectrum showed an increasing degree of peak broadening suggesting that *inter*-molecular paramagnetic broadening was occurring. This implies that ^{15}N-labeled MCP_{TM} and spin-labeled $MCP_{TM\text{-}Cys33}$ have formed heterodimers leading to spin label-mediated broadening of the signal from ^{15}N-bonded amide protons that are within a ~ 15 Å radius of the spin label. In particular G34 and L31 showed the largest amount of broadening, suggesting that these central residues are closest to the Cys33 residue that has been deriviatized with the spin label. The more extensive broadening observed in the central portion of the peptide indicates that the spectrum acquired at low OG concentration arises from a parallel or anti-parallel arrangement of monomers in the dimer.

To confirm that the NMR spectrum obtained at higher concentrations of OG corresponds to a monomeric form of MCP_{TM}, more OG detergent was added to the 1:0.8 (^{15}N-labeled: spin-labeled) ratio sample. As shown in Figure 34 (green), the chemical shifts changed to give the high concentration OG spectrum and peak intensities were uniformly recovered. Elimination of the paramagnetic broadening effect through the addition of detergent reflects the effect of diluting the peptide in the micellar phase, which shifts the monomer-dimer equilibrium towards the monomeric form. Therefore, the two NMR spectra obtained for MCP_{TM} at the different OG

concentrations appear to arise from a change in the oligomerization state of this peptide. Not only does this data provide the first evidence that the MCP TM segment can form an oligomer in OG micelles, but it also demonstrates that the 2D ^{1}H-^{15}N HSQC spectrum of this model system can be used to identify conditions that favor interactions between TM helices.

3.3.5 Evidence for a parallel arrangement of MCP_{TM} in the dimer

So far the NMR and spin-labeling data suggest at least a dimeric form of MCP_{TM} at low OG concentrations. To assess whether these dimers are parallel or antiparallel, a new peptide was synthesized with no ^{15}N-isotope and an additional Cys residue included just before the first M13 residue (Y21) and called $MCP_{TM\text{-}Cys20}$. After confirming by CD that this Cys residue did not introduce any structural changes in the peptide (data not shown) this peptide was derivatized with MTSL and co-dissolved with ^{15}N-labeled MCP_{TM} in 30mM OG. As shown by the 2D ^{1}H-^{15}N HSQC spectrum in Figure 35 a) there is selective broadening of peaks from residues closer to the N-terminus, with G23 almost completely disappearing from the spectrum. In order to quantitate the extent of broadening the nitroxide spin label was reduced and the HSQC recorded again (Figure 35 b). As expected all peak intensities were recovered, demonstrating that the broadening had been caused by the spin label.

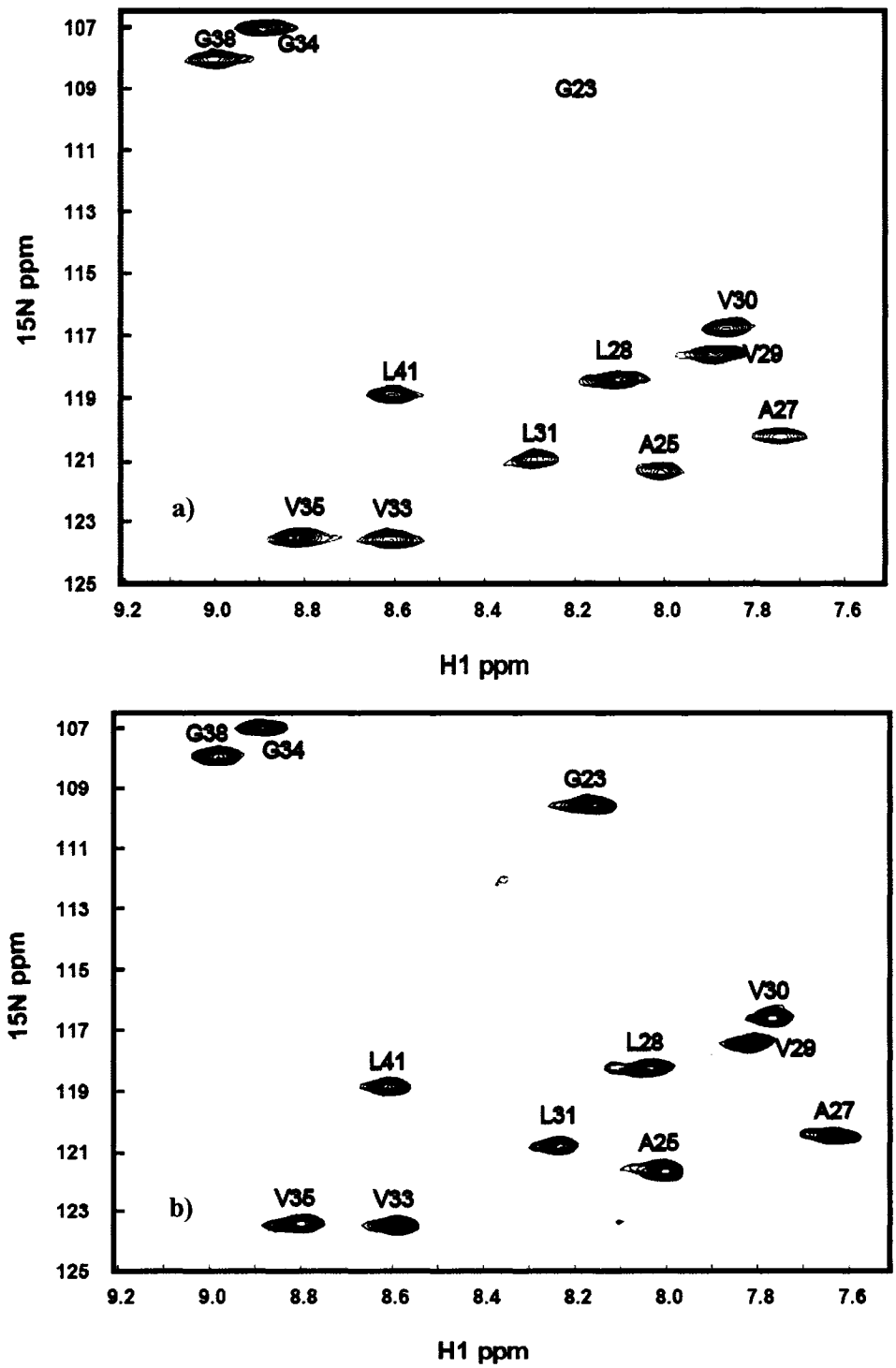


Figure 35: ^{1}H-^{15}N HSQC spectra of ~0.25 mM selectively ^{15}N-labeled MCP_{TM} and approximately the same concentration of spin-labeled $MCP_{TM\text{-}Cys20}$ peptide, pH 5.5, 10% D_2O, at 37 °C in 30 mM OG a) before and b) after reduction of the nitroxide radical with ascorbic acid.

The ratio of peak intensities in these HSQC spectra before and after reduction was used to quantitate the paramagnetic broadening effect as shown in Figure 36. This analysis showed that the spin-label specifically broadened the N-terminal resonances on the ^{15}N-labeled MCP_{TM} peptide, with the magnitude of the broadening decreasing rapidly for residues further away in the protein sequence. This effect leveled off around the center of the TM segment suggesting that residues beyond this point in the sequence were further than 15 Å away from the spin label. This pattern of *inter*-molecular broadening provides excellent evidence that the MCP_{TM} dimer is parallel in OG. The results obtained here are consistent with excimer formation studies done with related M13 peptide which showed that the dimers formed in SDS micelles are also parallel. This demonstrates that the dimer obtained in the ionic SDS detergent appears to be similar to the dimer in the non-ionic OG and that the OG detergent should be also be a suitable solvent for structure determination of this dimer.

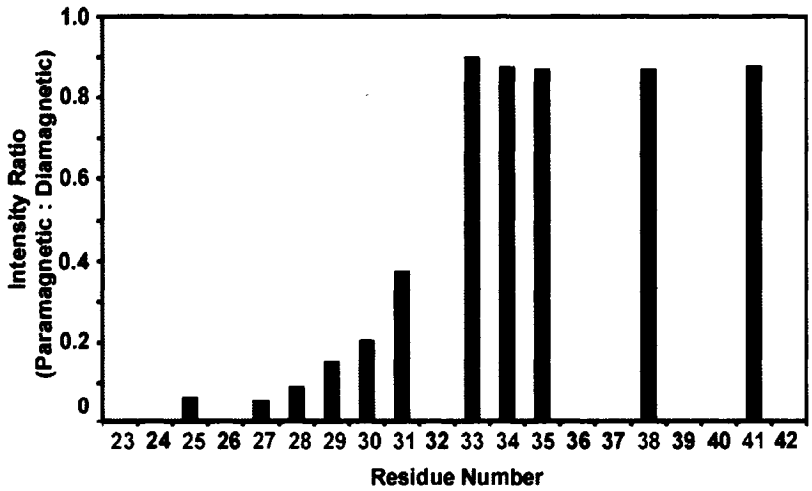


Figure 36: Inter-molecular paramagnetic broadening effect as determined from the peak intensity ratios for spectra shown in Figure 35 plotted with respect to residue number. ^{15}N-labeled amino acids are highlighted on the x-axis in red.

3. 4 Future work

According to equations 13 and 14, the paramagnetic broadening effect is inversely proportional to the sixth power of the distance between the unpaired electron on the spin label and the affected amide proton, and thus these relaxation data can be used as distance constraints in structure calculations (McConnell, 1967; Stemlicht & Wheeler, 1967). This type of long-range distance restraint has also provided a rapid approach for structure determination of larger proteins (Schmidt & Kuntz, 1984; Kuliopulos, 1987, Gillespie, 1997 a & b; Gaponenko, 2000). While the structure of the monomeric form of MCP has been solved by NMR spectroscopy (Papavoine *et al.*, 1998) no structure of the MCP dimer has yet been determined. Therefore it should be possible to use these paramagnetic restraints to determine its structure using this approach. This will open the door to structure determination of the dimer in various micellar environments to evaluate the impact of a detergent on the TM helix-helix interaction. Overall, the model system and methods established here provide a fast and effective way to screen detergents for compatibility with solution NMR studies of membrane protein structure, an investigation that we plan to pursue in the future.

3.5 Conclusion

Oligomerization of MCP_{TM} under a range of conditions was investigated using 2D ^{1}H-^{15}N HSQC spectroscopy in combination with site-directed spin labeling. Our results show that detergents generally recognized as denaturing (e.g. SDS), as

well as those known as non-denaturing (OG), are capable of maintaining a water-soluble form of the peptide while preserving the TM-helix homodimer interaction. Our spin labeling result also shows that the MCP_{TM} associates into a parallel homodimeric configuration at low detergent concentrations and becomes monomeric at higher detergent concentrations. Overall these results demonstrate that selectively ^{15}N-labeled MCP_{TM} can be used as a model system to study the ability of other detergents to support interactions between TM helices under conditions that are compatible with solution NMR.

Claims to Original Research

1. Development and implementation of the ^{15}N-edited BPP-LED pulse sequence to estimate the size of membrane protein – detergent complexes.

2. First characterization of NMR spectra for the transmembrane segment peptide from the M13 bacteriophage coat protein (MCP_{TM}) with both denaturing and non-denaturing detergents at a range of concentrations.

3. First demonstration that the MCP_{TM} peptide associates into a parallel homodimeric configuration at low concentrations of OG using spin labeling in combination with 2D ^{1}H-^{15}N HSQC spectroscopy.

4. Publication: Melnyk, R.A.; Partridge, A.W.; Yip, J.; Wu Y.; Goto N.K. & Deber, C.M. Polar residue tagging of transmembrane peptides. 2003, *Biopolymers*. 71:675-685.

5. Oral presentation: Wu, Y.; Goto, N. K. **Investigation of Detergents for Structural Characterization of Membrane Proteins by Solution NMR Spectroscopy,** June 3rd 2005, Ottawa-Carleton Chemistry Institute Day (OCCI), Ottawa, Ontario.

6. Poster Presentation: i) Wu, Y.; Goto, N. K. **Detergent screening and evaluation for the structural study of membrane proteins by solution NMR**, Sept., 25-26, 2004, MOOT

XVII NMR Symposium. Queen's University, Kingston, Ontario. ii) Wu, Y.; Goto, N. K. **Investigation of detergent for structural characterization of a dimeric transmembrane segment by solution NMR**, May 9th 2005, 1st Annual McGill Biophysical Chemistry Symposium. McGill University, Montreal, Quebec.

References

Abragam, A., The principles of Nuclear Magnetism. *Clarendon Press, Oxford* **1961**.

Altieri, A. S.; Hinton, D. P.; Byrd, R. A., Association of biomolecular systems via pulsed field gradient NMR self-diffusion measurements. *J. Am, Chem. Soc.* **1995,** 117, 7566-7567.

Andrec, M.; Prestegard, J. H., Quantitation of chemical exchange rates using pulsed-field-gradient diffusion measurements. *J. Biomol. NMR* **1997,** 9, (2), 136-150.

Arora, A.; Abildgaard, F.; Bushweller, J. H.; Tamm, L. K., Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. *Nature Structural Biology* **2001,** 8, (4), 334-338.

Arora, A.; Tamm, L. K., Biophysical approaches to membrane protein structure determination. *Curr. Opin. Struct. Biol.* **2001,** 11, 540-547.

Bales, B. L., Fluorescence quenching of pyrene by copper(II) in sodium dodecyl sulfate micelles. Effect of micelle size as controlled by surfactant concentration. *J. Phys. Chem.* **1995,** 99, 15153-15162.

Bax, A.; Ikura, M., An efficient 3D NMR technique for correlating the proton and ^{15}N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly $^{15}N/^{13}C$ enriched proteins. *J. Biomol. NMR* **1991,** 1, 99-104.

Berliner, L. J.; Grunwald, J.; Hankovszky, H. O.; Hideg, K., A novel reversible thiol-specific spin label: papain active site labeling and inhibition. *Anal. Biochem.* **1982,** 119, 450-455.

Bezzobotnov, V. Y.; Borbély, S.; Cser, L.; Faragó, B.; Gladkih, I. A.; Ostanevich, Y. M.; Sz., V., Temperature and concentration dpendence of properties of sodium dodecyl sulfate micelles determined from small-angle neutron scattering experiments. *J. Phys. Chem.* **1988,** 92, 5738-5743.

Bloembergen, N.; Purcell, E. M.; Pound, R. V., Relaxation effects in nuclear magnetic resonance absorption. *Phys. Rev.* **1948,** 73, 679-712.

Buck, M.; Schwalbe, H.; Dobson, C. M., Characterisation of conformational preferences in a partly folded protein by heteronuclear NMR spectroscopy: Assignment and secondary structure analysis of hen egg-white lysozyme in trifluoroethanol. *Biochemistry* **1995,** 34, 13219-13232.

Cantor, C. R.; Schimmerl, P. R., Biophysical chemistry. *W. H. Freeman Company* **1980**.

Cavanagh, J., et al. , Protein NMR Spectroscopy: Principles and Practice. . *San Diego: Academic Press* **1996**.

Choma, C.; Gratkowski, H.; Lear, J. D.; DeGrado, W. F., Asparagine-mediated self-association of a model transmembrane helix. *Nat. Struct. Biol.* **2000,** 7, 161-166.

Chou, J. J.; Baber, J. L.; Bax, A., Characterization of phospholipid mixed micelles by translational diffusion. *J. Biomolecular NMR* **2004,** 29, 299-308.

Chou, J. J.; Kaufman, J. D.; Stahl, S. J.; Wingfield, P. T.; Bax, A., Micelle-induced curvature in a water-insoluble HIV-1 Env peptide revealed by NMR dipolar coupling measurement in a stretched polyacrylamide gel. *J. Am, Chem. Soc.* **2002,** 124, 2450-2451.

Claridge, T. D. W., High Resolution NMR Techniques in Organic Chemistry *Amsterdam: Pergamon* **1999.**

Clore, G. M.; Gronenborn, A. M., Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. *Crit. Rev. Biochem. Biol. Biol.* **1989,** 24, 479–564.

Cohn, E. J.; Conant, J. B., The molecular weight of proteins in phenol. *PNAS* **1926,** 12, 433-438.

Cross, T. A.; Tsang, P.; Opella, S. J., Comparison of Protein and DNA Structure in fd and Pf1 Bacteriophages. *Biochemistry* **1983,** 22, 721-726.

Deber, C. M.; Khan, A. R.; Li, Z.; Joensson, C.; Glibowicka, M.; Wang, J., Val → Ala mutations selectivity alter helix-helix packing in the transmembrane segment of phage M13 coat protein. *Proc. Natl. Acad. Sci. USA* **1993,** 90, 11648-11652.

Debye, P., Polar Molecules. *Dover pulications, New York.* **1945.**

Delaglio, F.; Grzesiek, S. V., G. W.; Zhu, G.; Pfeifer, J.; Bax, A., NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995,** 6, 277-293.

Dingley, A. J.; Mackay, J. P.; Chapman, B. E.; Morris, M. B.; Kechel, P. W.; Hambly, B. D.; King, G. F., Measuring protein self-association using pulsed-field-gradient NMR spectroscopy: application to myosin light chain 2. *J. Biomol. NMR* **1995,** 6, 321-328.

Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-kay, J. D. K., L. E. , Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry* **1994,** 33, 5984-6003.

Fernández, C.; Hilty, C.; Wider, G.; Güntert, P.; Wüthrich, K., NMR Structure of the integral membrane protein OmpX. *J. Mol. Biol.* **2004,** 336, 1211-1221.

Fisher, L. E.; Engelman, D. M.; Sturgis, J. N., Detergents modulate dimerization, but not helicity, of the glycophorin A transmembrane domain. *J. Mol. Biol.* **1999,** 293, 639-651.

Fisher, L. E.; Engelman, D. M.; Sturgis, J. N., Effect of detergents on the association of the Glycophorin A transmembrane helix. *Biophysical Journal* **2003,** 85, 3097-3105.

Fleming, K. G., Standarding the free-energy change of helix-helix interactions. *J. Mol. Biol.* **2002,** 323, (563-571).

Gaponenko, V.; Howarth, J. W.; Columbus, L.; Gasmi-Seabrook, G.; Yuan, J.; Hubbell, W. L.; Rosevear, P. R., Protein global fold determination using site-directed spin and isotope labeling *Protein Science* **2000,** 9, 302-309.

Gibbs, S. J.; Johnson, C. S., A PFG NMR experiment for accurate diffusion and flow studies in the presence of eddy currents. *J. Magn. Reson.* **1991,** 93, 395–402.

Gillespie, J. R.; shortle, D., Characterization of long-range structure in the denatured state of staphylococcal nuclease. 1. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.* **1997a,** 268, (158-169).

Gillespie, J. R.; Shortle, D., Characterization of long-range structure in the denatured state of staphylococcal nuclease. 2. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol.* **1997b,** 268, 170-184.

Girvin, M. E.; Rastogi, V. K.; Abildgaard, F.; Markley, J. L.; Fillingame, R. H., Solution structure of the transmembrane H+-translocating subunit c of the F1FO ATP synthase. *Biochemistry* **1998,** 37, 8817-8824.

Gounarides, J. S.; Chen, A.; Shapiro, M. J., Nuclear magnetic resonance chromatography: applications of pulse field gradient diffusion NMR to mixture analysis and ligand-receptor interactions. *J. Chromatogr B Biomed Appl* **1999,** 725, 79-90.

Gratkowski, H.; Lear, J. D.; DeGrado, W. F., Polar side chains drive the association of model transmembrane peptides. *Proc. Natl. Acad. Sci. U.S.A.* **2001,** 98, 880-885.

Greenfield, N.; Fasman, G. D., Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **1969,** 8, 4108-4116.

Haigh, N. G.; Webster, R. E., The major coat protein of filamentous bacteriophage f1 specifically pairs in the bacterial cytoplasmic membrane. *J. Mol. Biol.* **1998,** 279, 19-29.

Hayashi, S.; Ikeda, S., Micelle size and shape of sodium dodecyl sulfate in concentrated NaCl solutions. *J. Phys. Chem.* **1980,** 84, 744-751.

Hennel, J. W.; Klinowski, J., Fundamentals of Nuclear Magnetic Resonance. *Longman, Essex, UK* **1993.**

Henry, G. D.; Sykes, B. D., Assignment of amide H-1 and N-15 NMR resonances in detergent-solubilized M13 coat protein - a model for the coat protein dimer. *Biochemistry* 31, 5284-5297.

Hiyama, Y.; Niu, C.; Silverton, J. V.; Bavoso, A.; Torchia, D. A., Determination of ^{15}N Chemical Shift Tensor via ^{15}N-^{2}H Dipolar Coupling in Boc-glycylglycyl[^{15}N]glycine Benzyl Ester. *J. Am. Chem. Soc.* **1988,** 110, 2378-2383.

Hwang, P. M.; Choy, W. Y.; Lo, E. I.; Chen, L.; Forman-Kay, J. D.; Raetz, C. R.; Prive, G. G.; Bishop, R. E.; Kay, L. E., Solution structure and dynamics of the outer membrane enzyme PagP by NMR *Proc. Natl. Acad. Sci. USA* **2002,** 99, 13560-13565.

Johnson, B. A.; Blevins, R. A., NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomolecular NMR* **1994,** 4, 603-614.

Kay, L. E.; Keifer, P.; Saarinen, T., Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **1992,** 114, 10663-10665.

Kay, L. E.; Torchia, D. A.; Bax, A., Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **1989,** 28, 8972-8979.

Koradi, R.; Billeter, M.; Wuthrich, K., MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **1996,** 14, 51-5, 29-32.

Kosen, P. A., Spin labeling of proteins. *Methods Enzymol* **1989,** 177, 86-121.

Kosen, P. A.; Scheek, R. M.; Naderi, H.; Basus, V. J.; Manogaran, S.; Schmidt, P. G.; Oppenheimer, N. J.; Kuntz, I. D., Two-dimensional 1H NMR of three spin-labeled derivatives of bovine pancreatic trypsin inhibitor. *Biochemistry* **1986,** 25, 2356-2364.

Krueger-Koplin, R. D.; Sorgen, P. L.; Krueger-Koplin, S. T.; Rivera-Torres, I. O.; Cahill, S. M.; Hicks, D. B.; Grinius, L.; Krulwich, T. A.; Girvin, M. E., An evaluation of detergents for NMR structural studies of membrane proteins. *J. Biomolecular NMR* **2004,** 17, 43-57.
Krueger-Koplin, R. D.; Sorgen, P. L.; Krueger-Koplin, S. T.; Rivera-Torres, I. O.; Cahill, S. M.; Hicks, D. B.; Grinius, L.; Krulwich, T. A.; Girvin, M. E., An evaluation of detergents for NMR structural studies of membrane proteins. *J. Biomolecular NMR* **2004,** 17, 43-57.

Kuchel, P. W.; Chapman, B. E., Heteronuclear double-quantum-coherence selection with magnetic-field gradients in diffusion experiments. *J. Magn. Reson.* **1993,** 101, 53-59.

Kuliopulos, A.; Westbrook, E. M.; Talalay, P.; Mildvan, A. S., Positioning of a spin-labeled substrate analogue into the structure of delta 5-3-ketosteroid isomerase by combined kinetic,

magnetic resonance, and X-ray diffraction methods. *Biochemistry* **1987,** 26, 3927-3937.

Kyogoku, Y.; Fujiyoshi, Y.; Shimada, I.; Nakamura, H.; Tsukihara, T.; Akutsu, H.; Odahara, T.; Okada, T.; Nomura, N., Structural Genomics of Membrane Proteins. *Acc. Chem. Res.* **2003,** 36, 199-206.

Lee, S.; Mesleh, M. F.; Opella, S. J., Structure and dynamics of a membrane protein in micelles from three solution NMR experiments. *J. Biomolecular NMR* **2003,** 26, 327-334.

Li, Z.; Glibowicka, M.; Joensson, C.; Deter, C. M., Conformational states of mutant M13 coat proteins are regulated by transmembrane residues. *J. Biol. Chem.* **1993,** 268, 4584-4587.

Lipari, G.; Szabo, A., Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1.theory and range of validity. *J. Am. Chem. Soc.* **1982a,** 104, 4546-4559.

Lipari, G.; Szabo, A., Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.* **1982b,** 104, 4559-4570.

Lodish, H., et. al., Molecular cell biology (4th Edition). *W. H. Freema an Company* **2000.**

Longsworth, L. G., The mutual diffusion of light and heavy water. *J. Phys. Chem.* **1960,** 64, 1914-1917.

Lou, P.; Baldwin, R. L., Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **1997,** 36, 8413-8421.

MacKenzie, K. R.; Prestegard, J. H.; Engelman, D. M., A transmembrane helix dimer. *Science* **1997,** 276, 131-133.

McConnell, H. M., Magnetic Resonance in Biological Systems. *Pergamon, Oxford* **1967.**

Melnyk, R. A.; Partridge, A. W.; Deber, C. M., Transmembrane domain mediated self-assembly of major coat protein subunits from Ff bacteriophage. *J. Mol. Biol.* **2002,** 315, 63-72.

Missel, P. J.; Mazer, N. A.; Carey, M. C.; Benedek, G. B., Influence of alkali-metal counterion identity on the sphere-to-rod transition in alkyl sulfate micelles. *J. Phys. Chem.* **1989,** 93, 8354-8366.

Nakashima, Y.; Wiseman, R. L.; Konigsberg, W.; Marvin, D. A., Primary structure and sidechain interactions of PFL filamentous bacterial virus coat protein. *Nature* **1975,** 253, (5486), 68-71.

Nambudripad, R.; Stark, W.; Opella, S. J.; Makowski, L., Membrane-mediated assembly of filamentous bacteriophage Pf1 coat protein. *Science* **1991,** 252, 1305-1308.

Papavoine, C. H. M.; Christiaans, B. E. C.; Folmer, R. H. A.; Koning, R. N. H.; Hilbers, C. W., Solution structure of the M13 major coat protein in detergent micelles: A basis for a model of phage assembly involving specific residues. *J. Mol. Biol.* **1998,** 282, 401-419.

Pervushin, K., Impact of transverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology. *Q. Rev. Biophys.* **2000,** 33, 161-197.

Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K., Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. USA* **1997,** 94, 12366-12371.

Piotto, M.; Saudek, V.; Sklenar, V., Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **1992,** 2, 661-665.

Popot, J. L.; Engelman, D. M., Helical membrane protein folding, stability, and evolution. *Annu. Rev. Biochem.* **2000,** 69, 881-922.

Quina, F. H.; Nassar, P. M.; Bonilha, J. B. S.; Bales, B. L., Growth of sodium dodecyl sulfate micelles with detergent concentration. *J. Phys. Chem.* **1995,** 99, 17028-17031.

Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K., New perspectives for sequential NMR assignment of proteins using [^{15}N,^{1}H]-TROSY in triple-resonance experiments. *Proc. Natl. Acad. Sci. USA* **1998,** 95, 13585-13590.

Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K., NMR assignment and secondary structure determination of an octameric 110 kDa protein using TROSY in triple-resonance experiments. *J. Am. Chem. Soc.* **2000,** 122, 7543-7548.

Sattler, M.; Fesik, S. W., Use of deuterium labeling in NMR: overcoming a sizeable problem. *Structure* **1996,** 4, (11), 1245-1249.

Schiksnis, R. A.; Bogusky, M. J.; Tsang, P.; Opella, S. J., Structure and dynamics of the Pf1 filamentous bacteriophage coat protein in micelles. *Biochemistry* **1987,** 26, (5), 1373-1381.

Schmidt, P. G.; Kuntz, I. D., Distance measurements in spin-labeled lysozyme. *Biochemistry* **1984,** 23, (18), 4261-4266.

Schmidt, P. G.; Kuntz, I. D., Distance measurements in spin-labeled lysozyme. *Biochemistry* **1984,** 23, 4261-4266.

Schwaiger, M.; Lebendiker, M.; Yerushalmi, H.; Coles, M.; Gröger, A.; Schwarz, C. S., S. ; Kessler, H., NMR investigation of the multidrug transporter EmrE, an integral membrane protein. *Eur. J. Biochem.* **1998,** 254, 610-619.

Shon, K.-J.; Kim, Y.; Colnago, L. A.; Opella, S. J., NMR studies of the structure and dynamics of membrane-bound bacteriophage Pfl coat protein. *Science* **1991,** 252, 1303-1305.

Solomon, I.; Bloembergen, N., Nuclear magnetic interactions in the HF molecule. *J. Chem. Phys.* **1956,** 25, 261-266.

Sreerama, N.; Woody, R. W., Structural composition of beta I and beta II proteins. *Protein Sci.* **2003,** 12, 384-388.

Sreerama, N.; Woody, R. W., Computation and analysis of protein circular dichroism spectra. *Methods Enzymol* **2004,** 383, 318-351.

Stejskal, E. O.; Tanner, J. E., Spin diffusion measurements: spin echoes in the presence of time-dependent field gradient. *J. Chem. Phys.* **1965,** 42, 288-292.

Sternlicht, H.; Wheeler, E., Preliminary magnetic resonance studies of spin labeled macromolecules. In Magnetic Resonance in Biological Systems. *Pergamon, Oxford, U.K.* **1967,** 325-334.

Tamm, L. K.; Hong, H.; Liang, B., Folding and assembly of β-barrel membrane proteins. *Biochimica et Biophysica Acta* **2004,** 1666, 250-263.

Therien, A. G.; Deber, C. M., Oligomerization of a Peptide Derived from the Transmembrane Region of the Sodium Pump γ Subunit: Effect of the Pathological Mutation G41R *J. Mol. Biol.* **2002,** 322, (3), 583-590.

Tjandra, N.; Bax, A., Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **1997,** 278, 1111-1114.

Todd, A. P.; Cong, J.; Levinthal, F.; Levinthal, C.; Hubbell, W. L., Site-directed mutagenesis of colicin E1 provides specific attachment sites for spin labels whose spectra are sensitive to local conformation. *Proteins* **1989,** 6, (3), 294-305.

Vinogradova, O.; Sönnichsen, F.; Sanders, C. R., On choosing a detergent for solution NMR studies of membrane proteins. *J. Biomolecular NMR* **1998,** 4, 381-386.

White, S., http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html *Last modified: 8 June* **2006**.

White, S. H., Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys.*

Biomol. Struct. **1999,** 28, 319-365.

White, S. H.; Ladokhin, A. S.; Jayasinghe, S.; Hristova, K., Minireview: How membrane shape protein structure. *J. Biological Chemistry* **2001,** 276, 32395-32398.

Wider, G.; Dotsch, V.; Wüthrich, K., Self-compensating pulsed magnetic-field gradients for short recovery times. *J. Magn. Reson. Ser. A.* **1994,** 108, 255-258.

Wider, G.; Wüthrich, K., NMR spectroscopy of large molecules and multi-molecular assemblies in solution. *Curr. Opin. Struct. Biol.* **1999,** 9, 594-601.

Williams, K. A.; Glibowicka, M.; Li, Z.; Li, H.; Khan, A. R.; Chen, Y. M.; Wang, J.; Marvin, D. A.; Deber, C. M., Packing of coat protein amphipathic and transmembrane helices in filamentous bacteriophage M13: role of small residues in protein oligomerization. *J. Mol. Biol.* **1995,** 252, 6-14.

Williamson, M. P.; Havel, T. F.; Wüthrich, K., Solution conformation of protein inhibitor IIA from bull seminal plasma by ^{1}H numclear magnetic resonance and distance geometry. *J. Mol. Biol.* **1985,** 182, 295-315.

Wu, D. H.; Chen, A. D.; Johnson, C. S., An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses. *J. Magn. Reson. Ser. A* **1995,** 115, 260-264.

Zhou, F. X.; Cocco, M. J.; Russ, W. P.; Brunger, A. T.; Engelman, D. M., Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat. Struct. Biol.* **2000,** 7, 154-160.

Appendix

1. Pulse program for RAW-BPP-LED

```
/* Water suppression BPP-LED for diffusion measurement */

#include <standard.h>

static int   /* phase cycling */

            phi1[4] = {0,1,2,3},
              rec[4] = {0,2,0,2};

pulsesequence()
{

/* DECLARE AND LOAD VARIABLES */

    double    gt_G1, gt_G2, gt_GH2O, pw_GH2O,
              dt1,dt2,dt4,dt_G1,
              gzlvl_MAX,
              gz_factor,
              gzlvl_G1,
                gzlvl_G2,
              gzlvl_G3,
              gzlvl_GH2O,
              gzlvl_H2O,
              tpwr_GH2O,

              tpwr_G,                  /* power level for Gaussian pulse */
              gt1,
              tpwr_Gf,                /* fine power level for Gaussian pulse */
              pw_G,                   /* power width for Gaussian pulse */

              pwNlvl,
              pwN,
              JNH,

              diff_T,
              lambda;                      /* 1/4J H1 evolution delay */
```

```
/* LOAD VARIABLES */

    gt_G1 = getval("gt_G1");
    gt_G2 = getval("gt_G2");
    gt_GH2O = getval("gt_GH2O");

    tpwr_GH2O = getval("tpwr_GH2O");
    pw_GH2O = getval("pw_GH2O");

    tpwr_G = getval("tpwr_G");              /* Gaussian pulses strength */
    gt1 = getval("gt1");
    tpwr_Gf = getval("tpwr_Gf");            /* fine power for Gaussian pulse */
    pw_G = getval("pw_G");                  /* Gaussian pulse width */

    gzlvl_MAX = getval("gzlvl_MAX");
    gz_factor = getval("gz_factor");
    gzlvl_G2 = getval("gzlvl_G2");
    gzlvl_GH2O = getval("gzlvl_GH2O");
    gzlvl_H2O = getval("gzlvl_H2O");
    gzlvl_G1 = gz_factor*gzlvl_MAX;
    gzlvl_G3 = -1*gz_factor*gzlvl_MAX;

    dt1 = getval("dt1");
    dt2 = getval("dt2");
    dt4 = getval("dt4");
    dt_G1 = getval("dt_G1");

    pwNlvl = getval("pwNlvl");
    pwN = getval("pwN");
    JNH = getval("JNH");

    diff_T = getval("diff_T");                  /* diffusion time */
    lambda = 0.91/(4.0*getval("JNH"));


/* LOAD PHASE TABLE */

    settable(t1,4,phi1);
    settable(t2,4,rec);


/* BEGIN PULSE SEQUENCE */
```

```
rcvroff();
obspower(tpwr);
dec2power(pwNlvl);
txphase(zero);
dec2phase(zero);
delay(d1);                                  /* recycle delay */

obspower(tpwr_G);
obspwrf(tpwr_Gf);
shaped_pulse("Gauss_H2O", pw_G, zero, rof1, rof1);   /* Gaussian pulse */
obspower(tpwr);
obspwrf(4095.0);
zgradpulse(gzlvl_H2O, gt1);
delay(dt1);

rgpulse(pw, zero, 0.0, 0.0);
zgradpulse(gzlvl_G1, gt_G1);
txphase(t1);
delay(dt_G1);
rgpulse(2.0*pw, t1, 0.0, 0.0);
txphase(zero);
zgradpulse(gzlvl_G3, gt_G1);
delay(dt_G1);
rgpulse(pw, zero, 0.0, 0.0);

delay(diff_T);                              /* diffusion delay */

rgpulse(pw, zero, 0.0, 0.0);
zgradpulse(gzlvl_G1, gt_G1);
delay(dt_G1);

rgpulse(2.0*pw, zero, 0.0, 0.0);

zgradpulse(gzlvl_G3, gt_G1);
delay(dt_G1);
rgpulse(pw, zero, 0.0, 0.0);

zgradpulse(gzlvl_G2, gt_G2);
delay(dt2);

rgpulse(pw, zero, 0.0, 0.0);                /* watergate */
txphase(two);
zgradpulse(gzlvl_GH2O, gt_GH2O);
delay(dt2);
```

```
  obspower(tpwr_GH2O);
  rgpulse(pw_GH2O, two, 0.0, 0.0);
  obspower(tpwr);
  txphase(zero);
  rgpulse(2.0*pw, zero, 0.0, 0.0);
  obspower(tpwr_GH2O);
  txphase(two);
  rgpulse(pw_GH2O, two, 0.0, 0.0);

  obspower(tpwr);
  txphase(t2);

  zgradpulse(gzlvl_GH2O, gt_GH2O);
  delay(dt2);

  /*   dec2power(dpwr2);   */
  /* POWER_DELAY */
  rcvron();
 /* statusdelay(C,1.0e-4);   */
  setreceiver(t2);
}
```

2. Pulse program for ^{15}N-edited BPP-LED

/* ^{15}N-edited BPP-LED for diffusion measurement for membrane protein. Modified BPP-LED followed by 1D gNhsqc with Gauss_H2O at the beginning of the pulse program */

```
#include <standard.h>

static int   /* phase cycling */

  phi1[32] = {0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,3,3,3,3},
   phi3[2] = {0,2},
  phi9[32] = {0,0,1,1,2,2,3,3,0,0,1,1,2,2,3,3,0,0,1,1,2,2,3,3,0,0,1,1,2,2,3,3},
   rec[32] = {0,2,2,0,0,2,2,0,2,0,0,2,2,0,0,2,0,2,2,0,0,2,2,0,2,0,0,2,2,0,0,2};

pulsesequence()
{

/* DECLARE AND LOAD VARIABLES */

    double    gt_G1, gt_G2, gt_G4, gt0, gt1,gt5,gstab,
              dt1,dt2,dt4,dt_G1,
              gzlvl_MAX,
              gz_factor,
              gzlvl_G1,
              gzlvl_G2,
              gzlvl_G3,
              gzlvl_G4,
              gzlvl_G5,
              gzlvl0,
              gzlvl1,
              gzlvl2,
              gzlvl5,

              tpwr_G,                  /* power level for Gaussian pulse */
              gt_1,
              tpwr_Gf,            /* fine power level for Gaussian pulse */
              pw_G,               /* power width for Gaussian pulse */
              gzlvl_H2O,

              pwNlvl,
              pwN,
              JNH,
```

```
                tpwrs,
                tpwrsf_t,
                pwHs,
                compH,

                diff_T,
                lambda;                         /* 1/4J H1 evolution delay */

/* LOAD VARIABLES */

    gt_G1 = getval("gt_G1");
    gt_G2 = getval("gt_G2");
    gt_G4 = getval("gt_G4");
    gt0 = getval("gt0");
    gt1 = getval("gt1");
    gt5 = getval("gt5");
    gstab = getval("gstab");

    gzlvl_MAX = getval("gzlvl_MAX");
    gz_factor = getval("gz_factor");
    gzlvl_G2 = getval("gzlvl_G2");
    gzlvl_G4 = getval("gzlvl_G4");
    gzlvl_G5 = getval("gzlvl_G5");
    gzlvl5 = getval("gzlvl5");
    gzlvl2 = getval("gzlvl2");
    gzlvl1 = getval("gzlvl1");
    gzlvl0 = getval("gzlvl0");

    gzlvl_G1 = gz_factor*gzlvl_MAX;
    gzlvl_G3 = -1*gz_factor*gzlvl_MAX;


    tpwr_G = getval("tpwr_G");          /* Gaussian pulses strength */
    gt_1 = getval("gt_1");
    tpwr_Gf = getval("tpwr_Gf");        /* fine power for Gaussian pulse */

    pw_G = getval("pw_G");              /* Gaussian pulse width */
    gzlvl_H2O = getval("gzlvl_H2O");

    dt1 = getval("dt1");
    dt2 = getval("dt2");
    dt4 = getval("dt4");
    dt_G1 = getval("dt_G1");
    pwNlvl = getval("pwNlvl");
```

```
        pwN = getval("pwN");
        JNH = getval("JNH");

        diff_T = getval("diff_T");                    /* diffusion time */

        lambda = 0.91/(4.0*getval("JNH"));

        tpwrsf_t = getval("tpwrsf_t");
        pwHs = getval("pwHs");
        compH = getval("compH");

    /* LOAD PHASE TABLE */

        settable(t1,32,phi1);
        settable(t3,2,phi3);
        settable(t9,32,phi9);
        settable(t2,32,rec);

    /* INITIALIZE VARIABLES */
        tpwrs = tpwr - 20.0*log10(pwHs/(compH*pw*1.69));      /* selective H2O
one-lobe sinc pulse needs 1.69 times more */
        tpwrs = (int) (tpwrs);

    /* BEGIN PULSE SEQUENCE */

        rcvroff();
        obspower(tpwr);
        dec2power(pwNlvl);
        txphase(zero);
        dec2phase(zero);
        delay(d1);                                    /* recycle delay */

        obspower(tpwr_G);
        obspwrf(tpwr_Gf);
        shaped_pulse("Gauss_H2O", pw_G, zero, rof1, rof1);   /* Gaussian pulse */
        obspower(tpwr);
        obspwrf(4095.0);
        zgradpulse(gzlvl_H2O, gt_1);
        delay(dt1);

        rgpulse(pw, zero, 0.0, 0.0);
        zgradpulse(gzlvl_G1, gt_G1);
        delay(dt_G1);
        txphase(t1);
```

```
rgpulse(2.0*pw, t1, 0.0, 0.0);
txphase(zero);
zgradpulse(gzlvl_G3, gt_G1);
delay(dt_G1);
rgpulse(pw, zero, 0.0, 0.0);

delay(diff_T);                          /* diffusion delay */

rgpulse(pw, zero, 0.0, 0.0);
zgradpulse(gzlvl_G1, gt_G1);
delay(dt_G1);

rgpulse(2.0*pw, zero, 0.0, 0.0);
zgradpulse(gzlvl_G3, gt_G1);
delay(dt_G1);

zgradpulse(gzlvl0, gt0);
delay(lambda - gt0);

sim3pulse(2.0*pw, 0.0, 2.0*pwN, zero, zero, zero, 0.0, 0.0);
zgradpulse(gzlvl0, gt0);
delay(lambda - gt0);
txphase(one);                           /* 90 H pulse and its phase is y */
rgpulse(pw, one, 0.0, 0.0);

zgradpulse(gzlvl_G2, gt_G2);
delay(dt2);

dec2phase(t3);
dec2rgpulse(pwN, t3, 0.0, 0.0);         /* in gNhsqc, calN */
dec2phase(t9);

delay(gt1 + 2.0e-4 - 2.0*pw);
dec2rgpulse(2.0*pwN, t9, 0.0, 0.0);
zgradpulse(icosel*gzlvl1, gt1);     /* coherence transfer selection gradient */
delay(2.0e-4 - 2.0*GRADIENT_DELAY);

txphase(zero);
dec2phase(two);

sim3pulse(pw, 0.0, pwN, zero, zero, two, 0.0, 0.0); /* Rance-Kay read-out */

dec2phase(zero);
zgradpulse(gzlvl5, gt5);
```

```
        delay(lambda - 1.3*pwN - gt5);

        sim3pulse(2*pw, 0.0, 2*pwN, zero, zero, zero, 0.0, 0.0);

        zgradpulse(gzlvl5, gt5);
        txphase(one);
        dec2phase(one);
        delay(lambda - 1.3*pwN - gt5);

        sim3pulse(pw, 0.0, pwN, one, zero, one, 0.0, 0.0);

        txphase(zero);
        dec2phase(zero);
        zgradpulse(1.5*gzlvl5, gt5);
        delay(lambda - 1.3*pwN - gt5);

        sim3pulse(2*pw, 0.0, 2*pwN, zero, zero, zero, 0.0, 0.0);

        zgradpulse(1.5*gzlvl5, gt5);
        delay(lambda - 0.65*pwN - gt5);
        rgpulse(pw, zero, 0.0, 0.0);
        delay((gt1/10.0) + 1.0e-4 +gstab - 0.65*pw + 2.0*GRADIENT_DELAY +
POWER_DELAY);
        rgpulse(2.0*pw, zero, 0.0, 0.0);
        txphase(t2);
        dec2power(dpwr2);                              /* POWER_DELAY */
        zgradpulse(gzlvl2, 0.1*gt1);                   /* decoding gradient */

        delay(gstab);
        rcvron();
        statusdelay(C,1.0e-4);
        setreceiver(t2);
```