UNIVERSITY OF CALIFORNIA, SAN DIEGO

# Applications of Nuclear Magnetic Resonance Spectroscopy to Membrane Protein Structure Determination

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Chemistry

by

Sangwon Lee

Committee in charge:

Professor Stanley J. Opella, Chair
Professor Senyon Choe
Professor Patricia A. Jennings
Professor Mauricio Montal
Professor John H. Weare

2007

The dissertation of Sangwon Lee is approved, and it is acceptable in quality and form for publication on microfilm:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2007

# Dedication

To my wife, Jungyuen and my son, Jaden,

who give me endless inspiration and happiness

# Table of Contents

# List of Symbols and Abbreviations

$^{1}H$ (Proton)
$^{2}H$ (Deuterium)
$^{13}C$ (Carbon-13)
$^{15}N$ (Nitrogen-15)
APS (ammonium persulfate)
Cα (alpha carbon)
C' (carbonyl carbon)
CMC (critical micelle concentration)
CSA (chemical shift anisotropy)
CSI (chemical shift index)
Da (1 Dalton = 1 atomic mass unit = mass of $^{1}H$)
DDM (n-dodecyl-β-D-maltoside)
DHPC (1,2-dicaproyl-1-sn-glycero-3-phosphocholine)
DNA (deoxy ribonucleic acid)
DPC (dodecylphosphocholine)
DTT (dithiothreitol)
EDTA (ethylenediamine tetraacetate)
Hα (alpha proton)
HIV (human immunodeficiency virus)
HSQC (heteronuclear single quantum coherence)
IPAP (in-phase / anti-phase)
IPTG (isopropyl β-thiogalactoside)
KSI (ketosteroid isomerase)
LB (Luria-Bertani)
LMPG (1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-RAC-(1-glycerol)])
NMR (nuclear magnetic resonance)
NOE (nuclear Overhauser enhancement)
OG (n-octyl-β-D-glucopyranoside)
PDB (protein data bank)
PISA (polarity index slant angle)
PISEMA (polarization inversion with spin exchange at the magic angle)
RDC (residual dipolar coupling)
RMSD (root mean squared deviation)
SDS (sodium dodecyl sulfate)
SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis)
$\tau_c$ (correlation time)
TEMED (N,N,N',N'-tetramethylenediamine)
TM (transmembrane)
TROSY (transverse relaxation optimized spectroscopy)

# List of Figures

# Acknowledgements

First of all, I would like to thank my advisor, Professor Stanley J. Opella, for giving me the opportunity to work with various projects and supporting me throughout the graduate research. The work presented in this Thesis would have not been accomplished without the help from former and current members of the Opella laboratory and I am grateful for having many interactions with them.

Drs. David H. Jones, Michael F. Mesleh, Alexander A. Nevzorov, and Gabriel A. Cook are thanked for spending infinate amount of time with me (whether it was inside the laboratory or outside) and sharing many thoughts (whether they were about science or life in general). Their endless friendship gave me the strength to move forward. I also want to thank other members in the laboratory for helping me and sharing their ideas; Dr. Anna A. De Angelis, Stanley C. Howell, Dr. Sang Ho Park, Dr. Che (Alex) Ma, Dr. David S. Thiriot, Dr. Ana C. Zeri, Lena Zagyanskyi, Dr. Chin (Albert) H. Wu, Dr. Christopher V. Grant, Dr. Anthony A. Mrse, and Dr. Neeraj Sinha.

Special appreciation goes to people outside the Opella laboratory; Dr. Francesca M. Marassi at the Burnham Institude for supporting and encouraging me; Dr. Timothy A. Cross, Richard C. Page, and Jacob D. Moore at the Florida State University for providing me with their invaluable proteins; Dr. Giuseppe Melacini and Dr. Xuemei Huang for helping me with the solution NMR facility at the UCSD; Professor Patricia A. Jennings and the members of her laboratory for encouragement; Dr. Alan S. Stern at the Rowland Institute for providing RNMRTK package and

giving us some tips on how to use it; Dr. Lewis E. Kay at the University of Toronto for providing the RDCA program; Dr. Logan W. Donaldson at the York University for sharing his CNS package for structure calculation; Dr. Keyang Ding at NIDDK for sharing the pulseprograms to measure RDCs; Dr. Junji Iwahara at NIH for sharing the pulseprograms to measure PREs; as well as my former advisor Professor Yangmee Kim at the Konkuk University, Seoul, Korea, who introduced me the world of NMR and structural biology.

Finally, I would like to thank my family for support and encouragement.

A part of work presented in Chapter 2 was published as "Structure and dynamics of a membrane protein with three experiments" by Lee, S., Mesleh, M.F. and Opella, S.J., *Journal of Biomolecular NMR*, **2003**, 26, 327-334.

# Curriculum Vitae

| | |
|---|---|
| 1996 | Bachelor of Arts, Chemistry, Konkuk University, Seoul, Korea |
| 1998 | Master of Science, Chemistry, Konkuk University, Seoul, Korea |
| 2007 | Doctor of Philosophy, Chemistry, University of California, San Diego |

# Publications

Page, R.C., Moore, J.D., Nguyen, H.B., Sharma, M., Chase, R., Gao, F.P., Mobley, C.K., Sanders, C.R., Ma, L., Sönnichsen, F.D., **Lee, S.**, Howell, S.C., Opella, S.J., Cross, T.A. (2006) Comprehensive evaluation of solution nuclear magnetic resonance spectroscopy sample preparation for helical integral membrane proteins. *Journal of Structural and Functional Genomics* **7**, 51-64.

Kochendoerfer, G.G, Jones, D.H., **Lee, S.,** Oblatt-Montal, M., Opella, S.J., Montal, M. (2004) Functional characterization and NMR spectroscopy on full-length Vpu from HIV-1 prepared by total chemical synthesis. *Journal of the American Chemical Society* **126**, 2439-2446.

Opella, S.J., Park, S., **Lee, S.,** Jones, D.H., Nevzerov, A., Mesleh, M.F., Mrse, A., Marassi, F.M., Oblatt-Montal, M., Montal, M., Strebel, K., Bour, S. (2004) Structure and function of Vpu from HIV-1 in *Viral Membrane Proteins: Structure, Function, and Drug Design*, Fischer W. (Ed), Kluwer, Academic Publishers, New York.

**Lee, S.,** Mesleh, M.F., Opella, S.J. (2003) Structure and dynamics of a membrane protein with three experiments. *Journal of Biomolecular NMR* **26**, 327-334.

Mesleh, M.F., **Lee, S.,** Veglia, G., Thiriot, D.S., Marassi, F.M., Opella, S.J. (2003) Dipolar waves map the structure and topology of helices in membrane proteins. *Journal of the American Chemical Society* **125**, 8928-8935.

**Lee, S.**, Kim, Y. (2004) Molecular dynamics simulations on β amyloid peptide (25-35) in aqueous trifluoroethanol solution. *Bulletin of the Korean Chemical Society* **25**, 838-842.

Lee, K., **Lee, S.,** Kim, Y., Park, N.G. (2003) Structures of neuropeptide gamma from goldfish and mammalian neuropeptide gamma, as determined by H-1 NMR spectroscopy. *Journal of Peptide Research* **61**, 274-285.

Oh, D., Shin, S.Y., **Lee, S.,** Kang, J.H., Kim, S.D., Ryu, P.D., Hahm, K.S., Kim, Y. (2000) Role of the hinge region and the tryptophan residue in the synthetic antimicrobial peptides, cecropin A(1-8)-magainin 2(1-12) and its analogues, on their antibiotic activities and structures. *Biochemistry* **39**, 11855-11864.

Kim, H.S., Park, C.H., Cha, S.H., Lee, J.H., **Lee, S.,** Kim, Y., Rah, J.C., Jeong, S.J., Suh, Y.H. (2000) Carboxyl-terminal fragment of Alzheimer's APP destabilizes calcium homeostasis and renders neuronal cells vulnerable to excitotoxicity. *The FASEB Journal* **14**, 1508-1517.

**Lee, S.**, Kim, Y. (1999) Solution structure of neuromedin B by H-1 nuclear magnetic resonance spectroscopy. *FEBS Letters* **460**, 263-269.

**Lee, S.**, Suh, Y.H., Kim, S., Kim, Y. (1999) Comparison of the structures of beta amyloid peptide (25-35) and substance P in trifluoroethanol/water solution. *Journal of Biomolecular Structure and Dynamics* **17**, 381-391.

Jhon, G.J., Park, S.Y., Han, S.Y., **Lee, S.**, Kim, Y., Chang, Y.S. (1999) Studies of the chemical structure of gangliosides in deer antler, Cervus nippon. *Chemical and Pharmaceutical Bulletin* **47**, 123-127.

Lee, K., **Lee, S.**, Jhon, G.J., Kim, Y. (1998) Tertiary structure of ganglioside GA1 determined by NMR spectroscopy. *Bulletin of the Korean Chemical Society* **19**, 569-575.

Shim, G., **Lee, S.**, Kim, Y. (1997) Motional properties in the structure of GlcNAc(β1,3)Gal(β)OMe studied by NMR spectroscopy and molecular modeling. *Bulletin of the Korean Chemical Society* **18**, 415-424.

**ABSTRACT OF THE DISSERTATION**

Applications of Nuclear Magnetic Resonance Spectroscopy

to Membrane Protein Structure Determination

by

Sangwon Lee

Doctor of Philosophy in Chemistry

University of California, San Diego, 2007

Professor Stanley J. Opella, Chair

Membrane proteins present a significant challenge to structural biology. Although more than 30% of the human genome encodes membrane proteins and about 60% of drugs target them, relatively few structures have been determined, highlighting the need for developing new approaches. An environment suitable for studying membrane proteins requires the presence of lipid molecules to solubilize the proteins, making it difficult to form a suitable crystal or produce resonance signals with narrow linewidths. Also, in most cases, the lack of long-range distance information obtained

from NOE measurements prevents defining the global fold of α-helical membrane proteins.

In this Thesis, methodological developments to determine membrane protein structures using solution NMR spectroscopy and results on proteins with single and multiple hydrophobic trans-membrane domains and amphipathic in-plane helices are described. Measurements of weak anisotropic interactions provide unique structural information that can complement distance information. The potential of the developed methods for high-throughput applications to structural genomics of membrane proteins is tested.

# 1 Introduction

## 1.1 Structures of Membrane Proteins

Membrane proteins are responsible for many important biological functions including those of transporters, pores, receptors, as well as various channels. These proteins mainly function as they transport small molecules or various ions through the membranes and initiate intra- and intercellular signaling pathways or help maintaining other cellular functions. The importance of membrane proteins is partially acknowledged by recognizing that roughly 30% of the genome sequenced to date encodes these kinds of proteins (Wallin and von Heijne, 1998) and more than half the drug targets pursued by pharmaceutical companies are related to these proteins (Torres et al., 2003).

Unfortunately, our understanding of how these proteins carry out their functions is still rather limited, and it is mainly due to the lack of detailed structural information. The structures of membrane proteins have been studied by various techniques, notably cryo-electron microscopy (CryoEM), X-ray crystallography, and nuclear magnetic resonance (NMR) spectroscopy. However, out of about 40,000 protein structures solved by these methods, there are less than 200 unique structures of membrane proteins currently deposited in the Protein Data Bank (PDB) (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html). Some examples of these structures are shown in Figure 1-1.

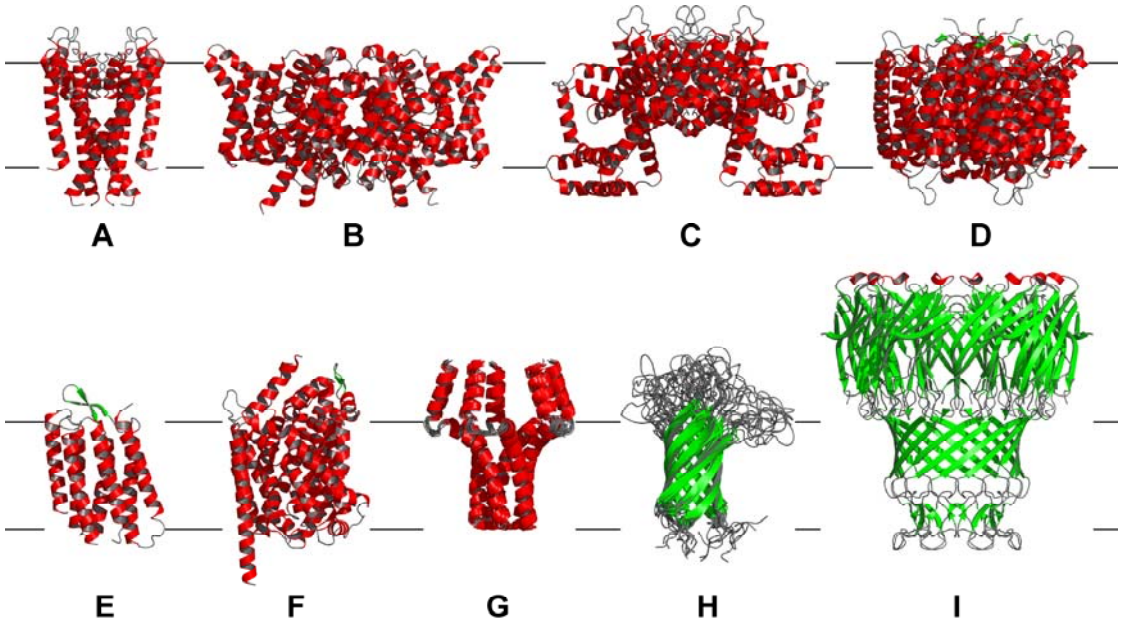


**Figure 1-1.** Some examples of three-dimensional structures of membrane proteins. The coordinates of structures were obtained from the protein data bank (http://www.rcsb.org/pdb/), of which accession codes are indicated within the parenthesis. Ribbon representations of structures were prepared using the software PyMOL (DeLano Scientific LLC, verion 0.99). (A) the 2.0 Å resolution structure of the bacterial $K^+$ selective channel KcsA (Zhou et al., 2001, 1K4C) (B) the 2.5 Å resolution structure of the voltage-gated $Cl^-$ channel (Dutzler et al., 2003, 1OTS) (C) the 3.2 Å resolution structure the voltage-gated potassium channel KvAP (Jiang et al., 2003, 1ORQ) (D) the 1.4 Å resolution structure of the ammonia channel AmtB (Khademi et al., 2004, 1U7G) (E) the 1.6 Å resolution structure of the bacterial light-activated proton pump bacteriorhodopsin (Luecke et al., 1999, 1C3W) (F) the 1.7 Å resolution structure of the Leucine transporter LeuT (Yamashita et al., 2005, 2A65) (G) the NMR ensemble structure of pentameric phospholamban (Oxenoid and Chou, 2005, 1ZLL) (H) the NMR ensemble structure of outer membrane protein OmpA (Arora et al., 2001, 1G90) (I) the 2.5 Å resolution structure of the mycobacterial porin MspA (Faller et al., 2004, 1UUN).

Despite a recent increase in the rate of determining membrane protein structures due to the developments in the fields of molecular biology and biochemistry, and technical advances in synchrotron X-ray beamlines for X-ray crystallography, high-field NMR spectroscopy and high-resolution CryoEM, the rate of membrane protein structure determination is still much slower compared to the rate for the soluble proteins. This is primarily due to the difficulties associated with the

overexpression, purification and refolding process as well as the sample preparation for structural studies in the presence of detergent molecules.

## 1.2 Membrane Protein Expression and Purification Strategy for Structural Studies

High-resolution structural studies of proteins generally require large quantities of highly pure materials. Especially, high resolution structural studies using NMR spectroscopy require yields of milligrams of purified protein per one liter of bacterial cell culture. This is mainly because the expense to use isotope-enriched media to introduce NMR-active nuclei such as $^{15}N$ and $^{13}C$ is relatively high and a sample for NMR experiments still requires high concentration of materials, although there are tremendous technological advances in mechanical development to improve the sensitivity of NMR experiments, such as high-field magnets, cryogenic probes (Kovacs et al., 2005), high temperature superconductors (Brey et al., 2006), and small diameter probes (Olson et al., 1995; Li et al., 2003).

However, overexpression of membrane proteins generally creates a toxic environment to the host cells, yielding extremely low expression level. One of the solutions to overcome the low expression level is to express the desired protein with a "fusion partner". Overexpressions of many membrane proteins have been successfully achieved by cloning the fusion partner with the targeted protein, with the utilized fusion partner being ketosteroid isomerase (KSI, Kuliopulos et al., 1994; Park et al.,

2003; Howell et al., 2005), a portion of Trp ΔLE 1413 (trpΔLE, Miozzari and Yanofsky, 1978; Staley and Kim, 1994), maltose binding protein (MBP, Berthold et al., 2003; Oxenoid and Chou, 2005), glutathione-S-transferase (GST, Smith and Johnson, 1988; Park et al., 2006a), or B-cell leukemia/lymphoma extra long (Bcl-XL, Thai et al., 2005). These fusion tags can be cleaved by enzymes or chemicals if any recognition sites are available in the expressed proteins. By implementing these fusion partners, the expression level and/or the expression locus are controlled by the fusion partners used rather than the target proteins themselves, thus reducing potential toxicity to the host cells.

Directing expressed protein to inclusion bodies by the use of a fusion partner certainly has been the most popular method over the years. This method was proven to overcome low expression levels as well as to achieve the high purity. When the target protein is expressed in the inclusion body, presumably in a highly aggregated state, isolating inclusion body makes it easier to obtain target proteins of high purity. Examples of fusion tag proteins that help the target membrane protein to be expressed in the inclusion body include trpΔLE and KSI. These fusion proteins have been successfully incorporated in a number of membrane proteins for NMR studies, including viral protein u (Vpu) (Ma et al., 2002; Park et al., 2003) from human immunodeficiency virus-1 (HIV-1), bacterial mercury transporter (MerF) (Howell et al., 2005), corticosteroid-hormone induced factor (CHIF), and phospholemman (PLM) (Crowell et al., 2003).

Although many membrane proteins can be expressed in inclusion bodies and

be refolded during the later stage of purifications, the so-called “detergent extraction” method seems attractive for preparing well-folded (and functional) protein samples. The detergent extraction method involves the isolation of the total membrane fraction and dissolving it in a detergent solution that preserve the proper folding and stability. This can be a preferred method of preparing a sample for membrane proteins because of the difficulties associated with the refolding process once they are expressed in the inclusion bodies, which are not necessarily present in the detergent extraction method. The expression of target proteins in the membrane fractions may be facilitated by lowering the post-induction temperature and reducing the rate of expression. When necessary, the detergent used for the solubilization can be exchanged for another through dialysis or during the other purification steps.

Unfortunately, even after devoted efforts over many decades, cloning, expression, purification and refolding protocols of membrane proteins still rely on ‘trial-and-errors’, meaning that the choices regarding host-vector combinations, locations of affinity tags and/or fusion tags, detergents used during the solubilization of membrane fraction, methods of cleaving the fusion tags and separation afterwards, and lipids used to solubilize a protein to fold it correctly have to be made, all of which will affect the efficiency of protein production and the quality of experimental data.

## 1.3 NMR Spectroscopy as a tool for membrane protein structure determination

NMR spectroscopy has progressed rapidly into one of the main methods to determine high resolution structures of biological molecules. The advantage of obtaining the experimental data in conditions close to a physiological environment makes NMR spectroscopy attractive for studying these molecules. Especially for membrane proteins, NMR experiments can be performed in various lipid environments that are necessary to solubilize these molecules. Many years of devoted efforts generated several invaluable systems that permitted the studies of membrane proteins using NMR spectroscopy, including phospholipid bilayers (Opella et al., 2001) and bicelles (Sanders and Schwonek, 1992; Vold et al., 1997; Losonczi and Prestegard, 1998). The possibility of performing experiments in a wide variety of lipids with different head groups and lengths of hydrocarbon chains is a strong point in NMR spectroscopy, because this provides the opportunity to study the structures of this important category of proteins in various environments. However, the method of sample reconstitution in the lipid environment varies depending on the choice of lipids, and different methods of sample preparation necessitate the optimization of different instrumentations and experimental methods. The correlation time and degree of alignment are the main factors that determine the spectroscopic method to use: samples in micelle solutions or in isotropic bicelles have relatively fast correlation time and can be studied by solution state NMR spectroscopy, and samples in large bicelles or in phospholipid bilayers have very slow correlation time with strong

alignment and can be studied by solid state NMR spectroscopy. Recent applications on membrane proteins in large bicelles promised a tremendous potential in structural studies using solid state NMR spectroscopy (De Angelis et al., 2006; Park et al., 2006a). The research described in the current thesis focuses on the methodological developments and applications in solution state NMR spectroscopy.

Traditionally, the primary resource for structure determinations using solution NMR spectroscopy has been $^{1}H$-$^{1}H$ distance restraints converted from measured nuclear Overhauser effects (NOEs) (Wuthrich, 1986). With their intensities being proportional to the inverse sixth power of the internuclear distances, NOEs provide valuable through-space information that correlates nuclei that are close in space. However, the accuracy of NOE measurements decreases as the internuclear distance becomes longer, and the maximum applicable distance limit converted from NOE measurements is around 5 Å, even though a high level of deuteration (Gardner and Kay, 1998) can lead to a small increase in the distance limit by employing longer mixing times (Venters et al., 1995; Mal et al., 1998; Koharudin et al., 2003). As a result, NOEs between the structural domains are extremely rare to observe, simply because the distance between the domains are too far to be detected by conventional NOE measurements, making inter-domain relationships poorly defined. NOEs for α-helical membrane proteins that would define inter-helical distances are extremely difficult to obtain because of similar reasons. In addition, membrane proteins produce broad signals in NMR spectra due to the requisite presence of a lipid environment that causes relatively slow correlation time, and limited chemical shift dispersions because

of the high abundance of hydrophobic residues, all of which contribute to the inability to measure sufficient number of NOEs to define a global fold.

A number of methods have been proposed as alternatives to NOE measurements for structure determination. Especially, recently introduced residual dipolar couplings (RDCs) provide valuable orientational information that is not available otherwise, opening a new realm of research investigations for biomolecular structure determination using solution NMR spectroscopy (Bax et al., 2001). Measurements originating from anisotropic interactions between nuclei such as dipolar couplings and chemical shift anisotropy have been major sources for structure determination using solid-state NMR spectroscopy for quite some time (Opella et al., 1987), however, it was not until very recent years that developments on the theoretical aspects of anisotropic interactions, methods to measure such interactions, as well as interpretation of the measurements became a major issue in the solution NMR community. Nonetheless, such orientational constraints are now routinely measured and incorporated in the structure calculation for globular proteins and their complexes, emphasizing the degree of impact on their structure determination and the need for the developments on the measurements and analysis for membrane proteins.

In this Thesis, methodological developments and applications to determine the structures of membrane proteins using solution NMR spectroscopy are described. These include the developments of weakly aligned systems for measuring RDCs and applications and analysis on membrane proteins with multiple helices. The potential to apply the developments on the structural genomics of membrane proteins is tested.

Finally, the applicability of the developed methods on a large, complicated system is discussed.

# 2 Weakly Aligned System for Membrane Proteins

## 2.1 Introduction to Residual Dipolar Couplings

When two scalar coupled nuclei, such as N-H in the backbone of a protein, are under the influence of an external magnetic field, they behave just like simple bar magnets and their local magnetic fields affect each other. These dipole-dipole interactions, i.e., dipolar couplings, are manifested as the additions to the scalar coupling values and depend on the angle between the external magnetic field and the internuclear vectors in the following form:

$$D^{AB} = -\mu_0 \hbar \gamma_A \gamma_B / (8\pi^2 r_{AB}^3) < 3\cos^2\theta - 1 > \qquad (2.1)$$

where $D^{AB}$ is the dipolar coupling between spin A and spin B; $\theta$ is the angle between the internuclear vector and the external magnetic field; $\mu_0$ is the permittivity of free space; $\hbar$ is Planck's constant; $\gamma_A$ and $\gamma_B$ are the gyromagnetic ratios of nuclear spins A and B, respectively; and $r_{AB}$ is the distance between nuclear spins A and B. Note that the angular brackets in the equation represent time averaging. In isotropic solutions where the molecules are subject to the Brownian motion, the dipolar interaction in Equation (2.1) averages to zero, which results in narrow resonance lines in the spectra. However, when a small amount of anisotropic interaction is introduced to the system, the dipolar coupling in Equation (2.1) becomes non-zero due to incomplete averaging and provides information about the orientation of the AB internuclear vector with respect to the external magnetic field. In solution state NMR spectroscopy, the dipolar

interactions need to be averaged out to about $10^{-3}$ of their static values in order to be detected; hence the dipolar couplings in solution state NMR are called the residual dipolar couplings (RDCs). Assuming the molecule of interest is rigid, Equation (2.1) can be rewritten as

$$D^{AB}(\theta,\phi) = D_a\left[(3\cos^2\theta - 1) + \frac{3}{2}R\sin^2\theta\cos 2\phi\right] \qquad (2.2)$$

where $\theta$ and $\phi$ are polar angles that specify the orientation of the AB internuclear vector with respect to the molecular alignment frame defined in Figure 2-1 (Bax et al., 2001; Prestegard et al., 2000).

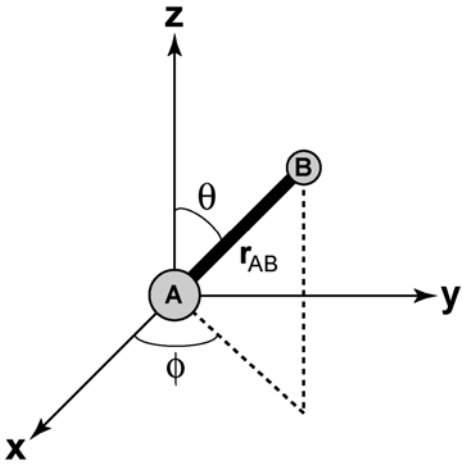


**Figure 2-1.** The definition of polar angles in the alignment frame.

The axial component of the alignment tensor, $D_a$ and the rhombicity, R, can be easily estimated by fitting the orientations of individual internuclear vectors to the dipolar coupling data if prior knowledge of the three-dimensional structure is available. Without structural information, estimation of the alignment tensor is generally carried out by plotting a histogram from the ensemble of dipolar coupling data and fitting it to

a powder pattern, if sufficient number of dipolar coupling data is available (Clore et al., 1998). Once the alignment tensor is defined, the orientation of the each vector can be determined by finding the polar angles from the Equation (2.2). However, it should be noted that, with a single set of dipolar couplings, four alignment frames yield the same dipolar couplings, obstructing the unique definition of orientations of individual internuclear vectors (Skrynnikov et al., 2000). This inherent 4-fold degeneracy can be solved by introducing different alignment media and modulating the molecular alignment (Ramirez and Bax, 1998).

Recently, a method was developed to incorporate the periodicity in the residual dipolar coupling data to the structure determination. This method, Dipolar Wave analysis (Mesleh et al., 2002; Mesleh et al., 2003; Mesleh and Opella, 2003), is a powerful tool to map α-helical structures onto NMR spectra of aligned samples. In solid state NMR, the periodicity of dipolar couplings and chemical shifts caused by anisotropic interactions from α-helical structures in completely aligned samples was recognized in PISEMA (polarization inversion spin exchange at the magic angle, Wu et al., 1994) spectra. The characteristic 'wheel-like' patterns of $^{15}N$-$^{1}H$ static dipolar couplings versus $^{15}N$ chemical shifts observed in samples mechanically aligned in lipid bilayers (Marassi and Opella, 2003) or magnetically aligned in large bicelles (De Angelis et al., 2004; Park et al., 2006b) were the basis of the interpretation of such data, i.e., PISA (polarity index slant angle) wheels (Marassi and Opella, 2000). Dipolar Waves describe the periodicity of dipolar couplings as a function of residue number, and since Dipolar Waves are the direct consequences of periodicity in α-

helices, similar sinusoid patterns are observed when static dipolar couplings from solid state NMR or residual dipolar couplings from solution state NMR for α-helical segments are plotted as a function of residue number, representing the convergence of solid-state and solution NMR approaches to structure determination. Although both PISA wheels and Dipolar Waves originate from the periodic nature of α-helices, the main goals of these approaches are different: while PISA wheels seek for obtaining the spectroscopic measurements, i.e., $^{15}$N-$^{1}$H static dipolar couplings and $^{15}$N chemical shifts for each amino acid residue by analyzing the PISEMA-type spectra, Dipolar Waves aim for extracting structural information *after* the spectroscopic data are measured. A particularly important aspect of the Dipolar Wave approach is that the α-helical segments within the molecule can be detected by searching for sinusoid patterns with periodicity of 3.6 in the dipolar coupling data, providing independent validation of geometry. The absolute orientations of helices with respect to the bilayer normal in completely aligned samples or the relative orientations of helices in the molecular frame in weakly aligned samples can be determined by fitting the dipolar couplings to sinusoids with periodicity of 3.6. Although it is possible to observe similar periodicity in the dipolar couplings measured from different nuclei pairs, such as Cα-Hα and Cα-C', the simplest measurements would be the ones related to N-H pairs.

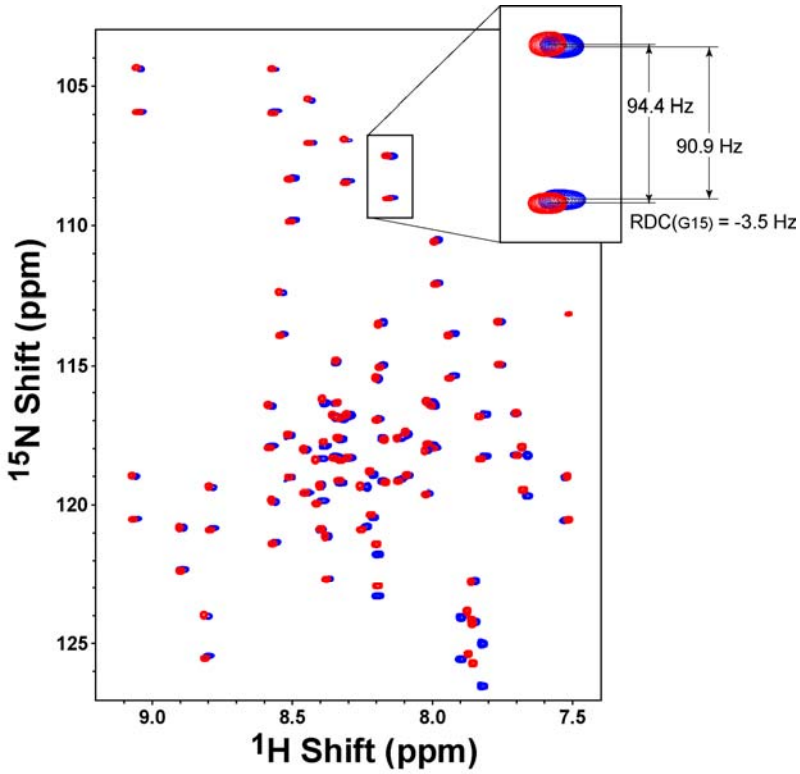


**Figure 2-2.** N-H RDCs for Pf1 coat protein in SDS micelles were measured using $^{15}N,^{1}H$-IPAP-HSQC spectra. This figure is made by overlaying 4 spectra, created by adding and subtracting in-phase and anti-phase spectra for isotropic and weakly aligned samples. Spectra colored red are from isotropic sample and spectra colored blue from weakly aligned sample in 7% polyacrylamide gel. The inset shows the zoomed region of residue G15 illustrating how to extract RDCs from these spectra.

One-bond $^{1}J_{NH}$ couplings can be simply measured by recording $^{15}N,^{1}H$-HSQC spectra without the 180 degree decoupling pulse in the middle of $t_1$ evolution time and computing the separations in $^{15}N$ dimension between the peaks detected from the same N-H pairs. However, doubling of resonances caused by the absence of the decoupling pulse leads the spectral overlapping, resulting an inability to measure most of the couplings. In the IPAP (in-phase/anti-phase) experiments (Ottiger et al., 1998), in-phase and anti-phase components of $^{15}N$ magnetization doublet are recorded separately, and addition and subtraction of in-phase and anti-phase components create

two sub-spectra containing only the upfield or downfield components so that the resulting spectra can be much simpler, increasing the number of measurable couplings. The application of this clever approach was not limited to the N-H pair, and it was adapted to many pulse programs that require more spectral simplicity (Ottiger et al., 1998; Yang et al., 1998; Wang et al., 1998; Permi et al., 2000). An example of N-H RDCs measured using $^{15}$N,$^{1}$H-IPAP-HSQC spectra are illustrated in Figure 2-2.

Modifications to the pulse sequences are necessary for weakly aligned samples in strained polyacrylamide gels (Ishii et al., 2001). Strong background signals appear at around $^{1}$H chemical shifts of 7.8 ppm and 7.0 ppm from $NH_2$ sidechains of the acrylamide. High concentrations of acrylamide in the polyacrylamide gel matrix produce strong background NMR signals in the experiments that utilize the magnetization transfer between $^{15}$N and $^{1}$H, such as $^{15}$N,$^{1}$H-HSQC or $^{15}$N,$^{1}$H-IPAP-HSQC, even though the acrylamide molecules are not isotope-labeled. These natural abundance $NH_2$ signals are often strong enough to prevent some of the resonances from weakly aligned proteins from being analyzed, especially when high polyacrylamide gel concentration is needed, and this modified pulse program effectively suppressed these unwanted signals. More recently, the sensitivity-enhanced version of $^{15}$N,$^{1}$H-IPAP-HSQC was published (Ding and Gronenborn, 2003), and approximately 30% gain in the signal-to-noise ratio was observed when compared to the 'original' $^{15}$N,$^{1}$H-IPAP-HSQC published by Bax group (Ottiger et al., 1998). Interestingly, the theoretical calculation predicted that the signals from $NH_2$ would not be detected from this pulse sequence assuming that $^{1}J_{NH}$ values of backbone amide

groups are similar to those of side chain $NH_2$ groups. Indeed, the $NH_2$ signals from the protein side chain or acrylamide are not detected in the spectra acquired using the sensitivity-enhanced $^{15}N,^{1}H$-IPAP-HSQC experiment.

## 2.2 Development of Weak Alignment

After the first demonstration that weak alignment can be achieved and that RDCs can be accurately measured (Tolman et al., 1995), a wide variety of media have been developed for this purpose, including the dilute liquid crystalline media (Tjandra and Bax, 1997), solutions of the filamentous bacteriophages fd (Clore et al., 1998) and Pf1 (Hansen et al., 1998), mixtures of cetylpyridinium halide and hexanol (Prosser et al., 1998; Barrientos et al., 2000), mixtures of n-hexanol and alkylethylene glycol (Ruckert and Otting, 2000), and purple membrane fragments (Koenig et al., 1999; Sass et al., 1999). The criteria for the ideal anisotropic media for solution NMR spectroscopy can be summarized as the following: (1) the alignment media must have minimal interaction with the biomolecule of interest (2) the alignment media must be nearly homogeneous (3) the alignment media must create sufficiently weak anisotropic environments to produce measurable signals. Unfortunately, most of the alignment media developed for globular proteins is incompatible with membrane proteins because of the strong interaction between the protein-lipid complex and the alignment media itself. A region of metal-binding protein was added to the target protein followed by the binding of lanthanide metals to induce weak alignment (Ma and Opella, 2000), and this method successfully avoided the strong interaction

between the lanthanide ion bound to the metal binding motif, i.e., alignment medium, and the protein-micelle complex. Although some measurable RDCs were obtained by using this method, the high mobility of the motif caused the measured values to have low magnitudes.

The strained polyacrylamide gel was introduced some time ago as one of the weak alignment media (Tycko et al., 2000; Sass et al., 2000), and a wide range of applications have been reported (Chou et al., 2001; Chou et al., 2002; Meier et al., 2002; Lee et al., 2003; Ulmer et al., 2003; Cierpicki et al., 2006). A general procedure to make weakly aligned samples using polyacrylamide gels is illustrated in Figure 2-3.

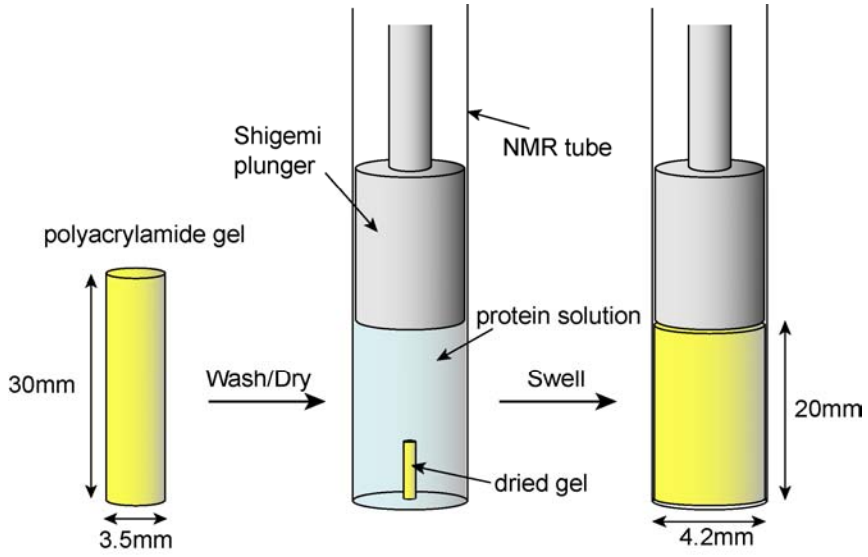


**Figure 2-3.** Graphical representation of a general procedure to make weakly aligned sample using polyacrylamide gels.

Polyacrylamide gels are made from a stock solution containing 30% w/v acrylamide and 0.8% w/v N, N'-methylenebisacrylamide. The polymerization reaction is initiated by adding 0.05% ammonium persulfate (APS) and 0.5%

N,N,N',N'-tetramethylenediamine (TEMED) and is completed within 15 minutes. After casting a polyacrylamide gel in a glass tube with a suitable diameter and let it polymerize, the polymerized gels are taken out, and unreacted acrylamide and crosslinker molecules are dialyzed against water for at least 2 hours. After cut to a desired length, the gels are dried at 35 °C overnight on a glass plate. The dehydrated gel is added to the protein solution in the NMR tube, and left to swell for more than 18 hours. The anisotropy of the protein solution in the polyacrylamide gel matrix is achieved by limiting the final length of the polyacrylamide gel during the swelling process using a Shigemi plunger. This is done by measuring the length of the final solution carefully and fixing the plunger with the Teflon tape and Parafilm before the swelling is complete. Because the final length is smaller than the initial length of the polyacrylamide gel and the final radius is bigger than the initial radius of the polyacrylamide gel, the 'pores' inside the polyacrylamide gel become oblate-shaped. The degree of alignment can be tuned by modifying the 'effective mean pore' of polyacrylamide gels (Holms and Stellwagen, 1991), and this can be achieved by adjusting the total acrylamide concentrations or the ratios between the acrylamide and the crosslinker, N, N'-methylenebisacrylamide.

Other approaches have been developed and published. For example, radial compression was applied instead of axial compression using a funnel-shaped device after swelling to a gel of which initial diameter was larger than the NMR tube (Chou et al., 2001). This method will make the pores inside the polyacrylamide gel prolate-shaped. When successfully applied, the same degree of radial compression

intrinsically yields two-fold stronger alignment than axial compression. The stronger alignment causes the broadening of the signals, and the pore size of the gels needs to be adjusted accordingly. This approach was applied for the viral protein 'u' (Vpu) from HIV-1 described in Chapter 3. The methods to avoid simple diffusion into polyacrylamide gels were also published such as copolymerization of acrylamide with proteins in micelles or diffusion of proteins using electrophoresis in an NMR tube (Jones and Opella, 2004).

These approaches can be applied to solutions with any kind of biomolecules if the solute is minimally interacting with the polyacrylamide gel matrix. In the case of membrane proteins, the protein solution must contain lipid or detergent molecules to solubilize the protein, and polyacrylamide gel matrix is effectively inert to the lipid or the detergent molecules so that the interaction between the protein-micelle complex and the gel matrix can be minimal. Figure 2-2 shows indirectly that the quality of NMR spectra is not perturbed after the protein-micelle complex is effectively aligned in a gel. It is noteworthy that, since it is critical to make a homogeneous aligned protein sample, it is also crucial to make a homogeneous polyacrylamide gel itself before it is incubated with the protein solution. Using a siliconized glass tube when casting the gel and a siliconized glass plate during the drying process makes the surface of the polymerized gel smooth and prevents the formation of air bubbles inside the dehydrated gels. The air bubbles inside the dehydrated gels might cause uneven swelling of the protein solution, resulting in broad linewidths or multiple peaks in the spectra.

The most important factor in making homogeneous aligned samples using the simple diffusion method would be the degree of swelling. In the case of globular proteins, the degree of swelling would primarily depend on the size of the protein itself, however, in the case of membrane proteins, the necessity to use a micelle solution to solubilize the protein adds complication: the solvent condition would affect the swelling ability in addition to the size of the protein. To optimize the swelling ability, or the degree of swelling, of different micelle solutions, a series of swelling tests were done with three detergents that are commonly used in NMR studies, i.e., sodium dodecyl sulfate (SDS), dodecylphosphocholine (DPC), and 1,2-dicaproyl-1-sn-glycero-3-phosphocholine (dihexanoyl phosphatidylcholine, DHPC). The polyacrylamide gels containing 5% w/v total acrylamide concentration and 0.07% w/v N, N'-methylenebisacrylamide (estimated effective mean pore radius of ~150 nm, Holms and Stellwagen, 1991) were used for these tests, and the results are summarized in Table 2-1. A total of 42 polyacrylamide gels were cast, cut to the same length before they were dried, then incubated with the micelle solutions in different conditions. The ratios between the lengths before and after the incubation with micelle solutions were measured.

This study provides the general information on what condition the NMR sample should have in order to make weakly aligned sample using polyacrylamide gels. The results in Table 2-1 indicate that, at least for the kinds of micelle solutions used in the tests, the concentration of the micelle solution is the most critical factor for the swelling of the polyacrylamide gel, which is somewhat expected. Elevated

temperature did not affect the relative ratio of swelling between different samples, but in some cases (especially with high concentration of detergents) it showed small increase in the swelling amount. It is interesting to find that there is a limit on the concentration of micelle solution that fully diffuses into a gel and that this concentration varies among different kinds of micelles. Previous studies suggested that the sample should contain adequately high detergent concentrations in order to get high-quality NMR spectra (McDonnell and Opella, 1993). The current test suggests the upper limit of the detergent concentrations that can be used in the isotropic sample to make a weakly aligned sample in a polyacrylamide gel, complementing the previous study which suggests the lower limit of detergent concentrations. This test also suggests that similar kinds of tests must be preceded when the detergent needed for an NMR sample is not listed in Table 2-1.

**Table 2-1.** Swelling abilities of detergent solutions to 5% polyacrylamide gels characterized by the percentage of the lengths. Final lengths were measured after 18 hours incubation at room temperature.

| | No Buffer No NaCl | 100mM Na Phosphate pH 5.8 | 100mM Na Acetate pH 4.0 | 100mM NaCl | 300mM NaCl | No Buffer Acidic soln. | No Buffer Basic soln. |
|---|---|---|---|---|---|---|---|
| 200mM SDS | 83.3 % | 92.6 % | 94.4 % | 92.6 % | 100 % | 70.4 % | 96.3 % |
| 500mM SDS | 88.9 % | 100 % | 98.1 % | 74.1 % | 81.5 % | 74.1 % | 77.8 % |
| 100mM DPC | 100 % | 100 % | 100 % | 100 % | 100 % | 100 % | 100 % |
| 300mM DPC | 85.2 % | 81.5 % | 79.6 % | 81.5 % | 85.2 % | 85.2 % | 85.2 % |
| 100mM DHPC | 100 % | 100 % | 88.9 % | 96.3 % | 100 % | 100 % | 100 % |
| 200mM DHPC | 81.5 % | 85.2 % | 85.2 % | 85.2 % | 81.5 % | 81.5 % | 85.2 % |

## 2.3 Applications to the Model Systems

The applicability of the residual dipolar coupling measurements of membrane proteins and the following Dipolar Wave analysis was tested using the model systems: Pf1 coat protein (Lee et al., 2003) and fd coat protein (Mesleh et al., 2003).

*Pf1 coat protein - Structure and Dynamics from Three Experiments*

In this section, the structure and dynamics of the membrane-bound form of Pf1 coat protein in SDS micelles are described with the results of three NMR experiments, two of which, the $^{15}N,^{1}H$-HMQC-NOESY and heteronuclear $^{1}H$-$^{15}N$ NOE experiments, have been applied previously to Pf1 coat protein in DPC micelles. The structure and dynamics of Pf1 coat protein in dodecylphosphocholine (DPC) micelles and in phospholipid bilayers have been reported; in these environments, Pf1 coat protein shows characteristics of a typical membrane protein with a long hydrophobic transmembrane helix and a short amphipathic in-plane helix (Schiksnis et al, 1987; Shon et al., 1991). The dynamics of the protein also have been described by molecular dynamics simulations and measurements on relaxation and line shapes (Tobias et al., 1995; Shon et al, 1991). However, the chemical shift differences between spectra obtained in DPC micelles and SDS micelles are sufficient to require the reassignment of the resonances. The sequential assignment of all backbone amide resonances was accomplished with a single two-dimensional $^{15}N,^{1}H$-HMQC-NOESY experiment (Shon and Opella, 1989) on a uniformly $^{15}N$ labeled isotropic sample. The heteronuclear $^{15}N$-$^{1}H$ NOE experiment is sensitive to the internal dynamics of

membrane proteins (Bogusky et al., 1987), because of their distinctive pattern of structured helical segments and mobile terminal and connecting loop segments. The results of these two experiments are complemented by measurements of RDCs from an $^{15}N,^{1}H$-IPAP-HSQC experiment (Ottiger et al., 1998).

Uniformly $^{15}N$-labelled Pf1 coat protein was prepared as described previously (Schiksnis et al., 1987). The weakly aligned sample was prepared by soaking a solution of 1 mM Pf1 coat protein in 500 mM SDS and 40 mM NaCl into a 7% acrylamide gel made from a stock solution containing 30% w/v acrylamide and 0.8% w/v N, N'-methylenebisacrylamide. The polyacrylamide gel was cast in a 3.5 mm-inner diameter glass tube, washed for about two hours, cut to a length of 25 mm, and dried at 35 °C overnight. The protein solution was then soaked into the dry gel overnight with the length of the gel restricted to 18 mm. Isotropic samples were prepared in the same way except for being cast in a 4.2 mm tube and having no restriction on the length of gel. The pH of the samples was adjusted to 4.0, and all experiments were performed at 313K.

The NMR experiments were performed on Bruker Avance 750 and DRX 600 spectrometers. The two-dimensional $^{15}N,^{1}H$-HMQC-NOESY spectrum was obtained on an isotropic sample using a 250 msec mixing time. The steady-state heteronuclear $^{15}N$-$^{1}H$ NOE measurements were also made on this sample using the experiment described by Farrow et al. (1994) with and without $^{1}H$ irradiation time of 2.2 seconds. $^{15}N,^{1}H$-IPAP-HSQC spectra (Ottiger et al., 1998) obtained with suppression of the $NH_2$ signals (Ishii et al., 2001) on isotropic and weakly aligned samples of protein-

SDS micelle complex were used to measure the N-H residual dipolar couplings. For the sample in the polyacrylamide gel, it was necessary to use a version of the $^{15}N,^{1}H$-IPAP-HSQC experiment modified for the suppression of the $NH_2$ resonances from acrylamide (Ishii et al., 2001). While the positions of the peaks shift slightly in the uncompressed gel samples of the protein relative to the isotropic samples (Figure 2-2), this does not affect the measurement of the N-H RDCs.

The amide region of the two-dimensional $^{15}N,^{1}H$-HMQC-NOESY spectrum of uniformly $^{15}N$ labeled Pf1 coat protein in SDS micelles is shown in Figure 2-4. The $^{1}H$-$^{1}H$ homonuclear mixing interval of 250 msec enabled the observation of cross-peaks from all of the backbone HN-HN NOEs between adjacent residues in the protein. The sequential NOE connectivities used to assign the resonances are represented with lines between cross-peaks. We previously applied this experiment to Pf1 coat protein in DPC micelles (Shon et al, 1991) and observed NOEs only between residues in the helical segments. The results obtained in SDS micelles differ in that it is possible to observe NOE connectivities between residues in the loop connecting the hydrophobic and amphipathic helices in addition to those in the helices. These additional NOEs observed among backbone amide sites in SDS micelles made it possible to 'walk' through the entire protein sequence and assign all of the backbone amide resonances with one experiment on a uniformly $^{15}N$ labeled sample. Similar results were obtained from a $^{15}N,^{1}H$-HSQC-NOESY experiment modified from two-dimensional FHSQC (Mori et al., 1995) with the expected improvement in the resolution in the indirect dimension. As the structures of an increasing number of membrane proteins are being

determined in various micelle, bicelle, and bilayer preparations, differences are being found in the structure and dynamics of residues that reside near the lipid head groups. In the longer term, this may become a compelling argument to perform all structure determinations of membrane proteins in lipid bilayers, which are the closest mimic to naturally occurring membranes. However, both micelles and bicelles provide model membrane environments compatible with native, functional membrane proteins and are useful for some NMR approaches to structure determination. Therefore, at present, there is considerable interest in determining structures in micelles by solution NMR spectroscopy, and comparisons between structures determined in various lipid environments will also be of interest until the effects of lipids on protein structure are better characterized.

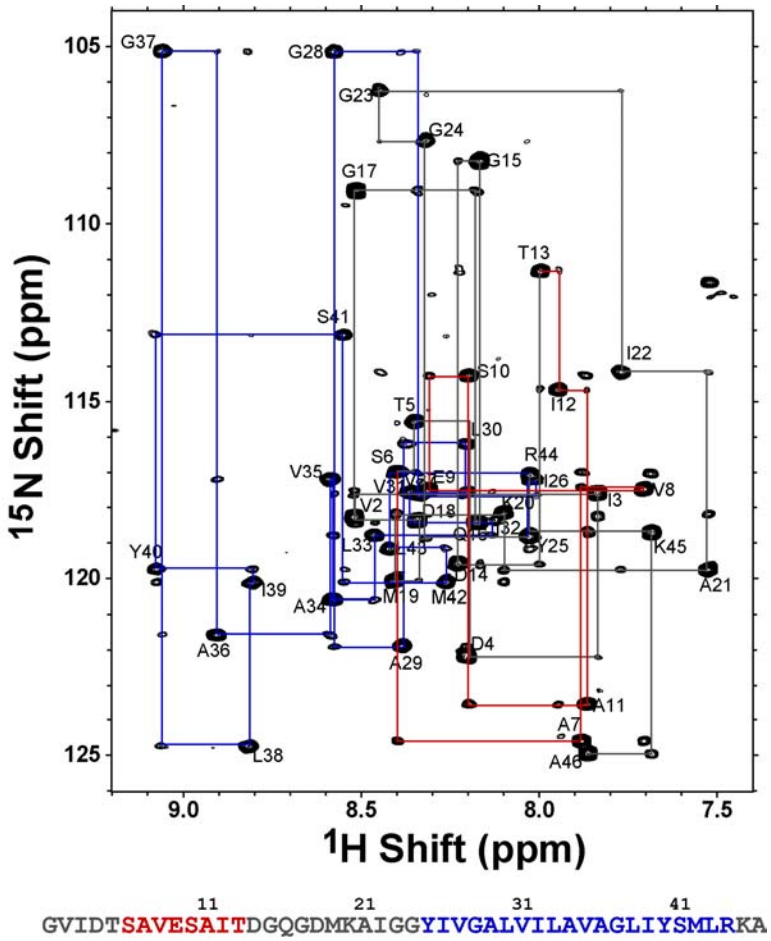


**Figure 2-4.** The resonance assignments of Pf1 protein in SDS micelles were complete using a single $^{15}N,^{1}H$-HMQC-NOESY. Lines connecting each of the resonances show the sequential HN-HN NOE connectivities between adjacent amino acid residues, of which colors indicate the regions of the amino acid sequence shown at the bottom. The spectrum was obtained using a mixing time of 250 ms, and a recycle delay of 1.5 with 512 t1 increments.

Backbone structure and dynamics are more obviously coupled in membrane proteins than they are in some other classes of proteins, and the heteronuclear $^{1}H$-$^{15}N$ NOE provides a remarkably direct and sensitive indicator of the local dynamics of membrane proteins in micelles (Bogusky et al., 1987). The experimental $^{1}H$-$^{15}N$ heteronuclear NOEs for all backbone amide sites of the membrane-bound form of Pf1

coat protein in SDS micelles are shown in Figure 2-5 (A). The values are taken from steady-state heteronuclear $^{1}H$-$^{15}N$ NOE spectra acquired using the method of Farrow et al. (1994) with and without 2.2 s of irradiation to saturate the $^{1}H$ magnetization. The regions with the slowest correlation times, as reflected in the positive $^{1}H$-$^{15}N$ heteronuclear NOE values, are highly correlated with the locations of the helices. The data in Figure 2-5 (A) show 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. These data also indicate that the amphipathic helix on the surface of the micelles has some motions that are not present in the trans-membrane hydrophobic helix. The motions of the amphipathic helix have been described in some detail for the membrane-bound form of fd coat protein (Almeida and Opella, 1997).

Dipolar waves are a direct consequence of the fact that the periodicity of α-helices maps to the measurement of anisotropic parameters in NMR spectroscopy (Mesleh et al., 2002; Mesleh et al., 2003; Mesleh and Opella, 2003). Figure 2-5 (B) plots the experimental $^{1}H$-$^{15}N$ RDCs of the backbone amide sites of Pf1 coat protein as a function of residue number. The fits of the RDCs to sinusoids with a periodicity of 3.6 residues characteristic of a α-helix are also shown. The quantitative fit of two sine waves demonstrates that Pf1 coat protein has two α-helical segments, between residues 6 and 13 and between residues 24 and 44. The average values and the magnitudes of the dipolar couplings of the waves fit to the amphipathic helix region and the transmembrane helix region are distinctly different, and can be used to

determine the relative orientations of the two helices in the same molecular frame of reference (Mesleh et al., 2002).

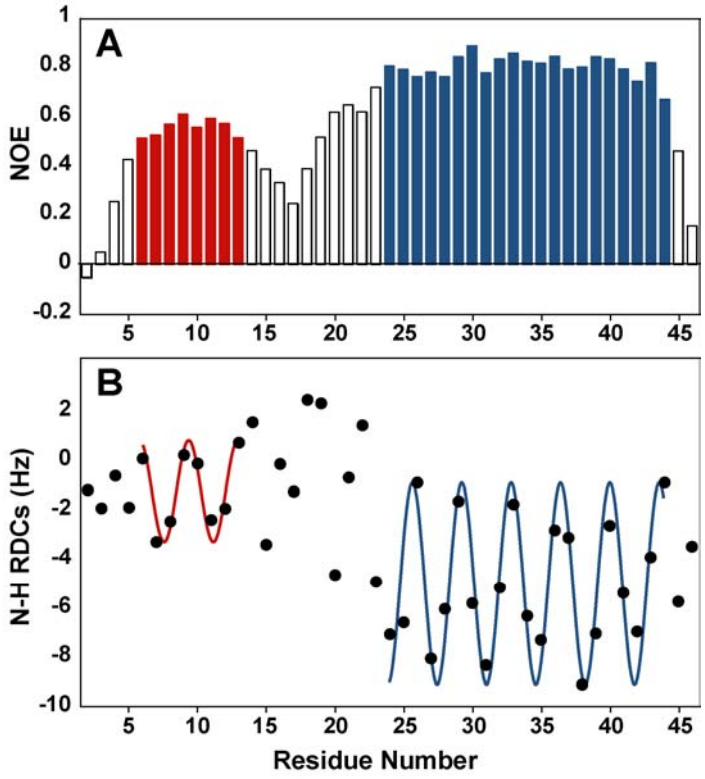


**Figure 2-5.** (A) Steady-state heteronuclear $^{1}H$-$^{15}N$ NOE data for Pf1 as a function of residue number along the sequence. (B) A plot of RDCs versus residue numbers for Pf1 in SDS micelle. Solid line represents the fitting of RDCs using sinusoid function (red; amphipathic helix, blue; trans-membrane helix).

Comparisons between measurements on isotropic and weakly aligned samples enable a remarkable amount of information to be obtained from just three experiments on a membrane protein in micelles. The sequential resonance assignments are obtained from the homonuclear $^{1}H$-$^{1}H$ NOEs observed in a single two-dimensional $^{15}N$,$^{1}H$-HMQC-NOESY spectrum, since the amide hydrogens in the polypeptide backbone are proximate in helices and in the turn between helices. The heteronuclear

$^{1}H$-$^{15}N$ NOEs differentiate between residues in structured helices and more mobile loop and terminal segments. Complementary information about structure (time average positions) and dynamics (time dependent positions) emerges from the first two experiments, and set the stage for the orientational parameters from the third experiment. The experiments directly show the positions of the helices in the sequence and their relative orientations with intervening and terminal residues that display greater mobility indicative of internal dynamics that are completely consistent with the previous molecular dynamics calculations (Tobias et al., 1995).

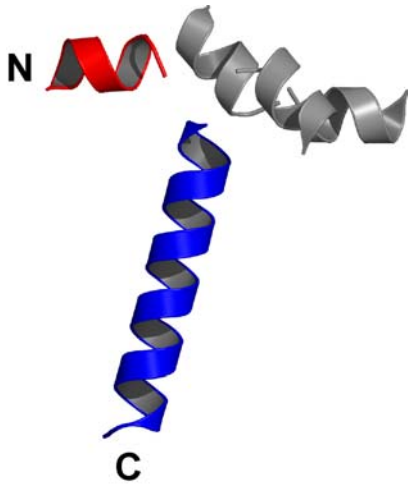


**Figure 2-6.** Backbone representation of the fold of the membrane-bound form of Pf1 coat protein in SDS micelles determined by the experimental data in Figures 2-5. Four possible orientations of in-plane helix respect to the transmembrane helix are shown. Selected model based on the solid-state NMR data is colored in red for in-plane helix and blue for transmembrane helix.

The Dipolar waves derived from the measurements of RDCs obtained by comparison of isotropic and weakly aligned samples, measured with the $^{15}N$,$^{1}H$-IPAP-HSQC experiment provide structural information in the form of the relative

orientations of the helices in the molecular frame. A structural model of the protein was built based on this orientational information, and it is presented in Figure 2-6. Due to angular degeneracies inherent in the use of RDCs, there are four families of structures defined by the four symmetry-related arrangements of the helices consistent with the experimental data. However, the correct arrangement of helices is determined by reference to the solid-state NMR data on a completely aligned sample in lipid bilayers (Shon et al., 1991), since this provides absolute orientational information.

*fd coat protein*

The fd coat protein is a typical 50-residue membrane protein with a long hydrophobic transmembrane helix, a shorter amphipathic in-plane helix connected to the transmembrane helix by a turn or loop, and a few mobile C- and N-terminal residues. The structure of this protein has been extensively studied in micelles (Cross and Opella, 1980; McDonnell et al., 1993; Almeida and Opella, 1997), lipid bilayers (Leo et al., 1987; Bogusky et al., 1988; Marassi et al., 1997; Marassi and Opella, 2003), and bacteriophage particles (Zeri et al., 2003) using NMR spectroscopy.

The previously determined structure of fd coat protein in SDS micelles by solution NMR spectroscopy primarily used distance restraints from NOE measurements that might be considerably time-consuming. RDC measurements in polyacrylamide gel followed by Dipolar Wave analysis will provide the orientational information that can be used in the structure determination instead of distance

information from NOE measurements that are not readily available. The applicability of the method on this protein is tested.

Solution NMR samples of the fd coat protein were made with 1 mM $^{15}$N-labeled protein in 500 mM SDS, 40 mM NaCl at pH 4.0. The samples of the fd coat protein in the SDS micelles were weakly aligned in a 7% w/v polyacrylamide gel containing 0.19% w/v N, N'-methylenebisacrylamide by soaking the protein solution into the dried gel and limiting the length of expansion in the NMR sample tube. N-H RDCs were measured by calculating the difference in one bond J couplings for backbone amide groups measured using $^{15}$N,$^{1}$H-IPAP-HSQC experiments (Ottiger et al., 1998) for an isotropic sample and an anisotropic sample. A modified $^{15}$N,$^{1}$H-IPAP-HSQC pulse program (Ishii et al., 2001) was used for the sample aligned in the polyacrylamide gel, which is designed for suppressing signals from $NH_2$ groups, permitting the amide resonances at around 7.8 ppm to be measured.

The experimental RDC data and fits to the sinusoids shown in Figure 2-7 (A) gives the secondary structures and relative orientations of the helices in the fd coat protein in SDS micelles. The regions having low RMSDs to the sinusoids with the periodicity of 3.6 suggest the presence of $\alpha$-helical secondary structures, and three such regions were identified for the fd coat protein in micelles by the Dipolar Wave analysis (Mesleh et al., 2002), two of which (H2 and H3) compose the transmembrane region. Because the Dipolar Wave analysis provides the structural information obtained from the dipolar couplings, quick and direct comparison between the structures from different environments is possible using the Dipolar Wave analysis,

when such experimental dipolar couplings are available. This is particularly interesting point for the fd coat protein, since a vast amount of NMR data has been accumulated for this protein. When N-H RDC data obtained from the weakly aligned samples in SDS micelles was compared to static N-H dipolar coupling data obtained from the completely aligned samples in the lipid bilayer, the Dipolar Wave analysis indicated that this protein has very similar properties in these environments. For example, the lengths of each of the helical segments, as detected by the small deviations of the dipolar coupling data from the sinusoid function with the periodicity of 3.6, are nearly identical in micelles and bilayers. As shown by the previous report (Mesleh and Opella, 2003), the Dipolar Wave analysis is a sensitive method to detect the helical regions within the protein, as well as the deviations from an ideal α-helical geometry such as kinks and curvature. Characterized by the change in the phase and the average amplitude over a small number of residues, a kink in the middle of the transmembrane helix was readily identified at the residue $Gly^{38}$. Interestingly, the same kink is found in the previously determined structures in the bilayer (Marassi and Opella, 2003), micelles (Almeida and Opella, 1997) and bacteriophage particles (Zeri et al., 2003).

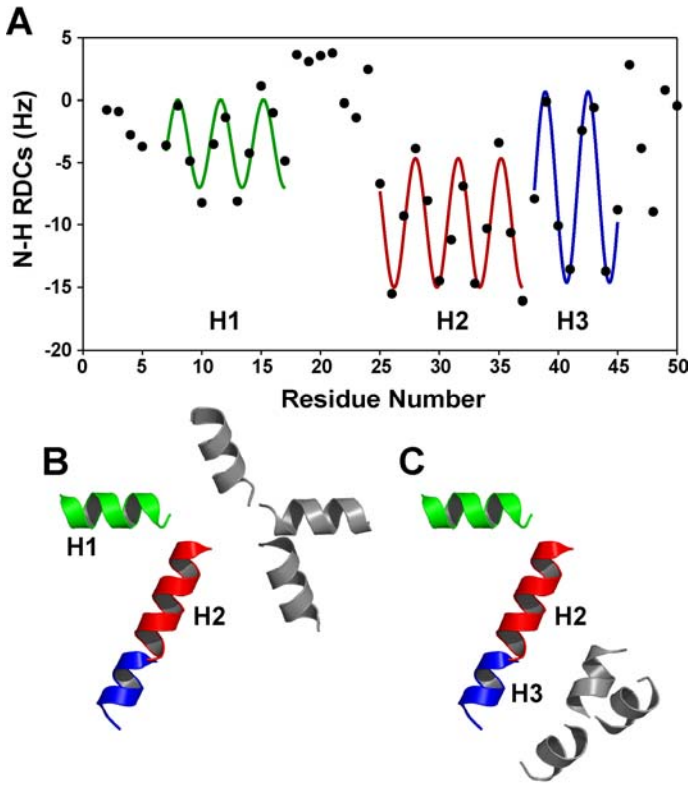


**Figure 2-7.** Structural information obtained from residual dipolar couplings for the fd coat protein in SDS micelles. (A) Plot of N-H RDCs as a function of residue numbers. The lines represent the fitted sinusoids for the individual helical segments. (B), (C) Models of the fd coat protein consistent with the data shown in (A). Degenerate structures are drawn for (B) helix 1 or (C) helix 3, when helix 2 is overlaid. Helical segments colored grey are discarded based on the structures determined previously.

The inherent four-fold degeneracy of RDC measurements yields a total of 16 possible models in the case of the fd coat protein, which has 3 helical segments, and this is illustrated in Figure 2-7 (B) and (C). The transmembrane domain with a kink at the residue $Gly^{38}$ is treated as two helical segments in this case. When helix 2 of each degenerate structure is overlaid, four possible orientations are obtained for helix 1, as shown in Figure 2-7 (B). Likewise, four possible orientations can be obtained for helix 3 when helix 2 is overlaid, as shown in Figure 2-7 (C). Although intuitive

knowledge can eliminate some of the possibilities of orientations especially for the helices separated by a few residues, the correct geometry can be obtained from additional information. In this case, of course, the correct geometry of the helical segments for the fd coat protein can be selected from the degenerate structures depicted in Figure 2-7 (B) and (C) by comparing them to the previously determined high-resolution structure (Marassi and Opella, 2003; Almeida and Opella, 1997).

The optimization of RDC measurements for membrane proteins and application on the model systems described in the current chapter provides a basis on how to measure anisotropic interactions using solution NMR and interpret such measurements. Dipolar Waves proved to be a sensitive method to detect multiple helices and obtain relative orientations within the molecule, which can be generally applied to membrane proteins. After the successful applications of developed methods on the model systems, the attention was shifted toward more challenging systems with biological significance.

A part of work presented in Chapter 2 was published as "Structure and dynamics of a membrane protein with three experiments" by Lee, S., Mesleh, M.F. and Opella, S.J., *Journal of Biomolecular NMR*, **2003**, 26, 327-334.

# 3 Residual Dipolar Coupling measurements of Vpu

## 3.1 Introduction

Human immunodeficiency virus (HIV) is a retrovirus that causes acquired immunodeficiency syndrome (AIDS), damaging the immune system and making the patients vulnerable to opportunistic infections. The Joint United Nations Programme on HIV/AIDS (http://www.unaids.org) and the World Health Organization estimate that, as of December 2006, a total of 39.5 million people are living with HIV worldwide. In 2006 alone, 4.3 million people were newly infected with HIV and about 3 million people died of AIDS-related illnesses.

HIV is a spherical shape particle with a diameter of roughly 120 nm. The virus particle consists of a conical capsid containing two viral RNA molecules along with a nucleocapsid and enzymes such as reverse transcriptase, protease, and integrase. The capsid is enclosed by matrix proteins which are surrounded by glycoproteins, gp120 and gp41, embedded in the phospholipid membranes. As HIV infects macrophage and CD4+ T cells, the glycoproteins on the surface of the virus particle bind the CD4 receptors as well as the chemokine coreceptors on the surface of the target cells. This is followed by fusion of the viral envelope with the cell membranes and the release of the viral capsid into the cells. The capsid is uncoated to release a viral nucleoprotein complex, making genomic single-stranded RNA reverse transcribed into double-stranded viral DNA, which is then transported into the cell nucleus where the viral

DNA is integrated into the host genome. After viral transcription, the RNA is spliced and transported into the cytoplasm. Viral mRNAs are translated and packaged, and the assembled core particle begins to bud from the cell surface. To assist the budding process, CD4 receptors degrade to release the coexpressed glycoproteins. The surface of the virion particle is coated with the glycoprotein complex as the particle is released from the cell surface and subject to a maturation process, during which polyproteins are cleaved into functional proteins and enzymes by HIV protease (reviewed in Frankel and Young, 1998). Although substantial efforts have been made toward developing and designing a vaccine for AIDS, no definite cure for HIV infection or AIDS has been found.

Viral protein "u" (Vpu) is an 81-residue membrane protein, and it is one of the six accessory proteins encoded by the HIV-1 genome (Strebel et al., 1988; Cohen et al., 1988). The term 'accessory proteins' comes from initial studies that showed their absence has little effect on viral replication; however, subsequent studies have shown that they can dramatically change the course and severity of viral infection (Bour and Strebel, 2000), thus have potential as therapeutic targets (Yu et al., 2005). Structural studies on Vpu have shown that it contains a single transmembrane helix and two amphipathic helices (Marassi et al., 1999). Interestingly, the two biological functions of Vpu are correlated with two distinct structural domains: the cytoplasmic domain of Vpu is responsible for CD4 degradation in the endoplasmic reticulum (ER), and the transmembrane domain of Vpu is related to the budding of new virus particles from the infected cells (Bour et al., 1995; Schubert et al., 1996a,b). When phosphorylated

by casein kinase II (Schubert et al., 1994), two conserved serine residues in the cytoplasmic domain, $Ser^{52}$ and $Ser^{56}$, are critically important for CD4 degradation through interaction with the human beta transducin-repeat containing protein (βTrCP, Margottin et al., 1998). However, virus release is only partially affected by mutation of these serine residues and requires the presence of the transmembrane domain (Schubert et al., 1996b; Paul et al., 1998). The ion-channel activities observed by single channel measurements on various Vpu constructs suggests that the transmembrane domain of Vpu forms the ion-conducting pores (Marassi et al., 1999; Ma et al., 2002; Montal, 2003), although how this activity is connected to virus release by the transmembrane domain of Vpu is still unclear. The three-dimensional structure of Vpu could lead to a better understanding of its mechanism, and eventually help designing an effective vaccine for HIV-1.

## 3.2 Sample Preparation and NMR Spectroscopy

Cloning, overexpression and purification of various Vpu constructs followed the protocol described previously (Marassi et al., 1999; Ma et al., 2002). Briefly, the cDNA encoding the target protein was ligated into the expression vector pMMHa containing a fusion partner trpΔLE (Miozzari and Yanofsky, 1978; Stanley and Kim, 1994) with a hexahistidine tag. After transformation into *E. coli* strain BL21(DE3) (EMD Biosciences, La Jolla, CA), the stable transformed clone was screened for high-level expression. Bacterial cell culture was performed using M9 minimal media

containing $^{15}$N-$(NH_4)_2SO_4$ and/or $^{13}$C-glucose as a sole nitrogen and/or carbon source. Expression of the target protein was achieved by inducing with 0.4 mM of isopropyl-β-D-thiogalactoside (IPTG) and incubating for four hours at 37 °C. The cells were subsequently harvested by centrifugation at 6,000 rpm for 30 min at 4 °C. The target proteins attached to the fusion partner trpΔLE are expressed as insoluble aggregates, i.e., inclusion bodies, and are not toxic to *E. coli* cells, allowing high-level expression. After isolating the inclusion body fraction, the fusion protein was purified with nickel affinity chromatography in denaturing conditions. Vpu was separated from the fusion partner by cyanogen bromide (CNBr) cleavage and further purified using reverse-phase high-pressure liquid chromatography (HPLC) or the combination of ion exchange chromatography, gel filtration, and HPLC.

Isotropic samples for Vpu were made by adding 400 mM 1,2-dicaproyl-1-sn-glycero-3-phosphocholine (DHPC) stock solution to the lyophilized protein powder and diluting the solution with water to the final detergent concentration of 100 mM. The pH of the sample was adjusted to 4.0 with a small amount of either hydrochloric acid or sodium hydroxide. Anisotropic samples were made by in two different ways: protein solution was weakly aligned in compressed and stretched polyacrylamide gels. For a sample in a compressed polyacrylamide gel, the polyacrylamide gel with 6% w/v total acrylamide concentration was cast in a medium-walled NMR tube (524-PP, Wilmad-LabGlass, Buena, NJ) with an inner diameter of 3.5 mm, cut to a length of 3 cm, dialyzed against water overnight, and dried overnight at 35 °C on a siliconized glass plate. The protein-micelle solution was then incubated with the dried gel for

about 2 days and the final length of the gel was set to 2 cm using the Shigemi plunger fixed with Teflon tape and Parafilm while the protein solution diffused into the gel. For radial compression, a funnel-shaped device (Chou et al., 2001) with an inner diameter of 5 mm was used to cast a polyacrylamide gel with 6% w/v total acrylamide concentration. After polymerization was complete, the gel was cut to a length of 2 cm, dialyzed against water, and dried overnight at 35 °C. The gel was then incubated with the protein solution in the funnel-shaped device for more than two days, and the gel containing the protein solution was 'squeezed' into the open-ended NMR tube using a piston with a plug.

NMR experiments were performed using a Bruker DRX600 with a TXI-cryoprobe, a Magnex 750 with a Bruker AVANCE console or a Magnex 800 with a Bruker AVANCE console at the University of California, San Diego. N-H RDCs were measured using a sensitivity-enhanced version of $^{15}N,^{1}H$-IPAP HSQC (Ding and Gronenborn, 2003). All NMR data were processed with NMRPipe (Delaglio et al., 1995) and analyzed with SPARKY (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco).

## 3.3 RDC Measurements and Dipolar Wave Analysis

Despite its relatively small size, Vpu challenges the structural studies because of the inability to form crystals for x-ray crystallography or broad resonance linewidths in NMR spectra. Because structural domains of Vpu are functionally independent, polypeptides containing the amino acid sequence regions of different

structural domains were expressed in *E. coli* and purified to aid the structure determination of the full-length protein. NMR samples for Vpu constructs containing just the cytoplasmic domain, $Vpu_{28-81}$, just the transmembrane domain, $Vpu_{2-30+}$, as well as the full-length protein, $Vpu_{2-81}$ were prepared with a high level of purity, and the $^{15}N,^{1}H$-HSQC spectra of the truncated constructs showed no significant chemical shift differences from the full-length protein, suggesting that the local environments of each domain are not affected by the other (Marassi et al., 1999; Ma et al., 2002; Opella et al., 2004).

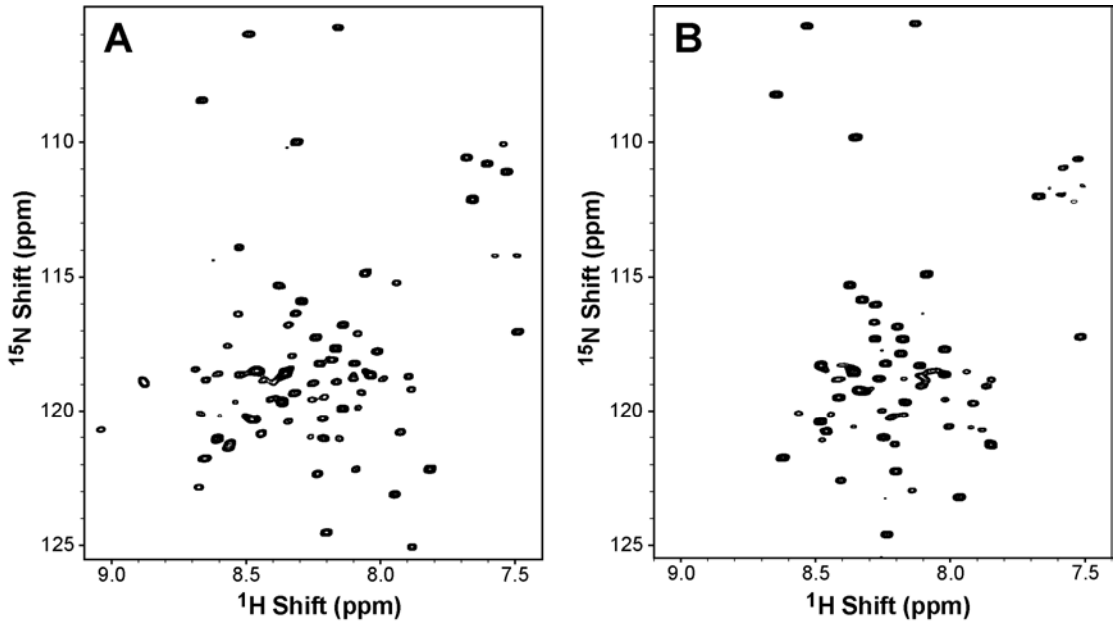


**Figure 3-1.** $^{15}N,^{1}H$-HSQC spectra of (A) $Vpu_{2-81}$ and (B) $Vpu_{28-81}$ in 100 mM DHPC micelles. Spectra were obtained at 323K, pH 4.0. Note that the majority of the resonances from $Vpu_{28-81}$ overlap with the resonances for $Vpu_{2-81}$, suggesting no variation in the local environments between these constructs.

The measurement of RDCs is a key step toward the structure determination of Vpu using solution state NMR spectroscopy, as described in Chapter 2. Although significant efforts have been made to measure any "long-range NOEs" that would

define the position of each helix, only small- and medium-range NOEs were observed, which help define the secondary structure elements. Inability to extract long-range distance information has lead to the development of new methods to determine the structure, i.e., measuring and analyzing RDCs. Orientational information obtained from dipolar coupling measurements play a critical role in membrane protein structure determination either by solid state or solution state NMR spectroscopy (De Angelis et al., 2005). A high-resolution structure of the trans-membrane domain of Vpu, $Vpu_{2-30+}$, was determined using the static dipolar couplings measured in various phospholipid bilayers or large bicelles (Park et al., 2003; Park and Opella, 2005; Park et al., 2006b), however, the orientations of cytoplasmic domain helices with respect to the lipid bilayer normal still need to be determined. For solution NMR spectroscopy, dipolar coupling measurements require the development of suitable alignment media that do not interfere with the interaction between the protein and micelle solution yet sufficiently align the protein. A region of metal-binding protein was added to utilize the lanthanide-induced alignment to obtain dipolar couplings (Ma and Opella, 2000). Although this method successfully avoided unnecessary interactions between the protein and the alignment center as well as any structural perturbation that might occur when the extrinsic metal-binding motif is attached, the high mobility of the motif causes a small degree of alignment, as reflected by the small magnitude of measured RDCs.

Strained polyacrylamide gels (Tycko et al., 2000; Sass et al., 2000) proved to be a weak alignment media that can be applied to many membrane proteins in micelles

(Chou et al., 2002; Lee et al., 2003; Mesleh et al., 2003; Cierpicki and Bushweller, 2004; Oxenoid and Chou, 2005). However, optimizing the anisotropic sample conditions for Vpu was not a trivial process. Incomplete soaking of Vpu in DHPC micelles into different concentrations of polyacrylamide gels resulted in samples with heterogeneous environments, yielding poor spectral quality. Polymerization of acrylamide in the presence of protein in micelles (Jones and Opella, 2004) was not applicable to Vpu because its behavior is highly sensitive to pH and the presence of TEMED lead to the precipitation. After realizing the heavy dependence of detergent concentrations on the degree of soaking into dried polyacrylamide gels (see Table 2-1), the concentration of DHPC in the NMR sample was lowered to 100mM from 175mM that was previously used. This simple but critical finding lead to high-quality anisotropic samples that yielded reproducible RDC measurements.

Figure 3-2 shows N-H RDCs for $Vpu_{28-81}$, plotted as a function of residue number. The data colored in blue (named N-H $RDC_{compressed}$ hereafter) are obtained using a polyacrylamide gel with axial compression, i.e., a 'compressed' polyacrylamide gel, as an alignment medium whereas the data colored in red (named N-H $RDC_{stretched}$ hereafter) are obtained using a polyacrylamide gel with radial compression, i.e., a 'stretched' polyacrylamide gel. The data clearly suggest that there are two distinct helical regions in the cytoplasmic domain, as judged by low deviations from the fitted sinusoids with the exception being a few residues of which the resonances are overlapped with signals from other residues in the NMR spectra. The helical regions detected with this method are from residues $Arg^{30}$ to $Asp^{51}$ and $Glu^{59}$ to

Met[70], and these helical regions are separated by a linker region containing two phosphorylation sites, Ser[52] and Ser[56]. Remarkably, the sinusoids fitted to N-H $RDC_{compressed}$ and the sinusoids fitted to N-H $RDC_{stretched}$ are mirror images about an RDC value of zero. Furthermore, close inspection of the data reveals that even the data for the regions corresponding to the linker and termini from both alignments are mirror images to each other, although they are not fitted to the sinusoid function with the periodicity of 3.6. This suggests that the values for these regions are highly reproducible and the N-H vectors within these regions have preferred orientations.

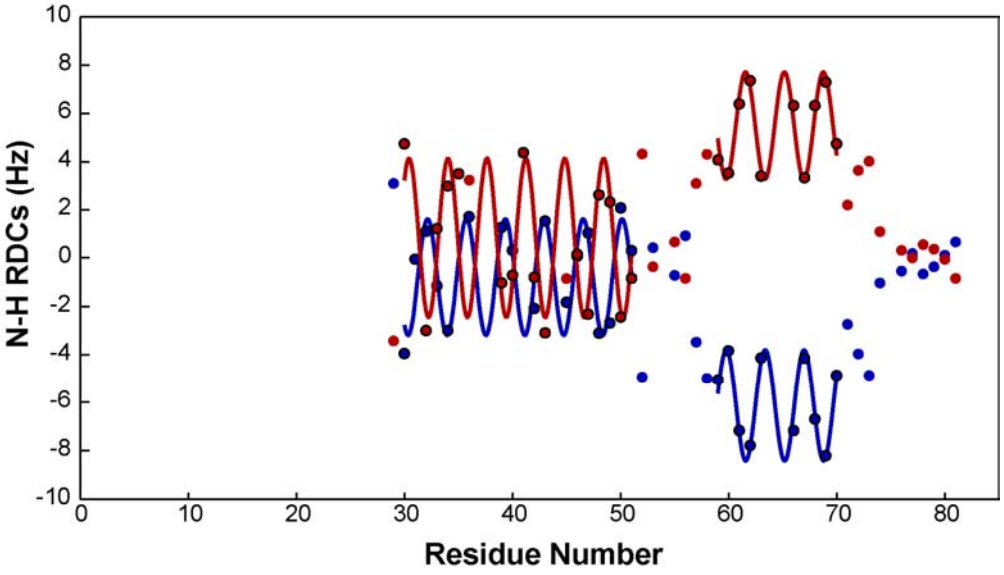


**Figure 3-2.** N-H RDCs for $Vpu_{28-81}$ are plotted as a function of residue number. The data colored in blue are RDCs measured using a polyacrylamide gel with axial compression, i.e., a 'compressed' polyacrylamide gel, as an alignment medium, whereas the data in red are measured using a polyacrylamide gel with radial compression, i.e., a 'stretched' polyacrylamide gel, as an alignment medium. Solid lines represent fitted sinusoids for helical regions.

Relative orientations of each of the helices can be determined using Dipolar Wave analysis (Mesleh et al., 2003; Mesleh and Opella, 2003), and the structures

calculated by this method are shown in Figure 3-3. Due to the inherent 4-fold degeneracy of the alignment axis calculated using RDC measurements from a single set of alignment (Skrynnikov et al., 2000; Al-Hashimi et al., 2000), four relative orientations for two helical segments are possible. Interestingly, no significant differences are observed when the calculated structures from each data set are compared; degenerate orientations of the second helix with respect to the first helix calculated using N-H $RDC_{compressed}$ overlap with the degenerate orientations calculated using N-H $RDC_{stretched}$. This result is not unexpected because the two data sets mirrored each other about zero are an effect of a simple 180 degree rotation of an alignment axis, becoming one of the degenerate alignment axes. Therefore, although this is an intriguing result, the combination of current data did not resolve 4-fold degeneracy. Manipulating the molecular alignment tensor by adding electrostatic interactions, i.e., including various charged acrylates, in the polyacrylamide gel matrix was inapplicable to Vpu due to protein precipitation. Making an assumption of possible electrostatic interactions between the lipid head groups and the charged residues in amphipathic helices may extract a plausible structure out of these degenerate structures; however, further experimental data will be required to support the assumptions.

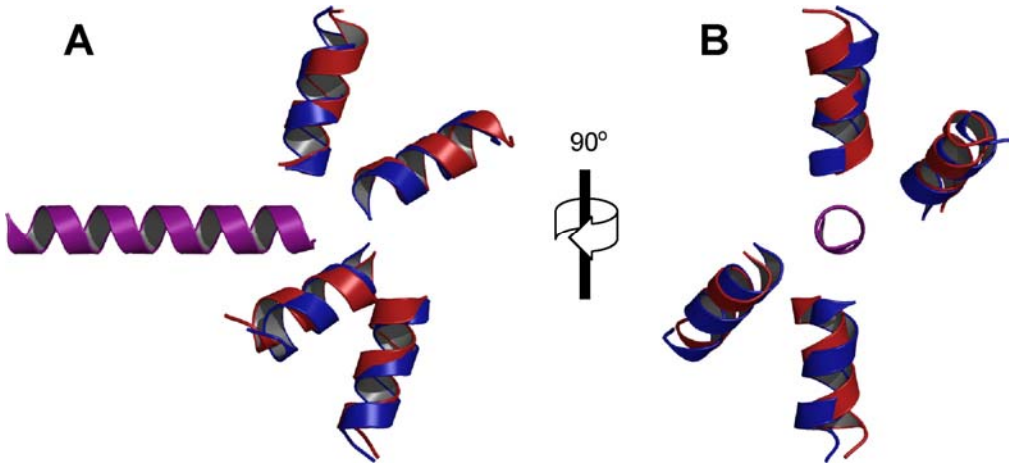


**Figure 3-3.** The structures of $Vpu_{28-81}$ calculated using the N-H RDCs shown in Figure 3-2. Degenerate orientations of the second helix calculated using the data from a 'compressed' gel are colored in blue, and degenerate orientations calculated using the data from a 'stretched' gel are colored in red. The structures are overlaid with respect to the first helix, colored in purple.

Obtaining multiple "unique" alignments will be an important and difficult task to determine the structure of Vpu. Developing improved alignment media for membrane proteins and optimizing sample conditions for creating unique alignments needs to be continued. N-H RDCs for $Vpu_{2-81}$ were measured using the same method as $Vpu_{28-81}$, and the same helical regions within the cytoplasmic domain in addition to a transmembrane helix were detected for $Vpu_{2-81}$ by the Dipolar Wave analysis. However, similar 4-fold degeneracy, yielding 16 possible orientations of helices, prevented defining the structure for $Vpu_{2-81}$. The phosphorylated form of $Vpu_{28-81}$ (p-$Vpu_{28-81}$) was separately prepared by phosphorylating purified $Vpu_{28-81}$ using casein kinase 2 and further purified by HPLC. N-H RDCs of p-$Vpu_{28-81}$ were then measured using the same technique as described above, after the backbone resonances were assigned using triple resonance experiments. However, similar orientational

ambiguities were also found for p-$Vpu_{28-81}$, making it impossible to compare the structures of $Vpu_{28-81}$ and p-$Vpu_{28-81}$.

The information obtained from RDC measurements, as can be seen from the Equations (2.1) and (2.2), does not contain translational information; therefore, the locations of helical regions in Figure 3-3 were arbitrarily set. The development and optimization of sample conditions to obtain experimental data that contain the translational information are necessary to complement the orientational information from RDC measurements to extract the correct fold from the degenerate structures. Acquiring experimental data for $Vpu_{2-81}$ or $Vpu_{28-81}$ using solid state NMR spectroscopy to define the absolute orientation of the helices in the cytoplasmic domain will also need to be continued.

# 4 The Structure of Rv1761c, a Membrane Protein from *Mycobacterium tuberculosis* - An Example on Rapid Structure Determination of Helical Membrane Proteins

## 4.1 Introduction

Tuberculosis (TB) is a deadly, highly contagious, air-born disease caused by mycobacteria, primarily *Mycobacterium tuberculosis* (*M. tuberculosis*). It is estimated that about one-third of the world's population is currently infected with TB bacillus with new infections occurring at the rate of one per second and approximately 2 million deaths are caused by TB every year (WHO facts sheet, 2006). This is not confined to developing nations because people with compromised immune systems due to immunosuppressive drugs, substance abuse, or HIV are also at risk. Since the complete genomic sequence of the best characterized strain of *M. tuberculosis*, H37Rv, has been determined (Cole et al., 1998), a structural genomics consortium was formed with the goal of determining the three-dimensional structures of all proteins from *M. tuberculosis* (http://www.doe-mbi.ucla.edu/TB/). This information can subsequently be used to develop novel approaches for the diagnosis and treatment of TB. As with most structural genomics initiatives, initial target selection process for this consortium chooses globular proteins and eliminates the proteins that are expressed in the inclusion bodies or membrane fractions. The *M. tuberculosis* genome contains as

many as 1200 open reading frames coding for membrane proteins and this eliminates a significant number of proteins that mediate many important cellular functions including those of ion channels, various transporters and receptors. To provide the structural information on these proteins, the structural genomics consortium specifically for the membrane proteins from *M. tuberculosis* was separately organized (http://magnet.fsu.edu/~mprotein/MBPweb/TBgrant.html). The goal of this consortium is not only to solve the structures of membrane proteins discarded in the precedent consortium but to develop requisite techniques specific for membrane proteins. These techniques can subsequently be applied to all structural genomics initiatives. Figure 4-1 illustrates the general strategy used in the current study of NMR-based structural genomics for membrane proteins.

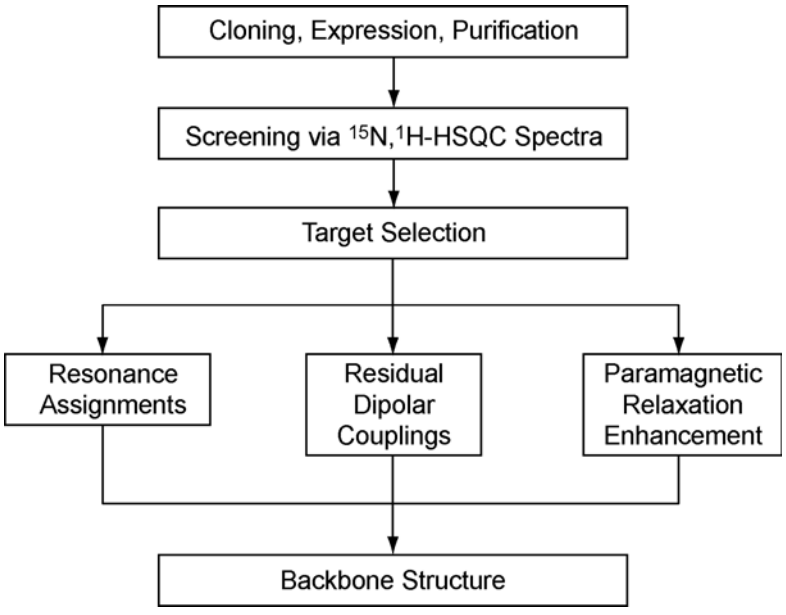


**Figure 4-1.** Overview of the strategy applied in structural genomics for membrane proteins using NMR spectroscopy.

There has been tremendous progress toward developing methods for rapid structure determination using solution NMR spectroscopy over the last five years. The time required for acquiring the data was significantly reduced using methods such as non-linear sampling schedules followed by maximum entropy reconstruction (Hoch and Stern, 2001; Rovnyak et al., 2004), projection-reconstruction (Kupce and Freeman, 2004), G-matrix Fourier Transform NMR spectroscopy (Kim and Szyperski, 2003), and covariance NMR spectroscopy (Bruschweiler and Zhang, 2004). Also, residual dipolar coupling (RDC) (Bax et al., 2001) measurements were developed to replace or complement the distance measurements from 'long-range' NOEs that are hard to obtain or sometimes not available at all. Specific applications of RDC measurements for helical membrane proteins were developed and Dipolar Wave analysis, the fitting of measured RDC values to sinusoid pattern as a function of residue number, represents a powerful method for extracting the structural information for the helical segments within a protein molecule (Mesleh et al., 2003; Mesleh and Opella, 2003). The applicability of this method was tested as described in Chapter 2. However, as discussed in Chapter 3, additional measurements must be made to complement the RDC measurements because of the ambiguities inherent with angular constraints. Paramagnetic relaxation enhancement (PRE) measurements (Battiste and Wagner, 2000; Donaldson et al., 2001; Liang et al., 2006; Iwahara et al., 2003) have become a powerful method for improving the quality of structures or to determine the structure with limited data. The combination of RDCs and PREs has the potential to facilitate the rapid structure determination of helical membrane proteins, and consequently to

play a crucial role in the initial stage of structural genomics by NMR. The purpose of the work presented in this chapter, at least in part, is to test the possibility of rapid structure determination of a membrane protein, using Rv1761c from *M. tuberculosis* as a model system.

## 4.2 Introduction to Paramagnetic Relaxation Enhancement

Paramagnetism has long been recognized by the NMR community (Kosen, 1989), but the development of practical paramagnetic probes and the applications for proteins that do not possess an intrinsic paramagnetic center have not progressed until very recent. The growing number of successful applications of this powerful approach has been assisted by the development of methods for introducing paramagnetic probes, measuring the spectroscopic information, as well as the utilization of the information in structural calculations (Hubbell et al., 1998; Donaldson et al., 2001; Battiste and Wagner, 2000; Iwahara et al., 2004).

The relationship between the transverse PRE ($R_2^{sp}$) and the inter-spin distance is expressed in Equation (4.1), the Solomon-Bloembergen equation (Solomon and Bloembergen, 1956):

$$R_2^{sp} = \frac{1}{r^6}\left[\frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \hbar^2 \gamma_H^2 g_e^2 \beta^2 S(S+1)\left(4\tau_c + \frac{3\tau_c}{1+\varpi_H^2\tau_c^2}\right)\right]$$

$$= \frac{K}{r^6}\left(4\tau_c + \frac{3\tau_c}{1+\varpi_H^2\tau_c^2}\right) \qquad (4.1)$$

where r is the distance between the unpaired electron and the nuclear spin; $\mu_0$ is the vacuum permeability; $\hbar$ is Plank's constant; $\gamma_H$ is the gyromagnetic ratio of proton; $g_e$ is the electron g factor; $\beta$ is the Bohr magneton; S is the spin number for proton; $\omega_H$ is the Larmor frequency of the proton nuclear spin; $\tau_c$ is the correlation time for the electron-nuclear interaction; and K is the constant, $1.23 \times 10^{-44}$ $m^6$ $s^{-2}$. The technique has been developed and optimized for either calculating or measuring $R_2^{sp}$. Pulse sequences have been developed for the direct measurements of the transverse relaxation rates for amide protons (Donaldson et al., 2001; Iwahara et al., 2003) or aliphatic protons (Iwahara et al., 2004), as well as the longitudinal relaxation rates for amide protons (Iwahara et al., 2003). The caveat of these measurements is the requirement of long experimental time, not only because of the low sensitivity but also because of the need for acquiring many time points to obtain $T_2$, since this is done by fitting the peak intensities to the exponential decay. This was partly solved by manipulating the equation for the exponetial fit so that two time points are sufficient to determine resonably accurate $T_2$ values (Iwahara et al., 2004).

An alternative approach is to measure the intensities of the $^{15}N,^{1}H$-HSQC spectra for a paramagnetic sample, $I_{\text{paramagnetic}}$, and a diamagnetic sample, $I_{\text{diamagnetic}}$, and convert their intensity ratios into $R_2^{sp}$ using the following relationship (Battiste and Wagner, 2000):

$$\frac{I_{\text{paramagnetic}}}{I_{\text{diamagnetic}}} = \frac{R_2 \exp(-R_2^{sp} t)}{R_2 + R_2^{sp}} \qquad (4.2)$$

where $R_2$ is the intrinsic transverse relaxation rate and t is the total evolution time used during the INEPT period in the $^{15}N,^{1}H$-HSQC experiment. The intrinsic $R_2$ can be estimated from the half-height linewidth. The estimation of $R_2^{sp}$ from the intensity ratios of $^{15}N,^{1}H$-HSQC spectra, followed by the conversion into the distance restraints to calculate structures has been successfully applied to several proteins (Battiste and Wagner, 2000; Liang et al., 2006).

Application of these Equations (4.1) and (4.2) requires prior knowledge regarding the intrinsic transverse relaxation rate $R_2$, calculated from half-height linewidth, and the global correlation time $\tau_c$. These values were systematically varied in Figure 4-2 where calculated distances were plotted against the intensity ratios from the $^{15}N,^{1}H$-HSQC spectra. Figure 4-2 (A) plots the calculated distance for several different linewidths and clearly demonstrates that the distances are not strongly dependent on the measured linewidths. However, when plotting the calculated distance at a fixed linewidth and varying the global correlation time, as in Figure 4-2 (B), there is a significantly larger effect on the distances calculated from the transverse PREs. This illustrates the requirement for careful examination/interpretation of the dynamics when utilizing Equation (4.1).

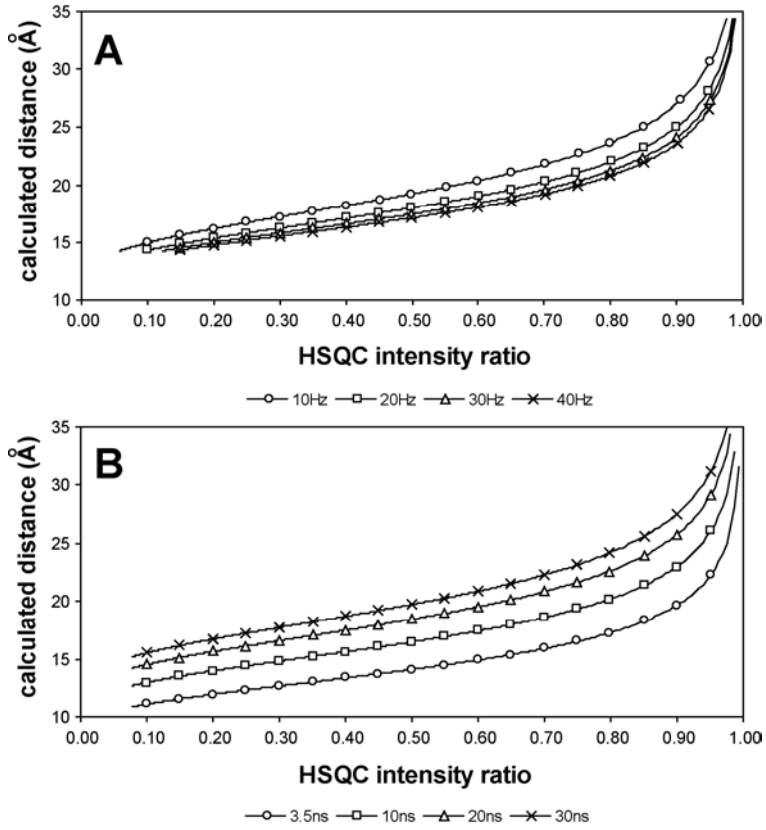


**Figure 4-2.** Distances calculated from HSQC intensity ratios by Equation (4.1) and (4.2). (A) with various linewidths and a fixed global correlation time, τc, of 20 ns or (B) with various τc and a fixed linewidth of 15 Hz.

## 4.3 Materials and Methods

*Sample preparation*

Rv1761c was cloned, expressed, and purified as published (Korepanova et al., 2005; Page et al., 2006). For weakly aligned samples, the protein solution in dodecylphosphocholine (DPC) micelles was soaked into dried polyacrylamide gels with 5.5% total acrylamide and 0.15% w/v N,N'-methylenebisacrylamide (Lee et al., 2003; Jones and Opella, 2004) made from a 30% stock solution (National Diagnostics,

Atlanta, GA). Negatively charged polyacrylamide gels were prepared by mixing the neutral acrylamide solution with a solution containing 2-acrylamido-2-methyl-1-propanesulfonic acid (AMPS) and N,N'-methylenebisacrylamide. Polymerization reaction was initiated by adding 0.05% ammonium persulfate (APS) and 0.5% N,N,N',N'-tetramethylenediamine (TEMED) in medium-walled NMR tubes (524-PP, Wilmad-LabGlass, Buena, NJ) siliconized with SIGMACOTE (Sigma-Aldrich, St. Louis, MO). The polymerized gels were removed from the tube by applying gentle pressure, cut to a desired length and dialyzed against water for 2 days to remove unreacted reagents. After dehydration for one day at 37 °C on glass plates pretreated with SIGMACOTE, the dried gel was incubated with the protein-micelle solution. Axial compression was achieved by limiting the final length of the polyacrylamide gel during the swelling process.

Site-directed mutagenesis was performed to introduce cysteine residues using the standard protocol. Three different cysteine mutants were generated: F30C, S10C, and S48C, which represent the cysteine-substituted constructs on residues $Phe^{30}$, $Ser^{10}$, and $Ser^{48}$, respectively. Spin labeling was performed on the cysteine mutants using (1-oxyl-2,2,5,5-tetramethyl-$\Delta^3$-pyrroline-3-methyl)methanethiosulfonate (MTSL, Toronto Research Chemicals Inc., Canada) as well as a diamagnetic analog of MTSL: (1-acetyl-2,2,5,5-tetramethyl-$\Delta^3$-pyrroline-3-methyl)methanethiosulfonate (dMTSL, Toronto Research Chemicals Inc., Canada). Briefly, purified protein powder was dissolved in 8 M urea, 20 mM Tris-HCl pH 8.5, 5 mM DTT to make about 1 mM protein solution. Following the complete solubilization, DTT was removed using a

PD-10 desalting column (GE Healthcare Bio-Sciences Corp., Piscataway, NJ). MTSL or dMTSL stock solution in acetonitrile was added to the final concentration of 10 mM and the solution was incubated in the dark for 30 minutes at room temperature. In order to ensure complete reaction, an additional 10 mM of the spin labeling reagent was added after 30 minutes and the sample was incubated overnight at room temperature. Unreacted spin labels were removed using PD-10 desalting column and the sample was dialyzed for 2 days against deionized water using a Slide-A-Lyzer Dialysis Cassette (MWCO = 3.5 kDa, Pierce Biotechnology, Inc., Rockford, IL) to remove urea. The resulting precipitates were collected by centrifugation and lyophilization.

The construct containing the amino terminal $Cu^{2+}/Ni^{2+}$ binding (ATCUN) motif was cloned according to the method published previously (Donaldson et al., 2001). Expression and purification of the construct was done using the same procedure as the sample without the ATCUN motif. NMR samples for the construct were prepared by dissolving the protein powder in 100 mM DPC, divided into two samples, and incubated with $NiSO_4$ (diamagnetic) or $CuSO_4$ (paramagnetic) overnight at room temperature. Unbound metal ions were removed by incubating with Chelex-100 beads (Bio-Rad Laboratories, Hercules, CA) followed by centrifugation to remove resin.

*NMR experiments*

All NMR experiments were performed using a Bruker DRX600 equipped with a TXI-cryoprobe or a Magnex 800 with a Bruker AVANCE console at the University of California, San Diego. All the triple resonance experiments used for backbone resonance assignments, i.e., HNCA, HN(CO)CA, HNCACB and HNCO were modified to adapt non-linear sampling schedules. A sampling schedule containing 600 random points was generated using COAST version 1.2 (Rovnyak et al., 2004) from a matrix of 40×64 real points for $^{15}$N and $^{13}$C, respectively. The spectra acquired using the non-linear sampling schedules were processed using RNMRTK (Hoch and Stern, 2001) and all other spectra were processed with NMRPipe (Delaglio et al., 1995). All NMR data was analyzed with SPARKY (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). N-H RDCs were measured using a sensitivity-enhanced version of $^{15}$N,$^{1}$H-IPAP HSQC (Ding and Gronenborn, 2003). The steady-state heteronuclear $^{15}$N-$^{1}$H NOE experiment was acquired using the standard parameters (Farrow et al., 1994). The transverse PREs for amide protons were measured by either the pulse program previously published (Donaldson et al., 2001; Iwahara et al., 2003) or calculated from the $^{15}$N,$^{1}$H-HSQC intensity ratios between diamagnetic and paramagnetic samples using Equation (4.2) (Battiste and Wagner, 2000).

*Structure calculations*

Dipolar Wave analysis was performed as previously published (Mesleh and Opella, 2003). Fitting of the sinusoid pattern to the measured N-H RDC values in the

helical regions as a function of residue numbers was performed using the standard non-linear minimization function in MATLAB (version 6, MathWorks). PRE distances were obtained using Equation (4.1) and assuming the $\tau_c$, the correlation time for the electron-nuclear interaction, is close to the global rotational correlation time of the target protein. The $\tau_c$ of 20 ns was used based on the previous report for a β-barrel protein OmpA in DPC. Distances were divided into 3 classes converted from the intensity ratios: 0.15 to 0.85, below 0.15 and above 0.85. The peaks that completely or nearly disappeared (intensity ratio below 0.15) were converted into a distance restraint of 10 ±6 Å. Intensity ratios above 0.85 were converted into a distance restraint of 30 ±6 Å. The intensity ratios between 0.15 and 0.85 were calculated into distances using the Equation (4.1) and (4.2), with ±4 Å bounds.

The standard input files for XPLOR-NIH (Schwieters et al., 2003) written in Python language were used for the structure calculations. Simulated annealing started from an extended template using the input restraints of PRE distances, N-H RDCs from the two different alignments, and dihedral angles ($\phi$,$\psi$) of (-65°,-40°) for the helical segments. The alignment tensors for RDC restraints were initially obtained using a Dipolar Wave analysis, by varying the alignment tensors while fitting the sinusoid to the data. After iterative optimization of the alignment tensor values, the structure with the smallest deviations from the experimental restraints was selected and used as the template for subsequent refinement. 100 structures were calculated and 30 structures were selected based on the pair-wise root mean square deviation (RMSD).

## 4.4 Results and Discussion

After target selection and extensive expression screening (Korepanova et al., 2005), screening via NMR spectra, i.e., $^{15}N,^{1}H$-HSQC in various conditions was performed (Page et al., 2006). The spectral quality, as judged by resolution of the spectrum, i.e., resonance linewidths and chemical shift dispersion, is directly related to the quality of subsequent experiments, thus to the quality of the final measurements. NMR structural studies of membrane proteins not only require that the protein is functional and well-folded in the selected detergent solution, but additionally that high quality NMR spectra are obtained to facilitate accurate data measurements. Since the molecular motions of protein-detergent complex strongly affect the resonance linewidths, care must be taken to choose the detergent-protein complex that has a correlation time within an acceptable NMR timescale in order to obtain narrow resonance linewidths. Recent applications have been developed for large biological systems such as transverse relaxation optimized spectroscopy (TROSY, Pervushin et al., 1997) and deuteration-selective methyl protonation (Tugarinov et al, 2004). These techniques have significantly increased the size limit of NMR spectroscopy (Sprangers and Kay, 2007). Unfortunately, the complicated dynamics of protein-detergent complexes prohibit the full benefit from these pulse sequence developments.

The present study utilized conventional pulse sequences and optimized sample conditions in order to achieve high quality NMR data. Quick detergent screening was employed using the detergents commonly used in NMR studies, e.g., sodium dodecyl sulfate (SDS), dodecylphosphocholine (DPC), and 1-myristoyl-2-hydroxy-sn-glycero-

3-[phospho-RAC-(1-glycerol)] (lysomyristoyl phosphatidyl glycerol, LMPG), and the best sample conditions were selected based on the linewidths, the number of peaks as well as chemical shift dispersion. In some cases, solubility in certain detergents was too low so that no further NMR experiments could be performed. In other cases, the number of signals in $^{15}N,^{1}H$-HSQC spectrum exceeded the number of backbone amide groups. The most likely explanation for this is the presence of conformational heterogeneity and conformational exchange (Krueger-Koplin et al., 2004). The $^{15}N,^{1}H$-HSQC spectra for three membrane proteins from *M. tuberculosis* in their best conditions are shown in Figure 4-3.

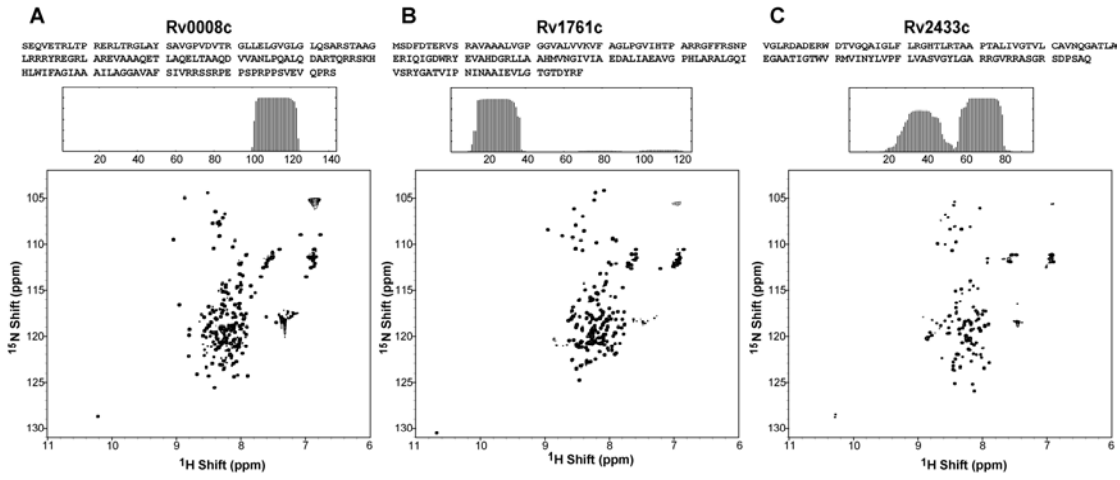


**Figure 4-3.** Examples on the $^{15}N,^{1}H$-HSQC spectra of membrane proteins from *M. tuberculosis* in various detergents. (A) Rv0008c in 250 mM SDS micelles (B) Rv1761c in 100 mM DPC micelles (C) Rv2433c in 150 mM LMPG micelles. All spectra were obtained at 323 K with pH adjusted to 4.0. The amino acid sequences of the proteins were shown on the top, with the diagram showing the predicted transmembrane regions by TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/).

100 mM DPC at 323 K, pH of 4.1 was selected as the detergent environment for Rv1761c based on the quality of the $^{15}N,^{1}H$-HSQC spectrum (Page et al., 2006)

and the ability to prepare weakly aligned samples in polyacrylamide gels. As discussed in Chapter 2, the limitation in the detergents and conditions that make homogeneous weakly aligned samples represents a significant impediment and severely limits the lipid mimetics that can be employed.

Earlier, it was demonstrated that for a small model system it is possible to completely assign backbone amide group resonances from a single 2D $^{15}$N,$^{1}$H-HMQC-NOESY experiment (Chapter 2). Unfortunately, as the size of the system increases, resulting in the number of peaks and the linewidths increasing, the spectra become more crowded, and the unambiguous assignments of peaks in 2D spectra are impossible. The obvious approach to increase resolution of experiments is to increase the dimensionality. In the current study, standard triple resonance experiments were implemented with non-linear sampling schedules and performed on a $^{13}$C,$^{15}$N-labeled sample. As shown by the previous reports (Rovnyak et al., 2004; Jones and Opella, 2006), incorporating non-linear sampling schedules results in a significant reduction in experimental time, but still provides the resolution and sensitivity necessary for resonance assignments of membrane proteins in detergent micelles. With the benefit of an excellent quality $^{15}$N,$^{1}$H-HSQC spectrum, over 90% of the resonances were assigned using only HNCA and HN(CO)CA (both acquired using non-linear sampling) along with the $^{15}$N,$^{1}$H-HSQC spectra from single amino acid selective labeled samples, i.e., $^{15}$N-Leu, $^{15}$N-Val, and $^{15}$N-Ala. HNCACB and HNCO experiments were also modified to incorporate non-linear sampling methods and used to confirm and complete the assignments.

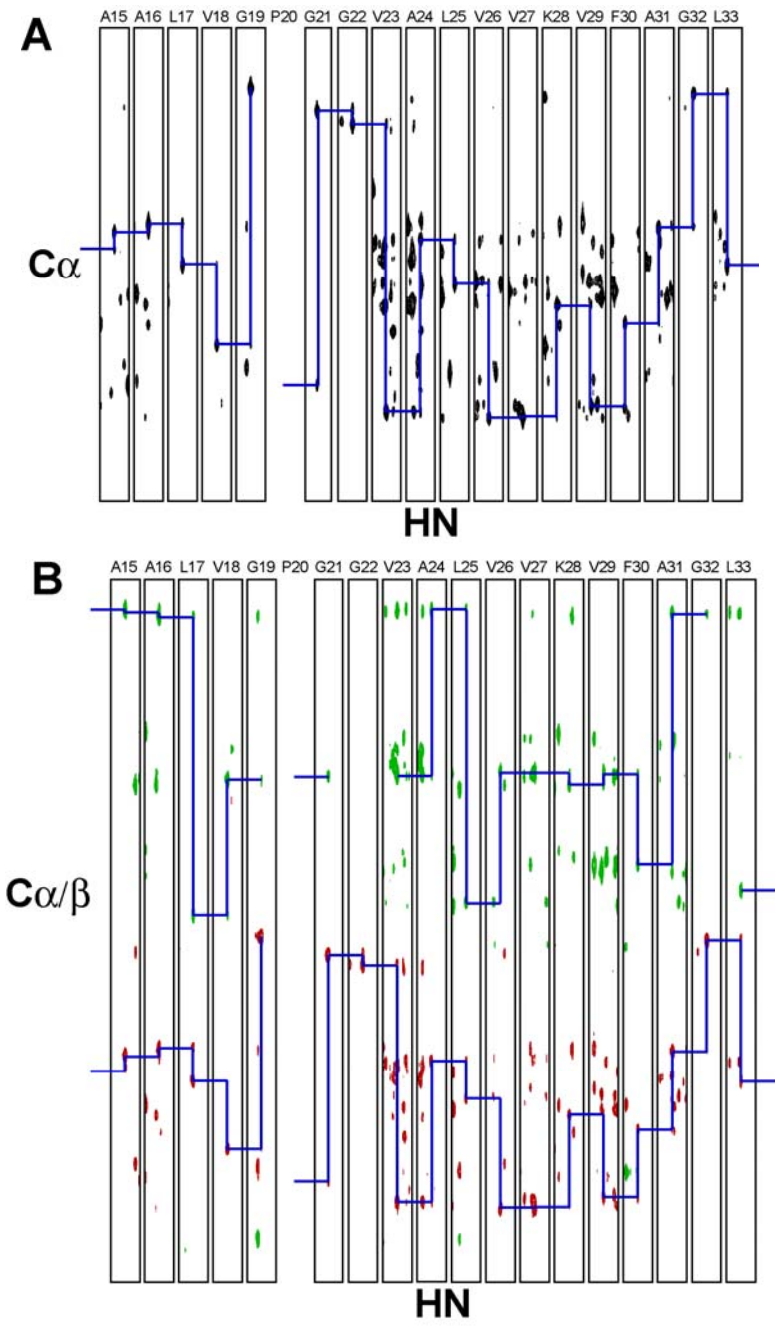


**Figure 4-4.** Strip plot of triple resonance experiments, (A) HNCA, (B) HNCACB, and (C) HNCO for the residues A15 to L33 of Rv1761c. The connectivities between the neighboring residues are shown for HNCA and HNCACB. All experiments were performed using the modified version from standard triple resonance experiments to implement non-linear sampling schedules, and processed with maximum entropy reconstruction method using RNMRTK ver.3.0. (D) $^{15}$N,$^{1}$H-HSQC spectrum of Rv1761c with assigned resonances labeled.

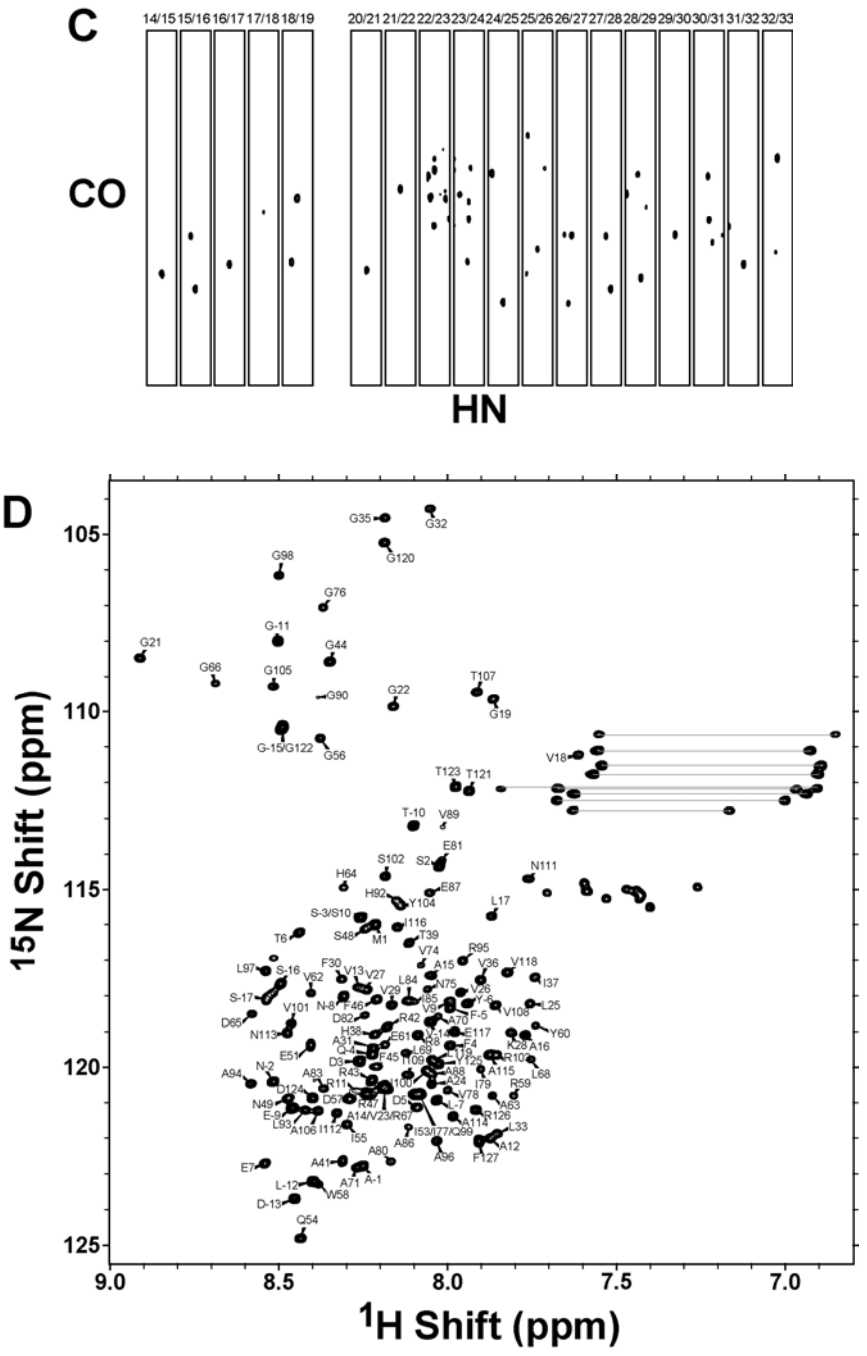


**Figure 4-4.** The triple resonance experiments and $^{15}N,^{1}H$-HSQC spectrum of Rv1761c, Continued.

Ideally, for small proteins, backbone resonance assignments can be completed by a walk-through of the Cα trace of HNCA and HN(CO)CA (see Figure 4). However, in the case of helical membrane proteins, the amino acid sequence consists primarily of hydrophobic residues, resulting in the limited chemical shift dispersion of the aliphatic carbons and/or amide protons and prohibiting unambiguous resonance assignments even for small, favorable systems. Since these samples are prone to aggregation, they are generally prepared at relatively low concentration, but the non-linear sampling schedules and maximum entropy reconstruction in the triple resonance experiments allows the number of scans to be increased to improve the sensitivity without making data acquisition prohibitively long or sacrificing the resolution.

The assigned $^{15}N,^{1}H$ HSQC spectrum is shown in Figure 4-4 (D) and contains the resonances from an extra 24 amino acid residues. This N-terminal 24-residue tag including polyhistidine residues was removed with Tobacco Etch Virus enzyme and the comparison of $^{15}N,^{1}H$-HSQC spectra (Figure 4-5) showed that, with the exception of the three N-terminal residues, all of the resonances perfectly match between the construct containing the polyhistidine residues and the one without the polyhistidine residues, confirming N-terminal tag is not affecting the protein structure. By removing the 24 residues from the protein, more simplified $^{15}N,^{1}H$-HSQC can be obtained, increasing the number of measurable peaks.

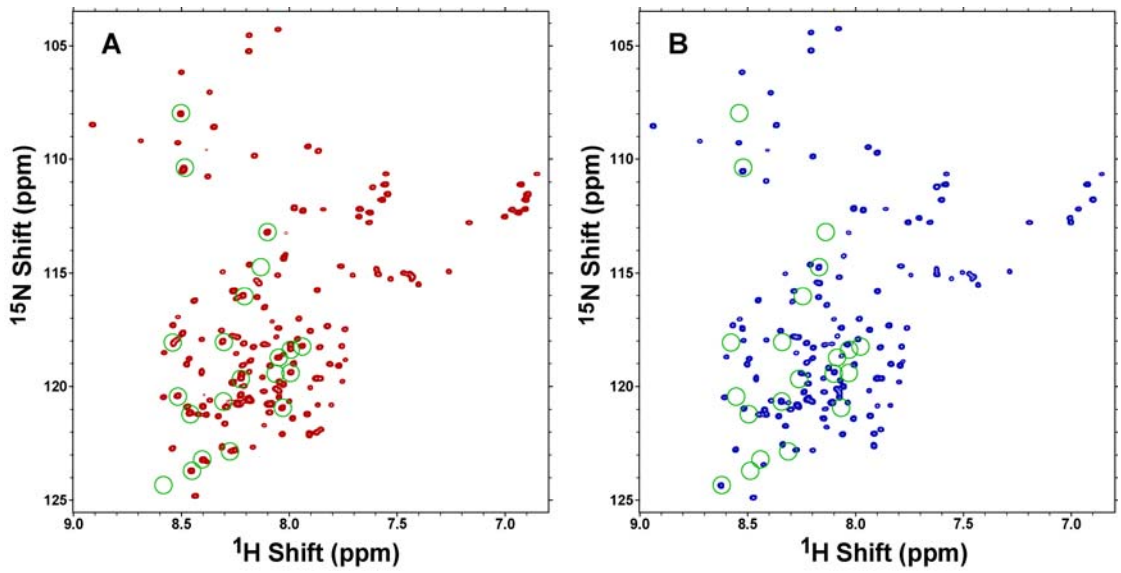


**Figure 4-5.** Comparison of $^{15}N,^{1}H$-HSQC spectra of Rv1761c constructs (A) containing N-terminal 24 extra amino acids including polyHis tag and (B) after cleaved by tobacco etch virus protease. Note that the most of the resonances are not affected by the cleavage reaction. Green circles highlight the resonances from extra residues cleaved by enzymes as well as three amino acids close to the cleavage site.

In parallel to the resonance assignment process, the sample was weakly aligned in polyacrylamide gels to obtain RDCs in as many "unique" orientations as possible. Although it is beneficial to obtain as many kinds of RDCs as possible other than ones for the N-H pair, such as Cα-Hα, Ca-C', and C'-N, it is extremely difficult and time consuming to measure any RDCs related to aliphatic carbons. Rather, it is more effective to focus on measuring N-H RDCs in different compositions of polyacrylamide gels, since this kind of alignment media seems to be the only media that can produce measurable RDCs for the membrane proteins in micelle or bicelle environments so far.

As mentioned previously, it is critical to optimize the isotropic sample condition so that the weak alignment in a polyacrylamide gel can be achieved. The

kind of detergent in the sample can be determined based on the quality of $^{15}N,^{1}H$-HSQC spectra, and the concentrations of detergents can be determined based on a guideline stated in Table 2-1 in Chapter 2. This is the reason for choosing 100 mM DPC micelles as the membrane-mimicking environment for Rv1761c and using the same method described in Chapter 2 that was successfully applied to align Rv1761c in a 5.5% polyacrylamide gel. In the present study, it was possible to obtain another set of alignments by adding small amounts of charged acrylate to the 'neutral' polyacrylamide gel. This is analogous to the previous studies where the addition of small amounts of charged molecules to bicelle alignment media (Ramirez and Bax, 1998) introduces some amount of electrostatic interaction. This creates an alternative alignment, reducing 4-fold orientational ambiguity that was discussed in Chapters 2 and 3. In the case of polyacrylamides, this can be achieved by the addition of charged acrylates such as: acrylic acid, 2-acrylamido-2-methyl-1-propanesulfonic acid (AMPS), 3-acrylamidopropyl-trimethylammonium chloride (APTMAC), N-(2-acryloamidoethyl)triethylammonium iodide (NAETEAI), and diallyldimethyl-ammonium chloride (DADMAC) (Meier et al., 2002; Cierpicki and Bushweller, 2004; Ulmer et al., 2003). The properties of polyacrylamide gels vary significantly after adding charged polyacrylates. For example, mixing charged acrylates with 'neutral' acrylamide significantly increases the swelling ability compared to the neutral acrylamide alone (Meier et al., 2002). This should decrease the effective pore size of a polyacrylamide gel and consequently increase the magnitude of RDCs. This needs to be compensated for by carefully varying total acrylamide concentrations,

crosslinker concentrations, or charged acrylate concentrations. Introducing charged acrylates also decreases the mechanical stability of the polyacrylamide gels, making it impossible to handle the gel after the dialysis against water when less than 4.5% of total acrylamide is used with AMPS.

The NMR data for Rv1761c is summarized in Figure 4-6. The secondary structure of the protein can be immediately noticed by aligning all the chemical shift indices (CSIs), intensity ratios from heteronuclear $^{15}N$-$^{1}H$ NOE, and N-H residual dipolar couplings. The chemical shift indices are obtained by plotting the difference between the observed chemical shifts and the standardized 'random coil' values (Schwarzinger et al., 2000). The steady-state heteronuclear $^{15}N$-$^{1}H$ NOE shown in Figure 4-6 (D) provides a direct and sensitive indicator of the local dynamics (Bogusky et al., 1987; Lee et al., 2003). As discussed in Chapter 2, the regions with the slow correlation times, indicated by high intensity ratios, are highly correlated with the locations of the helices that are predicted by CSI values shown in Figure 4-6 (A)-(C) or Dipolar Waves shown in Figure 4-6 (E). The local and globular dynamics of this protein can be determined in more detail by measuring the relaxation rates.

The Dipolar Wave analysis was performed on the N-H RDC data sets. This was accomplished by fitting the experimental data to a simple sinusoid with a periodicity of 3.6 (Mesleh and Opella, 2003). As expected, the helical regions detected by Dipolar Wave analysis match with the helical regions estimated by CSIs, because the Dipolar Wave analysis relies on the periodicity inherent in helical structures. The low deviations of the Dipolar Waves from the experimental data

suggest that the individual helical regions adopt ideal α-helical conformations. However, the TMHMM program estimated that this protein has a single transmembrane region of residues 14-37 (See Figure 4-3), which does not agree with the helical regions detected by CSIs or Dipolar Wave analysis. Instead, the Dipolar Wave analysis suggests there are total of six α-helical regions, two of which are included in the transmembrane region that was predicted by TMHMM, from residue 5 to 17 (Helix 1) and from residue 22 to 31 (Helix 2). This kind of conflicting result between the structure prediction program and the experimental data is not uncommon. The Dipolar Waves certainly suggest the presence of a kink around residue $Pro^{20}$, which is characterized by a sudden change in the phase, amplitude, and/or average value of the sinusoid (Mesleh and Opella, 2003). The presence of a kink in the middle of a transmembrane helix is also not uncommon, especially around proline residues (Chamberlain and Bowie, 2003). It is not clear at this point whether the transmembrane domain is comprised of all the residues in Helix 1 and 2, or only the part of those helices as predicted from TMHMM.

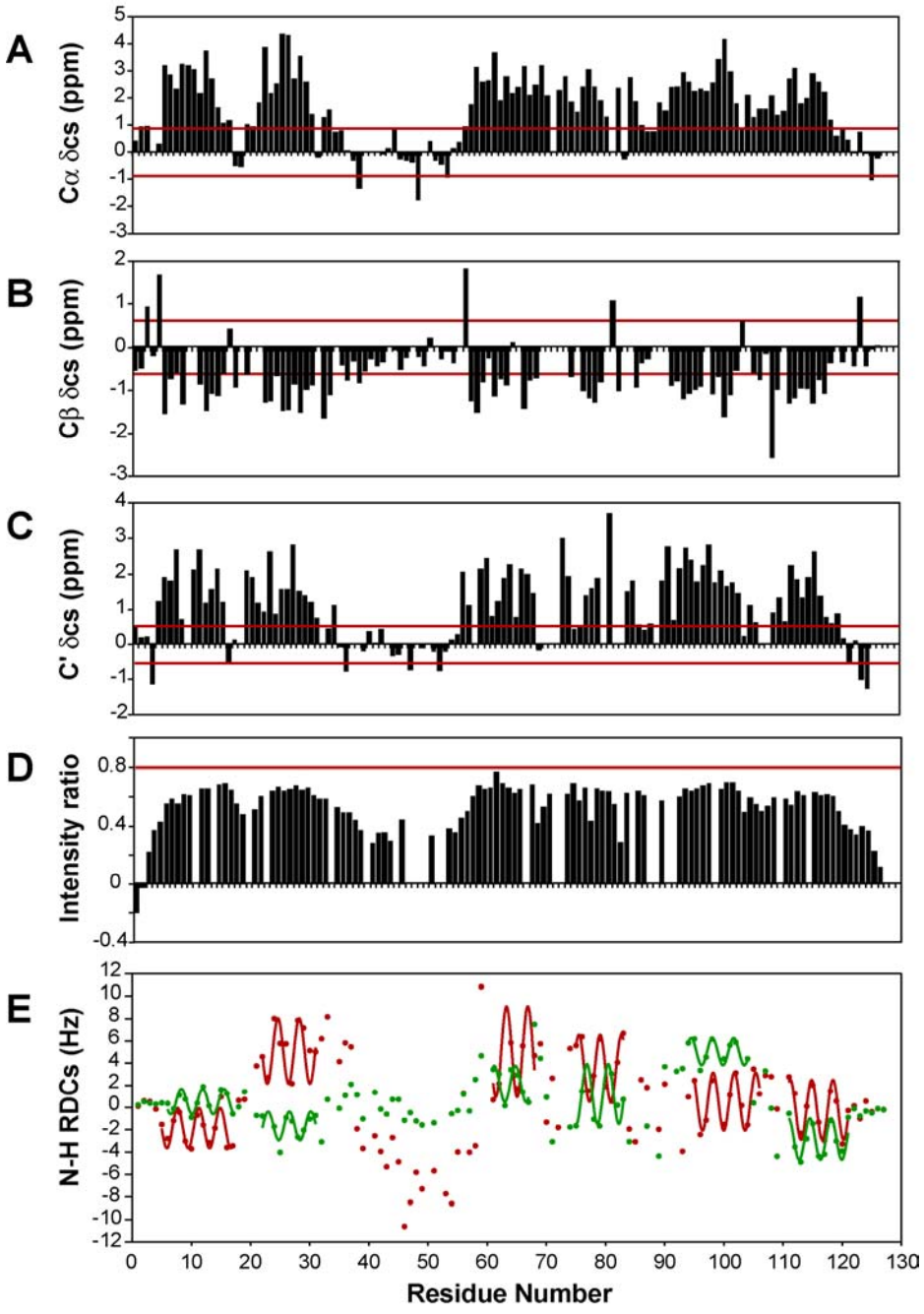


**Figure 4-6.** Summary of NMR data for Rv1761c in 100 mM dodecylphosphocholine micelles. (A) Cα chemical shift index (B) Cβ chemical shift index (C) C' chemical shift index (D) Relative intensity ratios from static heteronuclear $^{15}$N-$^{1}$H NOE experiments (E) N-H residual dipolar couplings measured either using a 5.5% polyacrylamide gel with 0.14% N,N'-methylenebisacrylamide (red) or a 5% polyacrylamide gel with 0.14% N,N'-methylenebisacrylamide and 1% 2-acrylamido-2-methyl-1-propanesulfonic acid (green). The best fitting sinusoids for helical regions are shown as solid lines.

The secondary structure information from the RDC data acquired using two different alignments is in agreement, as judged by the helical regions detected by the Dipolar Wave analysis, as shown in Figure 4-6. This suggests that the RDC data from different alignments resulted from the interaction (steric and/or electrostatic) between the protein-micelle complex and the polyacrylamide gel matrices, and not from the structural perturbation. This can also be confirmed by comparing the chemical shifts of the isotropic and anisotropic samples, as no significant chemical shift differences were observed. Perhaps the most notable feature in the current data is that the two alignments obtained from different kinds of polyacrylamide gels are not "degenerate", as detected by the Dipolar Wave analysis. The differences in the phase, average value and/or magnitude of the individual Dipolar Waves are immediately seen, and because the properties of the Dipolar Waves from these two alignments are different from each other, the alignment tensors from these data are "unique". Using these non-degenerate orientations, the angular degeneracy inherent in dipolar couplings can be reduced. These "unique" orientations are generated by the unique interactions between the solute and the alignment media, and this kind of approach has been applied to several globular proteins where many alignment media can be applied (Peti et al., 2002; Ulmer et al., 2003). As seen in Chapter 3, changing the direction of compressions, i.e., axial versus radial, merely caused the 180 degree rotation in the alignment tensors. In such case, the protein has the same orientation within the alignment frame for both axial and radial compression and consequently they have identical degenerate solutions, thus the angular degeneracy cannot be reduced. However, as shown in the

current chapter, the addition of electrostatic interactions by introducing the charged acrylate molecules yields an alignment tensor distinctly different from the alignment tensor obtained using a neutral polyacrylamide gel. Other kinds of acrylates have also been tested and each required separate optimization. It is best to collect the RDCs using as many orientations as possible in order to define the relative orientations of individual vectors more precisely. However, even if the angular degeneracy problem is solved this only determines the relative orientations of the helices and does not provide translational information.

The long-range distance information provided by PRE that may substitute NOE distance information represents one method to provide the requisite translational information and can potentially remove angular degeneracy from RDC measurements. The developments and applications of different spin probes as well as the incorporation of the measured values in terms of structure calculations have been published recently (Donaldson et al., 2001; Battiste and Wagner, 2000; Iwahara et al., 2004). Two kinds of approaches have been applied to introduce paramagnetic centers: amino-terminal $Cu^{2+}/Ni^{2+}$ binding (ATCUN) motif and site-directed spin labeling (SDSL).

The applicability of ATCUN motif ($NH_2$-Gly-Ser-His-Gly-) on PRE measurements has been previously tested on human ubiquitin (Donaldson et al., 2001). Paramagnetic $Cu^{2+}$ or diamagnetic $Ni^{2+}$ is known to be coordinated by four nitrogen atoms in the motif through the distorted square planar geometry. Figure 4-7 (A) and (B) show the $^{15}N$-$^{1}H$ HSQC spectra of ATCUN-Rv1761c with $Ni^{2+}$ and $Cu^{2+}$. High-

quality $^{15}N$-$^{1}H$ HSQC spectra were observed, and negligible chemical shift changes from $^{15}N$-$^{1}H$ HSQC spectra for the TEV-cleaved sample were seen (see Figure 4-5), indicating that neither $Ni^{2+}$-bound nor $Cu^{2+}$-bound ATCUN motif affects the overall structure of Rv1761c. The intensity ratios between the diamagnetic ($Ni^{2+}$-bound) and paramagnetic ($Cu^{2+}$-bound) samples were calculated and converted into distances using an equation similar to the Equation (4.1), except for the g factor ($g_{Cu}$=2.09). As reported previously, the concentration-dependent broadening caused by intermolecular interactions was observed, and this effect was significantly reduced at low (~0.4 mM) protein concentration. The PRE $R_2$ for amide protons were directly measured using the pulse sequence previously published (Donaldson et al, 2001; Iwahara et al., 2003). Despite the high-quality data was obtained, the structure calculations did not yield consistent structures. The main reason for this short coming was the mobility of the paramagnetic center and inability to fix its location during the calculations. The intrinsic flexibility of the paramagnetic centers (especially for the metal-binding tags) has been previously noted and requires complicated computational strategy in the structure calculations to account for this motion (Iwahara et al., 2004). While this strategy may seem attractive, a more rigid spin label was employed in the current study, since the goal was developing a method for rapid backbone structure determination.

Site-directed spin labeling has become a popular and straightforward approach to introduce an extrinsic paramagnetic probe in proteins (Hubbell et al., 1998). Once a single cysteine site is introduced on the target protein, a spin labeling reagent with

methanethiosulfonate group, e.g., MTSL, reacts with the sulfhydryl group from the sidechain of cysteine residue, generating a nitroxide sidechain. Figure 4-7 (C) and (D) show the $^{15}N$-$^{1}H$ HSQC spectra of Rv1761c spin-labeled on residue $Phe^{30}$ mutated to cysteine (F30C). In the previously reported approach, after the intensities of the $^{15}N$-$^{1}H$ HSQC spectrum were measured using the paramagnetic sample, ascorbic acid was subsequently added to reduce the nitroxide group and the intensities were measured again (Battiste and Wagner, 2000). This method is advantageous since it utilizes the same sample and eliminates the need to calibrate protein concentration. However, when this approach was applied to Rv1761c, ascorbate was unable to completely reduce the spin label, as judged by the comparison of intensities. This problem can be easily solved by applying a parallel spin labeling approach utilizing a diamagnetic analog of the spin labeling reagent, dMTSL (Liang et al., 2006). Several additional cysteine mutants including S10C and S48C were also generated and spin labeled. Unfortunately, the flexibility of $Ser^{48}$ prevented the location of the spin label from being defined, and the S10C mutation had an increased number of peaks in $^{15}N$,$^{1}H$-HSQC spectrum even after labeling with dMTSL, indicating the presence of conformational heterogeneity.

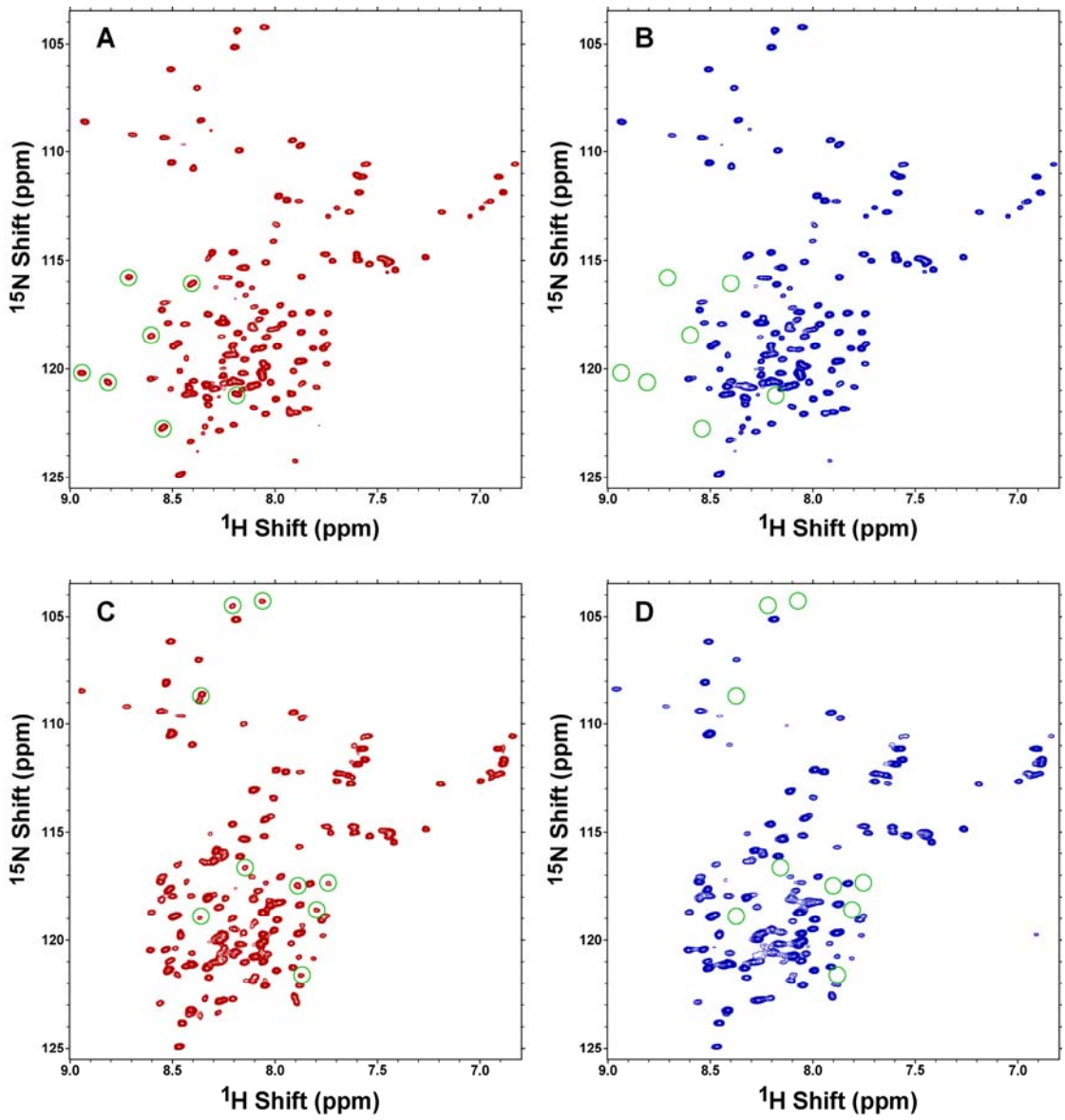


**Figure 4-7.** $^{1}H,^{15}N$-HSQC spectra of ATCUN-Rv1761c with ATCUN motif being (A) $Ni^{2+}$-bound and (B) $Cu^{2+}$-bound, and F30C-Rv1761c with (C) dMTSL attached and (D) MTSL attached to $Cys^{30}$. Green circles highlight the residues that are completely broadened out in the paramagnetic samples.

Figure 4-8 shows the measured $^{15}N,^{1}H$-HSQC intensity ratios between the paramagnetic sample and the diamagnetic sample, as well as the PRE distances converted using the Equations (4.1) and (4.2). As noted in Figure 4-2, the distances

calculated using Equation (4.1) depend on the global correlation time, $\tau_c$. Because the dynamics of the current target protein is not thoroughly studied, the exact value of the global correlation time is unknown. However, the dynamics of β-barrel membrane proteins PagP (Hwang et al., 2002) and OmpA (Arora et al., 2001; Tamm et al., 2003) have been studied quite extensively and the overall isotropic rotational correlation times were reported to be around 20 ns in both cases. Because both proteins were studied in DPC micelles, 20 ns of correlation time can be estimated for Rv1761c, assuming the overall rotational correlation time of protein-detergent complex is mostly governed by the detergent molecules. The precision of the calculated distances can be improved if the dynamics of individual backbone amide sites are incorporated into the distance conversion. Nonetheless, the rough approximation of the correlation time can be compensated by the range of input distance restraints. The structure calculations for Rv1761c used the distance restraints with ±4 Å bounds for the intensity ratios between 0.15 and 0.85, which is sufficient to cover the correlation time range of about ±10 ns as estimated from Figure 4-2 (B).

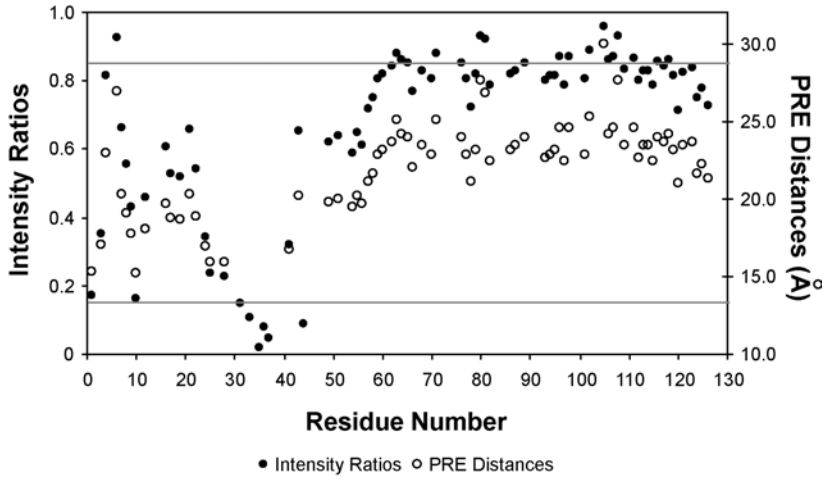


**Figure 4-8.** Plot of paramagnetic relaxation enhancement (PRE) measurements as a function of residue numbers for F30C-Rv1761c. Filled circles represent the intensity ratios between paramagnetic and diamagnetic samples measured from $^{15}N,^{1}H$-HSQC spectra and open circles represent the corresponding distances converted using Equations (4.1) and (4.2) assuming the global correlation time of 20ns. The lines indicate the categorized boundaries when used as the distance restraints in the structure calculations.

Calculated structures of Rv1761c are shown in Figure 4-9. The representative structure in (A) is a cartoon representation with the input PRE distances shown as grey dotted lines from the nitroxide side chain. The grey solid lines represent where the membrane bilayer (or the detergent micelles) would be located if helix 1 and helix 2 constitute a transmembrane domain as predicted by the transmembrane prediction program, TMHMM, and the thickness of the membrane (micelles) is estimated to be 20 Å. The actual transmembrane region can only be hypothesized at this point and it is possible that only a part of the helical structure is located inside the hydrophobic lipid bilayer. The absolute orientation of the transmembrane helices with respect to the lipid bilayer can not be determined with the current data and requires information

from alternative techniques such as solid-state NMR spectroscopy (Nevzorov et al., 2004).

Figure 4-9 (B) and (C) show 30 selected structures out of 100 calculated structures. When overlaid with the representative structure shown in Figure 4-9 (A), a backbone RMSD of 2.75 was obtained when all helical regions are overlaid, whereas backbone RMSD of 0.64 was calculated when helix 1 and helix 2 are overlaid. Since the major goal of current study was to develop a strategy to rapidly obtain the structural information for membrane proteins, the resolution of the current structures would be sufficient. This 'initial' structure can provide some insights as to how to design further experiments, e.g., where to place additional spin-labeled cysteine residues. Because of the lack of experimental constraints, the calculated structures were forced to satisfy the experimental constraints during the course of simulated annealing. Therefore, one should be very cautious not to ignore the possibility of other structures, and providing more PRE distances from other sites would solve any ambiguity that might be present. A caveat regarding the structure calculation is the potential for flexibility of side chain atoms in the spin label group. Heterogeneity in the side chain conformation can produce multiple distance ranges for the amide proton in the helical segments, yielding poor convergence among the calculated structures. As demonstrated by the crystal structures from the Hubbell group (Langen et al., 2000), the nitroxide spin label side chain (conventionally named 'R1') generally adopts ($X_1$,$X_2$) of ($g$,$g$) conformation, and the structure calculations for Rv1761c employed this spin label conformation during the refinement stage for better

convergence. PRE measurements from multiple sites would potentially alleviate this problem, as well as the computational strategy to account for the flexibility of the spin label group and more accurately determine the PRE distances (Iwahara et al., 2004).

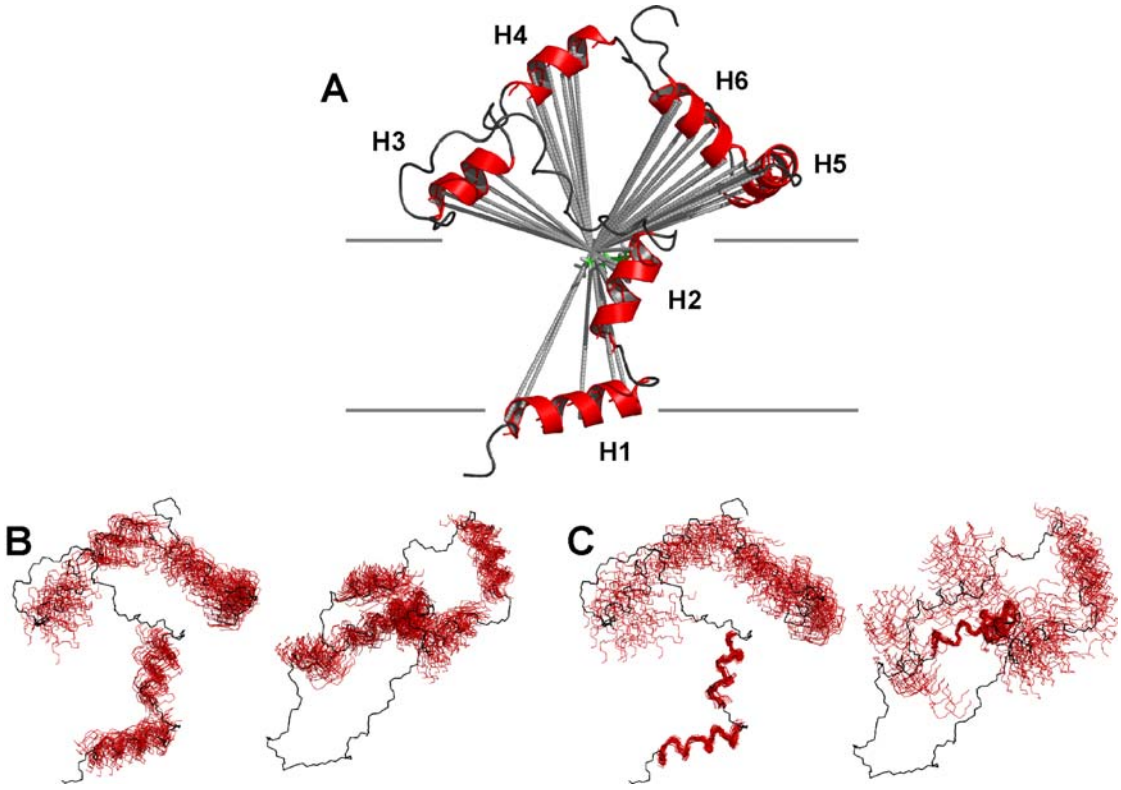


**Figure 4-9.** The backbone structure of Rv1761c calculated using orientational information from residual dipolar couplings and distance information from paramagnetic relaxation enhancement. (A) A cartoon representation of Rv1761c. The grey dotted lines show the input distances from the spin label MTSL drawn in stick. 30 selected structures are overlaid against the structure shown in (A) for the (B) helical regions or (C) transmembrane helices, i.e., helix 1 and helix 2. Side view and the top view are shown in (B) and (C). Only backbone atoms of helical regions are shown except the structure shown in (A), colored in black.

The quality of the calculated structures to the experimental data can be evaluated by the correlation plots in Figure 4-10. The correlation coefficients for both RDC data sets are close to 1.0 as shown in Figure 4-10 (A) and (B) and indicate there

is an excellent agreement between the experimental data and the calculated values from the structures. This demonstrates that the calculated structures satisfy the RDC data from both alignments, and that using these two different alignment media preserves not only the secondary structure information, as shown in Figure 4-6 (E), but also the tertiary structure information. A high correlation coefficient for the PRE distance restraints was also observed, as shown in Figure 4-10 (C), indicating that all of the measured distances from the calculated structures reside within the distance bound and satisfy the input distance restraints. As expected, some of the input distance restraints calculated from the intensity ratios above 0.85 deviate from the distance bounds of ±4 Å because these intensity ratios were converted into the distance restraints with the bounds of ±6 Å. Error bars in the correlation plot show the RMSDs for the 30 calculated structures, illustrating good convergence. Determining the distance bounds for this type of calculation can be difficult, because tight distance bounds might force the structure into the wrong conformation and the wide distance bounds might generate the structures with poor convergence. A prior knowledge of structure allows the converted distances to be calibrated and solves this problem (Battiste and Wagner, 2000; Liang et al., 2006) but this method is not feasible in the absence of a high resolution structure. In this case, iterative structure calculations with varying distance limits yielded ±4 Å to be the best bounds.

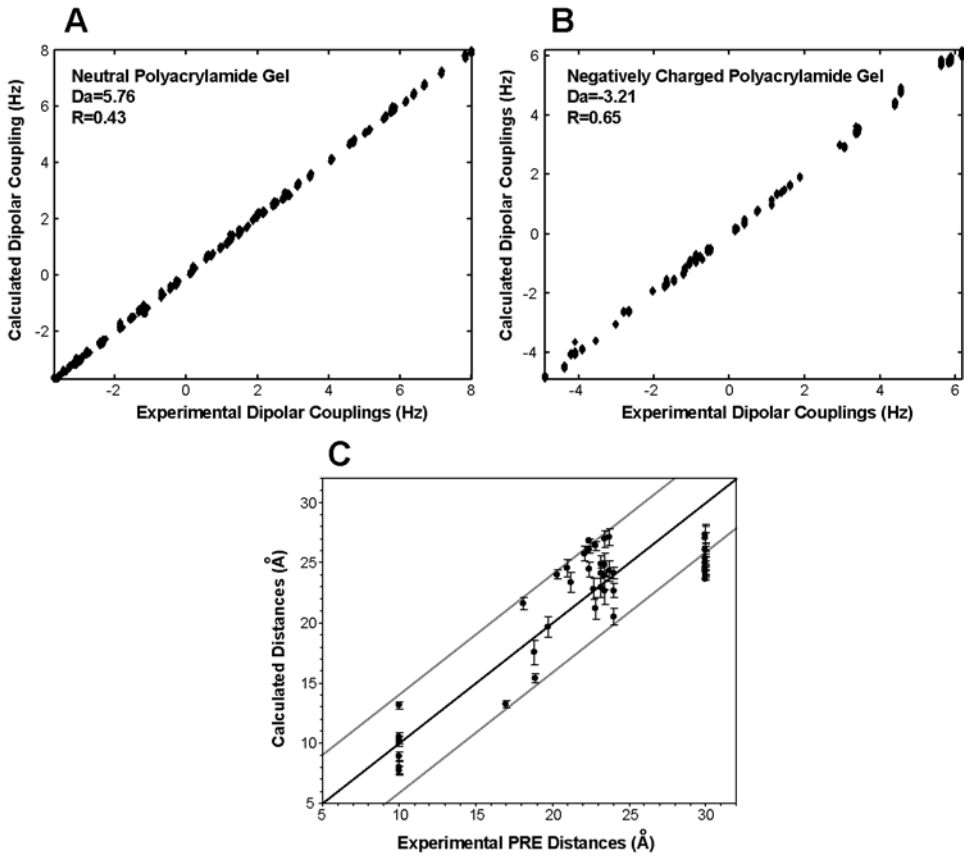


**Figure 4-10.** Correlation plots for the structure calculated using the information from residual dipolar coupling and paramagnetic relaxation enhancement measurements. Plots of the calculated dipolar couplings using 30 structures calculated by XPLOR-NIH versus the experimental dipolar couplings obtained using the alignment media of (A) neutral polyacrylamide gel or (B) negatively charged polyacrylamide gel. The calculations of dipolar couplings from the structures were done by RDCA (provided by Prof. L.E. Kay at the University of Toronto), the script written for MATLAB. (C) The plot of calculated distances versus experimental distances from paramagnetic relaxation enhancement. The error bars show the RMSDs within the 30 calculated structures, and the grey lines represent the distance range of ±4 Å.

An interesting issue is determining the number of PRE measurements required to define the backbone fold for this type of protein. As shown by the dotted lines in Figure 4-9 (A), the PRE distances in the flexible loop regions were excluded from the structure calculations, as these loop regions exhibit different dynamics from those of helical regions. The only meaningful PRE distances at this stage are from the rigid

helical regions and the structure calculations were performed with each helical segment treated as a "rigid body". The main role of PRE distances, in this study, was to roughly locate these rigid helical segments. Ideally, if the relative orientations of the helical segments were defined with a sufficient number of RDC measurements from multiple 'unique' alignments, one PRE measurement per helical segment would be sufficient to locate each helical segment. However, measuring a number of PRE distances per helical segment is much more straightforward than obtaining the unique alignments and measuring the types of RDCs required to define the orientations unambiguously.

Not only do the PRE measurements serve as translational restraints, but they also provide discrimination between the degenerate solutions from the RDC measurements. This is illustrated in Figure 4-11, where the structure calculated using XPLOR-NIH are overlaid on the degenerate solutions obtained from the Dipolar Wave analysis. The structures in ribbon representations are the degenerate solutions calculated by Dipolar Waves and are overlaid with respect to helix 1 (H1). Red ribbons represent the four degenerate solutions for helix 2 (H2) from N-H RDCs obtained using a neutral polyacrylamide gel as the alignment medium (red dots and lines from Figure 4-6 (E)). Green ribbons represent the degenerate solutions for helix 2 (H2) from N-H RDCs obtained using negatively charged polyacrylamide gel (green dots and lines from Figure 4-6 (E)). This figure proves that once the 'unique' orientations are obtained, the degeneracy inherent in RDC measurements can be reduced, which was reported previously using globular proteins (Ramirez and Bax,

1998). The Dipolar Wave analysis, by fitting the sinusoid to the data, can obtain the angles defining the relative orientations between the helices, and proves that the two unique alignments implemented in this study reduced the degeneracy by half. Interestingly, when the structure calculated from XPLOR-NIH, drawn in purple sticks, was overlaid to the helix 1 from Dipolar Wave analysis, the orientation of helix 2 from XPLOR-NIH matches with one of the degenerate solutions calculated from Dipolar Wave analysis. Moreover, the relative orientation of helices from XPLOR-NIH corresponds to one of the orientations from the 'reduced' degenerate solutions using two unique alignments. Note that the critical factor in determining a single orientation *among* the two 'reduced' degenerate solutions was PRE distances; PRE distances help reducing the orientational degeneracy as well as the translational degeneracy.

The agreement between the structure calculated by XPLOR-NIH and one of the degenerate structures calculated by Dipolar Wave analysis is not surprising because the quality of the calculated structure from XPLOR-NIH, shown in Figure 4-10, or the low RMSDs of the fitted sinusoid from the data, shown in Figure 4-6 (E), already suggest that the structures from both methods satisfy the RDC data from both alignments. Nonetheless, this proves the applicability of both methods in rapid structure determination.

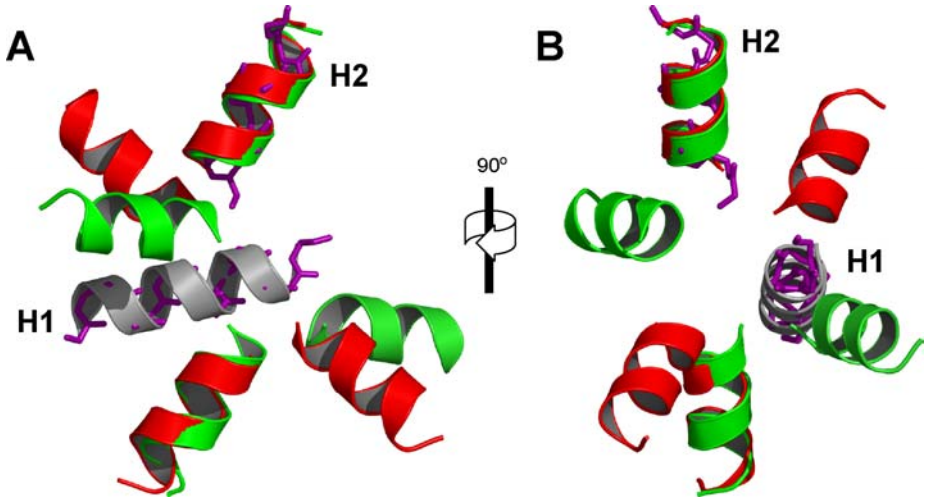


**Figure 4-11.** The structure calculated using PRE distances and RDC data matches with one of the degenerate structures from Dipolar Wave. Depicted are the structures predicted by Dipolar Wave (ribbon representations) and calculated by XPLOR-NIH (purple stick representation) for helix 1 (H1) and helix 2 (H2). Degenerate solutions are calculated using RDC data from the sample aligned in neutral and negatively charged polyacrylamide gels, and these are colored in red and green respectively. The structures are aligned with respect to helix 1.

## 4.5 Concluding Remarks

The current study is an example of the rapid structure determination of a helical membrane protein using minimal experimental data from solution NMR spectroscopy. A series of experiments were performed to extract the structural information necessary to define the backbone structure. Triple resonance experiments were performed with significantly reduced acquisition time to accelerate the resonance assignment process. The RDCs were measured from $^{15}$N,$^{1}$H-IPAP-HSQC experiments to provide the orientational information among the helical segments determined by the Dipolar Wave analysis. The PREs measured from the intensities of $^{15}$N,$^{1}$H-HSQC were used to provide distance information and locate the helical segments.

A similar approach has been applied on the β-barrel membrane protein, OmpA in DPC micelles (Liang et al., 2006; Cierpicki et al., 2006), although no application has been reported for the structure determination using only RDCs and PRE distances. In cases of β-barrel membrane proteins, it is easier to observe long-range NOEs between the backbone atoms in the β-sheets that form the hydrogen bond network and calculating the structure can be facilitated using the long-range hydrogen bond restraints. However, for the helical membrane proteins, such long-range NOEs that define the locations of the helices, i.e., inter-helical NOEs, are difficult to obtain because of the distance limitation of NOE itself. Although the selective protonation of the methyl group in highly deuterated background (Tugarinov et al., 2004) seem highly promising for obtaining such long-range NOEs, the applicability of this method to the helical membrane proteins containing repeated hydrophobic residues needs to be examined. Therefore, for now, structure calculations of helical membrane proteins need to rely on other kinds of restraints such as RDCs and PREs.

The choice of alignment media for membrane proteins is still quite limited because of the incompatibility of protein-detergent complex with most types of alignment media, and the development of other kinds of alignment media for membrane proteins is certainly needed. While attractive, *de novo* structure determination using only RDC data requires quite extensive measurements from different nuclei as well as from multiple alignments, and the development of suitable alignment media for the membrane proteins in detergent solutions or lipid environments should be continued. Parallel to this, the development of PRE distance

measurements, the improvement of the computational strategy to implement the data into the structure calculation, and the quest for a better paramagnetic spin probe need to be continued.

# 5 Applicability to a More Complex System - human Copper Transporter 1

## 5.1 Introduction

Copper is an essential element required for many important cellular functions. Because of its special redox property, it is involved in a variety of biological reactions such as photosynthesis and respiration, free radical eradication, neurological development, and iron homeostasis (reviewed in Harris, 2000; Llanos and Mercer, 2002). Like many other essential trace metals, excessive copper is toxic to the cell, and intercellular copper level is tightly regulated so that sufficient copper can be absorbed and excess copper can be secreted. The human diseases Menkes syndrome (copper deficiency) and Wilson disease (copper toxicosis) demonstrate the essentiality and potential toxicity of copper, and are caused by mutations in the P-type copper ATPases, ATP7A and ATP7B, respectively. Although some progress has been made to understand the mechanism of the intracellular copper transport through the metallochaperones (reviewed in Huffman and O'Halloran, 2001; Rosenzweig, 2002), the general mechanism of copper translocation across cellular membranes is not fully understood.

Copper transporters (CTR) are a family of proteins responsible for cellular copper uptake. Genetic studies in the yeast *Saccharomyces cerevisiae* led to the first identification of proteins involved in high-affinity copper uptake at the plasma

membrane of eukaryote cells, known as yCTR1 and yCTR3 (Dancis et al., 1994; Knight et al., 1996), and the human high-affinity copper transporter (hCTR1) was identified by functional complementation studies of *yCTR1* deletion mutant (Zhou and Gitschier, 1997). hCTR1 is a human high-affinity copper uptake transporter which consists of 190 amino acids and has three putative transmembrane segments and two methionine-rich metal binding motives. Models describing the oligomerization state of the protein have been proposed by several groups, which include dimer (Klomp et al., 2003), trimer (Lee et al., 2002) and dimer of trimers (Aller and Unger, 2006). Interestingly, the connection between the resistance to extracellular copper and the resistance to cis-diamminedichloroplatinum(II) (cisplatin), a platinum-based chemo-therapeutic agent has been observed (reviewed in Safaei and Howell, 2005), and the direct involvement of CTR1 in cisplatin transport was also demonstrated by showing the effect of *ctr1* deletion on cisplatin resistance as well as showing the competitive uptake between the copper and the cisplatin (Ishida et al., 2002). The principal goal of current study is to determine the structure of the high affinity copper transport protein hCTR1 as the first essential step towards understanding the molecular basis of its biological functions in copper homeostasis and the development of resistance to cisplatin in cancer chemotherapy.

In this chapter, an attempt to obtain structural information on the more complex system with significant biological functions is described. As a 190-residue oligomeric, multi-span helical membrane protein, hCTR1 presents great challenges for experimental methods of protein structure determination. Optimizing all of the

conditions required by the methodological developments that were discussed in Chapter 2, 3 and 4 will be crucial to the backbone structure determination. In other words, one needs to find a sample condition where (1) the sample can produce a $^{15}N,^{1}H$-HSQC spectrum with reasonable resolution and dispersion (meaning the sample is properly folded), (2) the sample can be weakly aligned in polyacrylamide gel so that the residual dipolar couplings (orientational information) can be measured, and (3) the sample can be spin-labeled without perturbing the structure so that the paramagnetic relaxation enhancements (distance information) can be measured. The applicability of the developed methods to a complex system such as hCTR1 will be discussed as well as the future direction of the development.

## 5.2 Methods and Materials

*Cloning of wt-hCTR1$_{TM}$, C161S-hCTR1$_{TM}$, C189S-hCTR1$_{TM}$, and C161S/C189S-hCTR1$_{TM}$*

Full-length hCTR1 cDNA was obtained from Professor Stephen B. Howell (University of California, San Diego). cDNA corresponding to the amino acid region 45-190 of hCTR1, hCTR1$_{TM}$ was PCR amplified from the DNA encoding full-length hCTR1 using the oligonucleotides 5'-GGAATTCCATATGACCTTCTACTTTGG and 5'-CCGCTCGAGATGGCAATGCTCTGTG. The PCR products were digested with the restriction enzymes *Nde* I and *Xho* I (New England Biolabs, Ipswich, MA), and purified by agarose gel electrophoresis using the protocol supplied with a Gel Extraction Kit (Qiagen, Valencia, CA). The DNA fragment was ligated into the

vector pET-31b(+) (EMD Biosciences, La Jolla, CA) which was cut by the same restriction enzymes and purified. The ligation product was transformed into DH5α competent cells (Invitrogen Corp., Carlsbad, CA) and the DNA sequence was verified. A PCR test with the same oligonucleotide primers on the ligation product was done to confirm if the ligation product contains the target PCR product. The ligation product was further confirmed by digesting with the restriction enzyme *EcoR* V to ensure the restriction digest yields the fragments of correct sizes. The final recombinant plasmid pEThCTR1$_{TM}$ contains the amino acid region 45-190 of hCTR1 plus amino acids LEHHHHHH at the C-terminus.

A single rare codon for *E. coli* was found within hCTR1$_{TM}$ at the position encoding Arg$^{90}$ (CGA) and was optimized by changing it to CGC using the standard site-directed mutagenesis protocol (QuikChange, Stratagene, La Jolla, CA) with the primers in Table 5-1. For the cysteine mutated variants, two cysteine sites, Cys$^{161}$ and Cys$^{189}$ were mutated into serine residues following the standard site-directed mutagenesis protocol with the primers given in Table 5-1.

**Table 5-1.** Oligonucleotides used in the site-directed mutagenesis for hCTR1$_{TM}$. Targeted mutation points are shown in bold. Primers are shown in a 5' to 3' direction.

| | |
|---|---|
| R90CGC forward | GGACTCAAGATAGCC**CGC**GAGAGCCTGCTGCG |
| R90CGC reverse | CGCAGCAGGCTCTC**GCG**GGCTATCTTGAGTCC |
| C161S forward | CCTACAACGGGTACCTC**AGC**ATTGCAGTAGCAGCAGG |
| C161S reverse | CCTGCTGCTACTGCAAT**GCT**GAGGTACCCGTTGTAGG |
| C189S forward | GGATATCACAGAGCAT**AGC**CATCTCGAGCACCACC |
| C189S reverse | GGTGGTGCTCGAGATG**GCT**ATGCTCTGTGATATCC |

Four hCTR1$_{TM}$ constructs were generated using the primers listed in Table 5-1: wt-hCTR1$_{TM}$, C161S-hCTR1$_{TM}$, C189S-hCTR1$_{TM}$, and C161S/C189S-hCTR1$_{TM}$. All mutations were done utilizing the *Pfu* polymerase-mediated reaction on the full-length plasmids. All the mutation constructs were verified by sequencing the plasmid using T7 promoter as a primer.

*Expression and Western blot analysis of hCTR1$_{TM}$ constructs*

The recombinant plasmids were transformed into *E. coli* competent cells C43(DE3) (Imaxio, France) and grown on LB agar plates containing 100 μg/mL carbenicillin. A single colony was picked and inoculated in 5 mL LB media with 100 μg/mL carbenicillin at 37 °C, 260 rpm for 2 hours. 500 μL of the culture was transferred into 100 mL M9 minimal media containing 3% LB media and incubated at 37 °C overnight, then 30 mL of the culture was transferred to 1 L M9 minimal media containing 3% LB media. The target protein expression was induced using 0.4 mM isopropyl-β-D-thiogalactoside (IPTG) when the $OD_{600}$ reached ~0.6, and was incubated for additional 4 hours at 37 °C. The cells were harvested by centrifugation at 12,000×g for 45 minutes at 4 °C using a Beckman Avanti J-20 XP centrifuge with a JLA-8.1 rotor. Isotope labeling was achieved by incorporating $^{15}$N-ammonium sulfate (Cambridge Isotope Laboratories, Inc., Andover, MA) as the sole nitrogen source.

Western blot analysis was performed on the whole cells using an antibody directed against C-terminal amino acid residues of hCTR1 (Novus biologicals, Inc., Littleton, CO) using the standard immunoblotting protocol. The samples were loaded

onto a 12% SDS-PAGE gel and electroblotted onto a polyvinylidene difluoride (PVDF) transfer membrane by applying 30 V for one hour. The blotted membrane was blocked using TBS-Tween20 buffer (25 mM Tris-Cl, pH 7.4, 150 mM NaCl, 0.05% Tween-20) containing 5% milk for 1 hour and incubated overnight with TBS-Tween20 buffer containing anti-C-terminal antibody diluted to 1:1,000. The membrane was washed with TBS buffer 3 times for 10 minutes and incubated for 1 hour with secondary ECL horseradish peroxidase (HRP)-conjugated anti-rabbit antibody (GE Healthcare Bio-Sciences Corp., Piscataway, NJ) diluted to 1:1,000. The bands were visualized using SIGMAfast 3,3-diaminobenzidine (DAB) tablets (Sigma-Aldrich, St. Louis, MO). InVision His-tag In-gel Stain (Invitrogen Corp., Carlsbad, CA) used to detect the proteins with hexa-Histidine tag followed the standard protocol.

*Purification of hCTR1$_{TM}$*

Cells containing the expressed target proteins were resuspended in 20 mL of lysis buffer (50 mM sodium phosphate, 1 M NaCl, 5 mM EDTA, pH 7.6) per gram of wet cells with protease inhibitor cocktail (cOmplete, Roche Applied Science, Indianapolis, IN). The resuspended cells were lysed by sonicating on ice for 4 minutes (5 seconds on, 5 seconds off) three times. Unlysed cells and inclusion body fractions were removed by centrifugation at 6,000×g for 30 minutes at 4 °C using a Beckman Avanti J-20 XP centrifuge with a JA-25.50 rotor. The membrane fraction containing the target protein was separated from the fraction containing soluble proteins by centrifuging the supernatant at 90,000×g for 3 hours at 4 °C using a

Beckman Optima L-100 XP Ultracentrifuge with a 45Ti rotor. The isolated membrane fraction was then dissolved in 1% w/v dodecylphosphocholine (DPC, Fos-Choline-12, Anatrace, Inc., Maumee, OH) in 20 mL of buffer A (50 mM sodium phosphate, 1 M NaCl, pH 7.6) per gram wet cells and incubated overnight at 4 °C with gentle agitation. The extracted solution was centrifuged at 90,000×g for 20 minutes at 4 °C to separate undissolved materials left in the membrane fractions.

After the selective extraction, the supernatant containing the target protein dissolved in DPC was incubated with Ni-NTA resin (Qiagen, Valencia, CA) pre-equilibrated with buffer A, for 1 hour at 4 °C with gentle agitation. The protein-DPC micelle mixture bound to Ni-NTA resin was loaded onto an empty column, and washed with >2 bed volumes of 0.2% w/v DPC solution containing 50 mM sodium phosphate, 1 M NaCl, and 5 mM imidazole at pH 7.6. Nonspecifically bound proteins were removed by washing with >2 bed volumes of buffer containing 0.2% w/v DPC, 50 mM sodium phosphate, 1 M NaCl, and 30 mM imidazole at pH 7.6. The target protein was eluted with buffer containing 0.2% w/v DPC, 50 mM sodium phosphate, 1 M NaCl, and 300 mM imidazole.

*NMR sample preparation and experiments*

5 mM EDTA was added to the sample purified with the nickel affinity chromatography, and the sample was dialyzed against water for ~2 days in a dialysis bag with the molecular weight cutoff of 1 kDa. The precipitated sample was collected by centrifugation followed by lyophilization. The NMR sample was made by directly

adding a stock solution of 100 mM DPC to the lyophilized protein powder. 10% $D_2O$ was added to the sample, and the pH was adjusted to 5.0.

Standard $^{15}N$,$^{1}H$-HSQC experiments were performed using a Bruker DRX 600 spectrometer with a TXI-cryoprobe or a Magnex 800 MHz spectrometer equipped with a Bruker AVANCE console at the University of California, San Diego. The obtained spectra were processed with NMRPipe (Delaglio et al., 1995), and analyzed with SPARKY (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco).

## 5.3 Results and Discussion

*Cloning, Expression, and Purification of $hCTR1_{TM}$*

hCTR1 consists of 190 amino acids with three putative transmembrane domains and is known to oligomerize in vivo or in vitro, posing one of the largest helical membrane proteins studied by NMR spectroscopy. The challenges associated with this protein are not limited to the sample preparations, i.e., difficulties of over-expressing the target protein in *E. coli*, purifying it to a high degree, and/or reconstituting in the membrane with proper folding, but also with the methods in terms of NMR spectroscopy. The methodological and mechanical advances in NMR spectroscopy including TROSY-related pulse sequences and high field magnet have increased the size limit of the systems for which NMR spectroscopy can be used. However, the complicated dynamics of a membrane protein-micelle complex as well as the limited chemical shift dispersions prevent the success in the studies of

membrane proteins even though the globular proteins with comparable sizes are routinely studied by NMR spectroscopy. The difficulties can often be avoided by designing the protein construct carefully and obtaining the 'partial' structure to aid the 'full' structure. This is analogous to making a 'truncated' protein in the X-ray crystallography to help the crystallization processes.

In the current study, the plasmid pEThCTR1$_{TM}$, shown in Figure 5-1 (A), was designed so that the target protein is expressed in *E. coli* with a hexa-Histidine ($His_6$) tag attached at the C-terminus for facilitating the purification with immobilized metal affinity chromatography. The target protein, hCTR1$_{TM}$, consists of residues 45-190 of the hCTR1, which contains all three transmembrane domains and one putative metal-binding motif located between TM2 and TM3, but not the other metal-binding motif located in the N-terminal region. In fact, the other metal-binding motif is located just before the N-terminus of the current construct. The $His_6$ tag was deliberately placed on the C-terminus to prevent any possible nonspecific interactions that can be present when the tag is placed on the N-terminus of the protein. The pET-31b(+) (EMD Biosciences, Inc., San Diego, CA) plasmid, originally developed to over-express the target protein with the ketosteroid isomerase (KSI) fusion protein, was used to design the plasmid pEThCTR1$_{TM}$, with the KSI-encoding region removed by the restriction enzymes *Nde* I and *Xho* I.

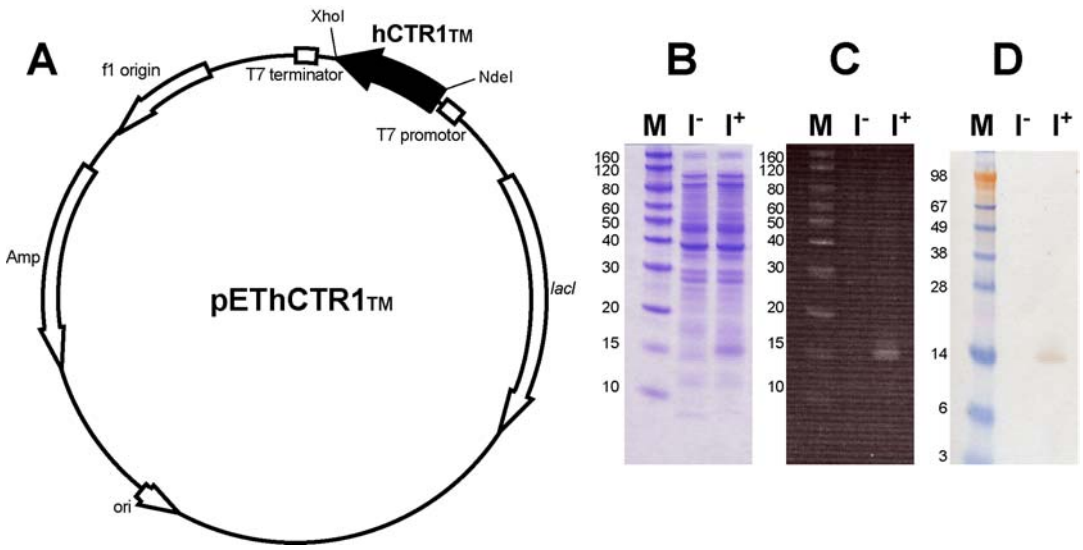


**Figure 5-1.** (A) A schematic diagram of the plasmid pEThCTR1$_{TM}$. The expression of hCTR1$_{TM}$ was confirmed by (B) SDS-PAGE, (C) InVision His-tag In-gel Stain, and (D) Western blot detected with anti-C-terminal hCTR1. Lanes marked as "M" indicate the protein standards, with the standardized molecular weights in kDa indicated next to the lanes. Samples containing the whole cells before the induction were labeled "I$^-$", and 4 hours after the induction, "I$^+$". The cells loaded onto each lane were 'normalized' according to $OD_{600}$ values.

After transformation into different strains of bacteria and optimization of growth conditions, C43(DE3) was found to give the expression level suitable for NMR studies. C43(DE3) and C41(DE3) are the mutant strains of *E. coli* strain BL21(DE3), that are known to survive when the expressed proteins are toxic to the host (Miroux and Walker, 1996). Although the majority of proteins over-expressed in these kinds of hosts seem to be localized in the inclusion body, it is still not clear how the hosts can tolerate the toxicity of the target proteins when they are over-produced. Nonetheless, it is quite important that the suitable host-plasmid combination should be tested in order to achieve high expression levels. For the current plasmid, when the same growth and expression protocol was applied, the host BL21(DE3) did not survive (the

solution becomes clear and $OD_{600}$ goes down) 2 hours after the induction at 37 °C, but C43(DE3) could survive 5 hours after the induction.

The bacterial culture performed with a rich media such as the Luria-Bertani (LB) broth did not yield expression in detectable level. However, when switched to M9 minimal media, the expression level went up. This is a fortunate case for an NMR study where minimal media is required for introducing isotopes using the conventional method. Further optimization was required to increase the expression level of target protein and the possibility of the presence of any rare codons in the DNA sequence was examined. Transformation of the plasmid into competent *E. coli* cells designed for overcoming the codon-bias problem, such as Rosetta 2 (DE3) (EMD Biosciences, Inc., San Diego, CA) yielded a slight (but noticeable) increase in the expression level, and a single rare codon was found at the position encoding $Arg^{90}$. Therefore, the codon CGA encoding $Arg^{90}$ was modified to a more 'abundant' codon CGC by standard site-directed mutagenesis. Varying IPTG concentrations from 0.2 mM to 1 mM did not change the expression level significantly with the optimum IPTG concentration being around 0.4 mM, but adding thiamine did help increase the expression level although the BL21 strain is not vitamin B1 deficient. Also, addition of 3-5% LB media in the minimal media helped increase the expression level as well as the growth rate.

Expression of $hCTR1_{TM}$ after the optimization is clearly visible in the SDS-PAGE gel shown in Figure 5-1 (B). $hCTR1_{TM}$ consists of a total of 154 amino acids (residues 45-190 of hCTR1 plus the amino acid sequence LEHHHHHH at the C-

terminus), and the molecular weight is calculated to be 17.4 kDa. The band that appeared at around 15 kDa after the induction with IPTG indicates that the expression level is high enough to pursue the structural studies with NMR spectroscopy. The cell concentrations loaded in each of the lanes were normalized according to the $OD_{600}$, meaning the SDS-PAGE shows the *relative* level of protein in the samples. Various induction times have also been tested to optimize the expression level, and significant reduction in $OD_{600}$ was observed after 6 hours of induction at 37 °C. Proteins with the exposed $His_6$ tag were detected using InVision His-tag In-gel Stain (Invitrogen, Corp., Carlsbad, CA). The InVision His-tag In-gel Stain consists of a fluorescent dye conjugated to a nickel-nitrilotriacetic acid (NTA) complex and allows the specific detection of proteins with a polyHis-tag when exposed to the ultraviolet radiation, as seen in Figure 5-1 (C). An immunoblot using the antibody raised against the C-terminal region of hCTR1 shown in Figure 5-1 (D) suggests that the protein expressed after the induction contains the same amino acid sequence as the C-terminal region of hCTR1, proving that the 15kDa band that appears after induction contains the target protein. Lowering the incubation temperature following induction is known to reduce the toxicity caused by the over-production of membrane proteins and localize the target protein within the membrane, thus increasing the final yield of the protein. For $hCTR1_{TM}$, inducing at 25 °C for 12-16 hours resulted in significantly higher cell density than inducing at 37 °C for 5 hours, however, an increase in the amount of purified protein per liter of culture was not observed.

Once a sufficient level of expression was achieved, the harvested cells were resuspended in the standard lysis buffer for *E. coli*, i.e., sodium phosphate buffer containing ethylenediamine tetraacetate (EDTA) and NaCl. The presence of protease inhibitors was necessary to keep the protein intact at this particular stage. The pellet containing the inclusion bodies and unlysed cells was separated from the solution containing the soluble proteins and the membrane proteins that reside in the lipid membranes by low speed centrifugation. A majority of the fraction that corresponds to $hCTR1_{TM}$ was found in the solution according to SDS-PAGE, suggesting $hCTR1_{TM}$ is expressed within the membrane even at high concentration. This is in agreement with the previous reports about the subcellular localization of hCTR1 (Lee et al., 2002; Eisses et al., 2005). Many attempts were made to isolate the target protein from the soluble proteins by solubilizing the membrane proteins using various detergents such as Triton X-100 and Empigen BB followed by purification using the metal affinity chromatography, but the purity and the yield were not high enough for NMR studies. The purity was increased when the total membrane fraction was separated from the solution using high speed centrifugation, and dissolved in a high concentration of detergent. For this purpose, the detergent used to extract membrane proteins from the membrane fraction should not be too harsh, because the purpose of the detergent extraction method is to obtain the membrane proteins in a native state, and harsh detergents might denature the protein. When a solution of the mild detergent n-dodecyl-β-D-maltoside (DDM) was used, extraction of $hCTR1_{TM}$ from the membrane

was very efficient, while a solution of n-octyl-β-D-glucopyranoside (OG) was not able to solubilize the target protein at all, as judged by SDS-PAGE.

$Ni^{2+}$ affinity chromatography in the presence of 0.2% DDM provided successful purification of $hCTR1_{TM}$ after optimizing the conditions such as the salt concentration, pH, detergent concentration, and imidazole concentration. However, when the final sample was collected and the $^{15}N$,$^{1}H$-HSQC spectrum was obtained, the presence of the detergent DDM had adverse effects on the spectrum, and the low critical micelle concentration (CMC) of DDM (~0.2 mM) prevented complete removal using dialysis. Dodecylphosphocholine (DPC) has been established as an excellent micelle system, and successfully applied to a number of NMR studies on membrane proteins, including Pf1 coat protein (Shon et al., 1991), diacylglycerol kinase (Oxenoid et al., 2004), phospholamban (Oxenoid and Chou, 2005), PagP (Hwang et al., 2002) and OmpA (Arora et al., 2001). When 1% DPC solution was added to the membrane fraction isolated by high speed ultracentrifugation and incubated overnight at 4 °C, a majority of $hCTR1_{TM}$ was solubilized and ready to be loaded onto the Ni-NTA resin. Because 1% DPC (~30 mM) was able to effectively solubilize $hCTR1_{TM}$ from the membrane fraction, it was possible that the NMR sample could be made under similar conditions. Nickel affinity chromatography was carried out in the presence of 0.2% DPC and the binding of $hCTR1_{TM}$ to Ni-NTA (Qiagen, Valencia, CA) resin was very effective even without the presence of reducing agent (see below) and any nonspecifically bound proteins were removed by 30 mM imidazole. The purification scheme is briefly illustrated in Figure 5-2 (A), and the SDS-PAGE of the

sample from the last step of purification is shown in Figure 5-2 (B) where the purity of the final sample is clearly seen. When the purified sample was loaded with a strong reducing agent DTT, the major band was detected at around 15 kDa corresponding to the size of $hCTR1_{TM}$ in the monomeric state, and a weak band at around 28 kDa was also detected. When the same sample was loaded onto an SDS gel without DTT, the strong band appeared at around 28 kDa, corresponding to the size of $hCTR1_{TM}$ in a dimeric state. This is in excellent agreement with the previous reports that show that full-length hCTR1 migrates as a monomer and a dimer in a non-reducing SDS gel (Eisses and Kaplan, 2002; Lee et al., 2002; Klomp et al., 2003). The current study shows that the oligomerization state of hCTR1 is preserved in $hCTR1_{TM}$ which is over-expressed in *E. coli* and purified. Note that the samples loaded on the SDS gel were purified by nickel affinity chromatography: the protein was dissolved in 0.2% DPC, 50 mM sodium phosphate pH 7.6, 1 M NaCl and 300 mM imidazole, before the lithium dodecyl sulfate (LDS) sample buffer (Invitrogen Corp., Carlsbad, CA) was added. This suggests that the 'native conformation' is not disrupted by the presence of DPC, which is a similar result to that reported for the pentameric phospholamban (Oxenoid and Chou, 2005).

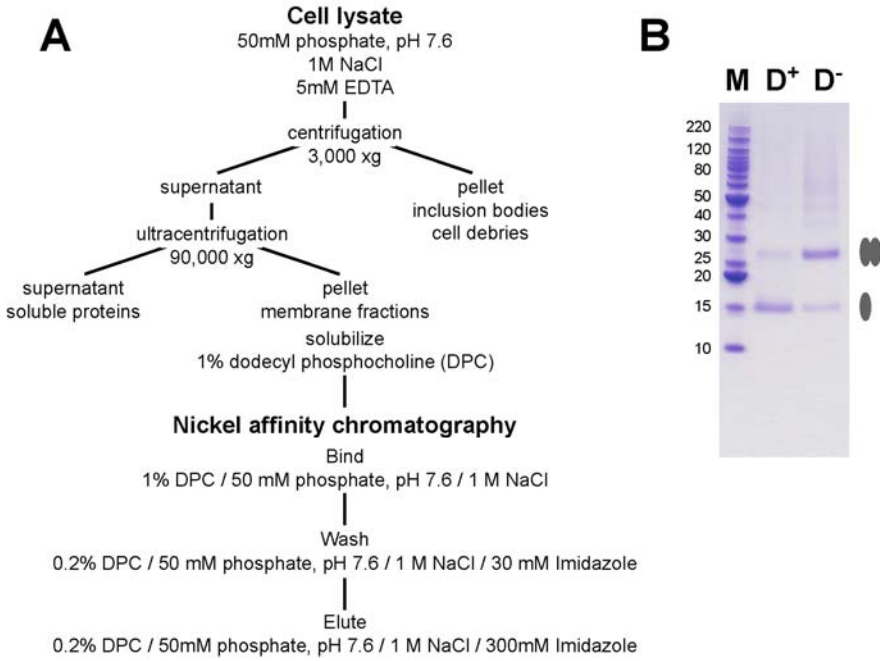


**Figure 5-2.** (A) Purification scheme of $hCTR1_{TM}$. (B) SDS-PAGE of $hCTR1_{TM}$ after $Ni^{2+}$ affinity chromatography in the presence of DTT ($D^+$) and absence of DTT ($D^-$). Ovals indicate the plausible oligomerization states of the protein. Benchmark protein standard (Invitrogen Corp., Carlsbad, CA), labeled "M", was used as a molecular weight standard with the standardized molecular weight values in kDa indicated next to the lane.

$^{15}N,^{1}H$-HSQC experiments were performed on the purified $^{15}N$-$hCTR1_{TM}$ sample. The best quality spectrum was acquired when the protein (~1 mM) was dissolved in 150 mM DPC, at pH 5.0, shown in Figure 5-3. The sample from nickel affinity chromatography was dialyzed for 2 days against water and lyophilized to powder. The NMR sample was made by adding the detergent solution directly to the lyophilized powder. The high detergent concentration was required to solubilize the high concentration of protein required for the NMR experiments, and the elevated temperature was required for narrow resonance linewidths. Interestingly, when reducing agent such as DTT or tris(2-carboxyethyl)phosphine hydrochloride (TCEP-

HCl) was added to the NMR sample, a $^{15}N,^{1}H$-HSQC spectrum of poor quality was obtained: the spectral quality of the $^{15}N,^{1}H$-HSQC spectrum changes according to the redox state of the cysteine residues. This result was highly reproducible. At this point, the reason for the poor quality spectra in the presence of reducing agent is not clear. Conformational heterogeneity in the presence of reducing agent might explain this result, but the addition of reducing agent and incubation for an extended time period did not help improve the spectral quality, implying that the poor spectral quality is not a result of proteins with different redox state. Poor spectral quality originating from reduced cysteine residues would potentially cause problems when the spin label is introduced to the cysteine residue, preventing it from forming intermolecular disulfide bonds.

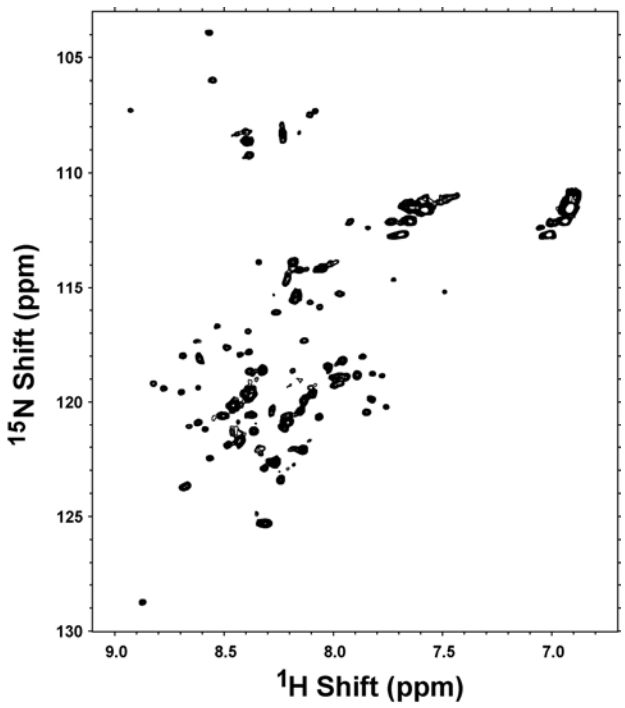


**Figure 5-3.** $^{15}N,^{1}H$-HSQC spectrum of $hCTR1_{TM}$ in 150 mM DPC at 323 K, pH 5.0.

Because different detergents create different environments for membrane proteins, many other detergents for NMR samples were tested, in an effort to find the detergent that would produce the high-quality spectrum in the presence of reducing agent. The detergents include 1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-RAC-(1-glycerol)] (LMPG), 1,2-dicaproyl-1-sn-glycero-3-phosphocholine (DHPC), n-octyl-β-D-glucopyranoside (OG), sodium dodecyl sulfate (SDS), and DDM (n-dodecyl-β-D-maltoside). Unfortunately, most of the tested detergents were not compatible with $hCTR1_{TM}$: low solubility of the protein in the tested detergents, LMPG, DHPC, and OG, yielded no detectable signals in the $^{15}N,^{1}H$-HSQC spectrum. The other detergents, DDM and SDS, solubilized the protein, but well-resolved $^{15}N,^{1}H$-HSQC spectra were not obtained in these detergens as compared to the spectrum in DPC.

*Mutation of Cysteine Residues and the Effect on the Oligomerization*

Introducing a single cysteine residue within the protein of interest seems necessary in order to apply the method described in previous chapter. To investigate the applicability of the method to the current system, the cysteine residues within the protein were substituted into serine residues. The presence of a single cysteine residue is required to add a nitroxide spin label to the cysteine residue, and, because there are two cysteine residues in the $hCTR1_{TM}$, $Cys^{161}$ and $Cys^{189}$, these residues must be mutated. The presence of two cysteine residues within the target protein makes it easier to decide where to place the spin labels. Since $Cys^{161}$ is located in the middle of the third transmembrane helix, labeling this residue with nitroxide could potentially

provide sufficient distance information for obtaining the backbone structure. In the current study, three mutants were generated to examine the effect of cysteine substitutions: two single amino acid substitution mutants, C161S and C189S, and a double substitution mutant, C161S/C189S. Adding the nitroxide spin label to a single mutant, i.e., C161S or C189S, will break the intermolecular disulfide bond that is associated with the cysteine residue left in the mutant, and the analysis on the double mutant C161S/C189S will reflect the result of spin labeling on either one of the cysteine residues.

All the mutants were generated by following the simple site-directed mutagenesis protocol, and were transformed into the same cell line, C43(DE3). The same expression protocol was applied to all mutants. Fortunately, all the mutants were expressed at high levels within membrane fractions, suggesting that the expression and purification protocol for the wild type can be applied to the mutants. Figure 5-4 (A) shows the SDS-PAGE result of the cysteine mutants and the wild type after nickel affinity chromatography purification.

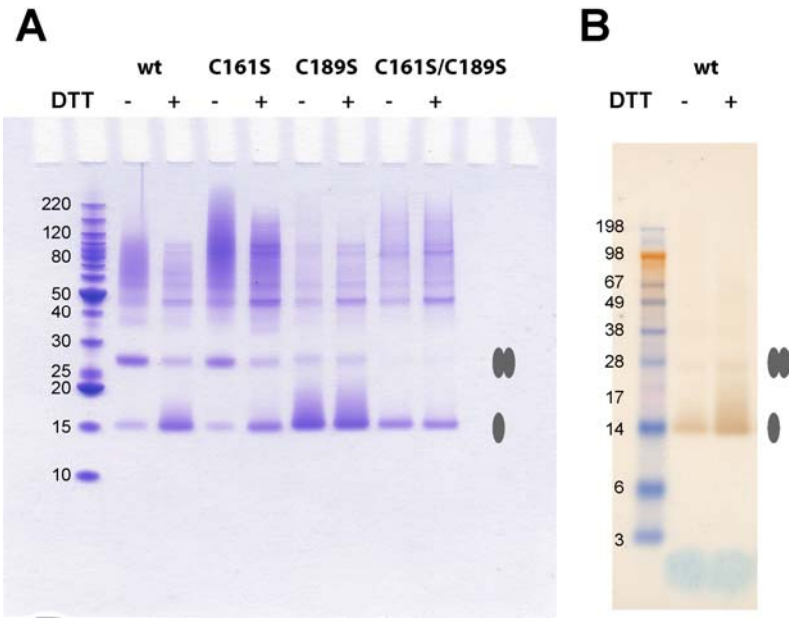


**Figure 5-4.** Various $hCTR1_{TM}$ cysteine mutants were analyzed by (A) SDS-PAGE and (B) Western blot detected using C-terminal hCTR1 antibody. The presence or absence of DTT was indicated as + or - on top of the figures. The grey ovals displayed to the right indicate the locations of the plausible oligomeric states of $hCTR1_{TM}$.

Clearly, the presence of the reducing agent DTT affects the oligomerization state of wild type and C161S, according to the results presented here. The intensities of the bands corresponding to the monomeric state of $hCTR1_{TM}$, as indicated as single grey oval in the figure, increase when DTT was added to the wild type or C161S. However, the intensities of the bands for C189S and C161S/C189S are not affected by the presence of DTT. This result indicates that the dimeric state of $hCTR1_{TM}$ is stabilized by $Cys^{189}$, which is located at the end of C-terminus. Using an antibody directed against the C-terminal residues of hCTR1, the bands corresponding to the monomeric state of $hCTR1_{TM}$ were detected as shown in Figure 5-4 (B). Unfortunately, this antibody failed to detect the dimer bands in the absence of DTT.

This result is partially in agreement with the previous studies from the Kaplan group that showed cysteine residues are responsible for the stable oligomerization of hCTR1, as detected from the DDM-solubilized membrane fraction (Eisses and Kaplan, 2002) or from the total membrane fraction (Eisses and Kaplan, 2005). However, these studies suggest that both of the cysteine residues are responsible for the oligomerization and they showed the synergistic effect on the copper transport, whereas the current result suggests that residue $Cys^{189}$ is solely responsible for the stable oligomerization. The relationship between the oligomerization through the cysteine residue(s) and the copper transport mechanism is still not clear, but the oligomerization seems to occur through $Cys^{189}$ in vitro, in DPC micelles. Other studies by Unger group suggested the oligomerization of hCTR1 is stabilized by the Gly-X-X-X-Gly (GG4) motif in the third transmembrane domain through helix packing (Aller et al., 2004), and later, the same group reported a 6-Å resolution projection structure of hCTR1 using electron microscopy, supporting the importance of the third transmembrane domain for multimerization and helix packing and the trimeric arrangement (Aller and Unger, 2006) that was initially suggested by Thiele group using crosslinking experiments (Lee et al., 2002). However, the trimeric arrangement cannot be explained by the formation of intermolecular disulfide bonds if only one of two cysteine residues, $Cys^{189}$ is responsible for the oligomerization, and Unger group also noticed the disulfide-dependent profile during the purification using gel filtration chromatography.

These intriguing, yet contradictory results directed the investigation toward NMR studies. The $^{15}N,^{1}H$-HSQC spectra for all the cysteine mutants were obtained and shown in Figure 5-5. The quality of $^{15}N,^{1}H$-HSQC spectra for these mutants will determine the feasibility of further NMR measurements, such as PRE measurements using nitroxide spin labels, as well as future directions. The results definitively show that the quality of $^{15}N,^{1}H$-HSQC spectra is strongly correlated with the oligomerization states of $hCTR1_{TM}$: high-quality $^{15}N,^{1}H$-HSQC spectra are obtained only for the wild type $hCTR1_{TM}$ or C161S. The quality of $^{15}N,^{1}H$-HSQC spectra for C189S or C161S/C189S shown in this figure is exactly the same as those for wild type $hCTR1_{TM}$ in the presence of DTT, and this proves that the spectral quality of the wild type is mostly affected by the redox state of $Cys^{189}$, not by the conformational heterogeneity. The $^{15}N,^{1}H$-HSQC spectra obtained for C189S or C161S/C189S exhibit the typical characteristics of those from unfolded proteins, as judged by intense peaks and limited chemical shift dispersion. This suggests that the mutation at the $Cys^{189}$ residue breaks the oligomerization of $hCTR1_{TM}$, causing the protein to be unfolded. The intermolecular disulfide bond formation was not reversible in this case: once DTT was added to the wild type $hCTR1_{TM}$, raising pH, adding hydrogen peroxide, adding a thiol-specific oxidizing agent N,N,N′,N′-Tetramethylazodicarboxamide (diamide) and/or re-purifying in the absence of DTT did not recover the spectral quality.

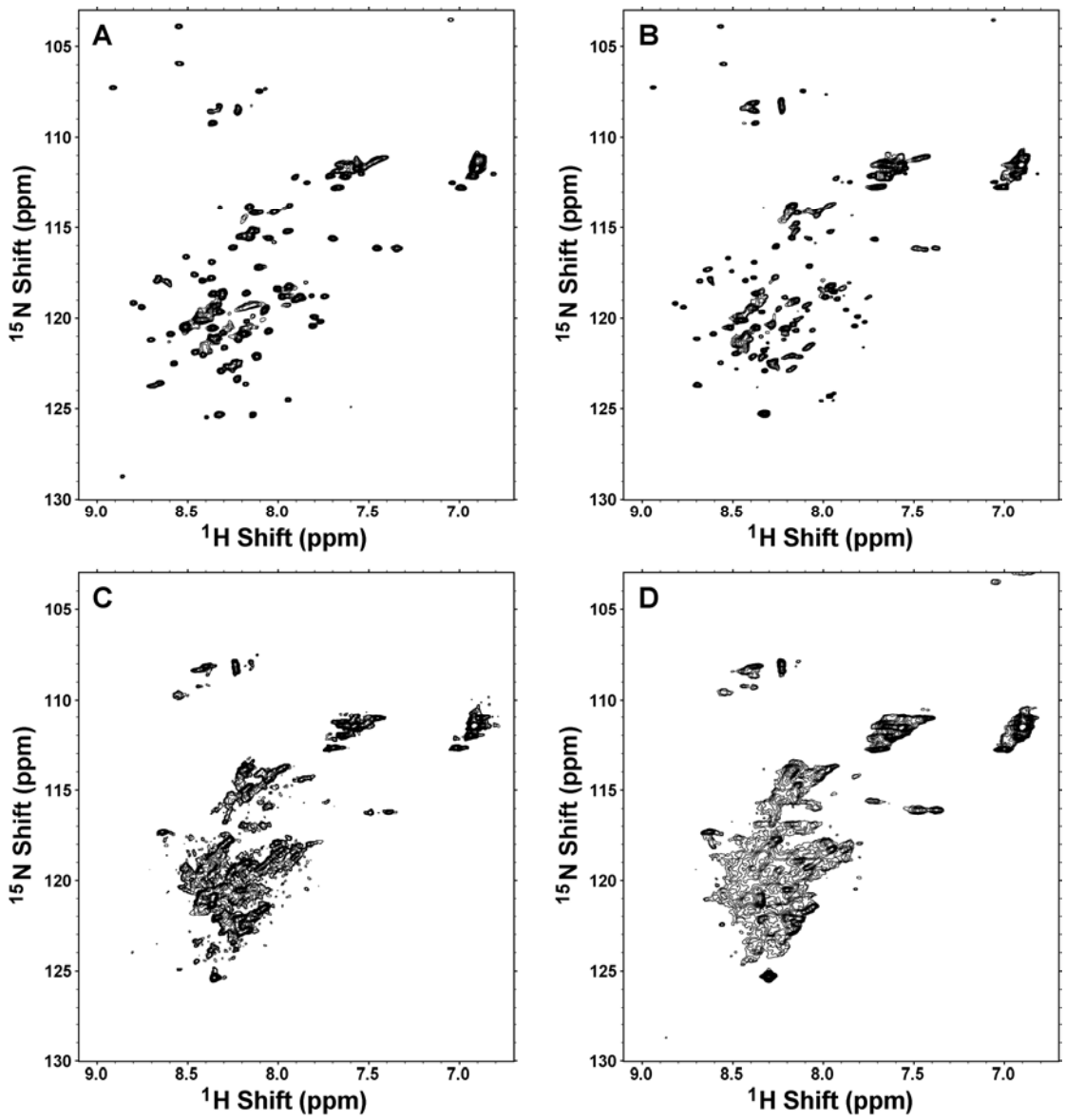


**Figure 5-5.** $^{15}N,^{1}H$-HSQC spectra of $hCTR1_{TM}$ cysteine mutants, (A) wild type, (B) C161S, (C) C189S, and (D) C161S/C189S. Samples were dissolved in 100 mM DPC, and the pH was adjusted to 5.0. All the spectra were acquired at 323 K.

This result is rather unexpected since the linewidths and intensities of NMR signals, indicatives of the motional properties of a protein, depend on the size of the system, and yet the $hCTR1_{TM}$ in the dimeric state (~60 kDa protein-micelle complex)

produces a high-quality $^{15}N,^{1}H$-HSQC spectrum without the need of deuteration and TROSY-like pulse sequences. Previous NMR studies on globular proteins of similar size such as maltose binding protein (42 kDa, Gardner et al., 1997), β-barrel membrane proteins (~50 kDa protein-micelle complex, Arora et al., 2001, Hwang et al., 2002), or a trimeric α-helical membrane protein with three transmembrane domains (~90 kDa protein-micelle complex, Oxenoid et al., 2004) have shown that the necessity of perdeuteration and TROSY-type pulse sequence for the proteins of these sizes. Although the factors related to the relaxation properties that affect resonance linewidths or intensities can be different in each case, the current study suggests that the quality of NMR spectra is affected by the population of properly folded proteins more than the size of the system. Unfortunately, attempts to express the current target protein in media with highly deuterated background yielded very little protein, making it impossible to pursue further NMR studies, and whether or not the high level of deuteration would help improve the spectral quality is currently unknown.

PRE measurements through the nitroxide spin label require the presence of a single cysteine residue, and the cysteine residue will not be able to form a intermolecular disulfide bond once it is spin-labeled: therefore, spin-labeled C161S or spin-labeled C189S will produce the $^{15}N,^{1}H$-HSQC spectra similar to the spectrum for C161S/C189S. Inability to form a stable intermolecular disulfide bond, thus producing unresolved $^{15}N,^{1}H$-HSQC spectrum, makes C161S or C189S inapplicable for the PRE measurements using nitroxide spin labels, and developments should be made toward the introduction of rigid paramagnetic centers without disrupting the

structural disulfide bonds. It is also an interesting area of development that can be used for the analysis of the PRE measurement on the homodimeric complex, since it should distinguish between intramolecular PRE and intermolecular PRE.

## 5.4 Concluding Remarks and a Plausible Model

The main purpose of the work presented in this chapter is to obtain structural information of hCTR1 with three transmembrane domains, and in order to achieve that goal, the methods described in the previous chapters were applied. In the course of applying these methods, the difficulties associated with the intrinsic properties of the protein were apparent, i.e., protein folding through the intermolecular disulfide bonds. By generating the necessary mutants, the cysteine residues responsible for the intermolecular disulfide bonds were identified. Although the current method cannot be applied to this complicated system, the result provided some insights about the structure of hCTR1. The suggested model based on the current data as well as the previous reports is depicted in Figure 5-6. In (A), a model shows the dimeric arrangement with the proteins connected through a intermolecular disulfide bond between $Cys^{189}$. Previous reports by Unger group point out the importance of the GG4 motif within the third transmembrane domain for helix packing (Aller et al., 2004), and suggest the third transmembrane domain may be located on the inside of the 'pore' based on the projection density map from the electron microscopy (De Feo et al., 2007). The locations of the first and second transmembrane domains are currently unknown, although N-terminal extracellular domain was reported to interact with itself

(Klomp et al., 2003). This self-association between N-terminal extracellular domains suggest the hypothetical model shown in (B), where the top view of the 'trimer of dimer' model is depicted based on the previously reported projection density map (Aller and Unger, 2006). The circles with the same color represent the transmembrane helices within the dimer assembled through the disulfide bond as shown in (A), and different colors represent different subunits of the dimers. The model shown here assumes that the third transmembrane domains are located on the inside of the 'pore', and the trimeric arrangement comes from the interactions between the N-terminal extracellular domains, which is excluded in the current study. Note that this model is based on the assumption that the dimerization occurs first through the intermolecular disulfide bond between the $Cys^{189}$ residues. It is possible that the other residues, such as the GG4 motif, contribute to the oligomerization process by assisting the helix packing. However, the current study shows that residue $Cys^{189}$ plays a critical role in the oligomerization and folding process, and it is impossible to explain the oligomerization through the intermolecular disulfide bonds with the trimer model that has been suggested previously. Obviously, more detailed information will be revealed once the high-resolution structure is obtained, and methodological development needs to be made in order to solve the structure.

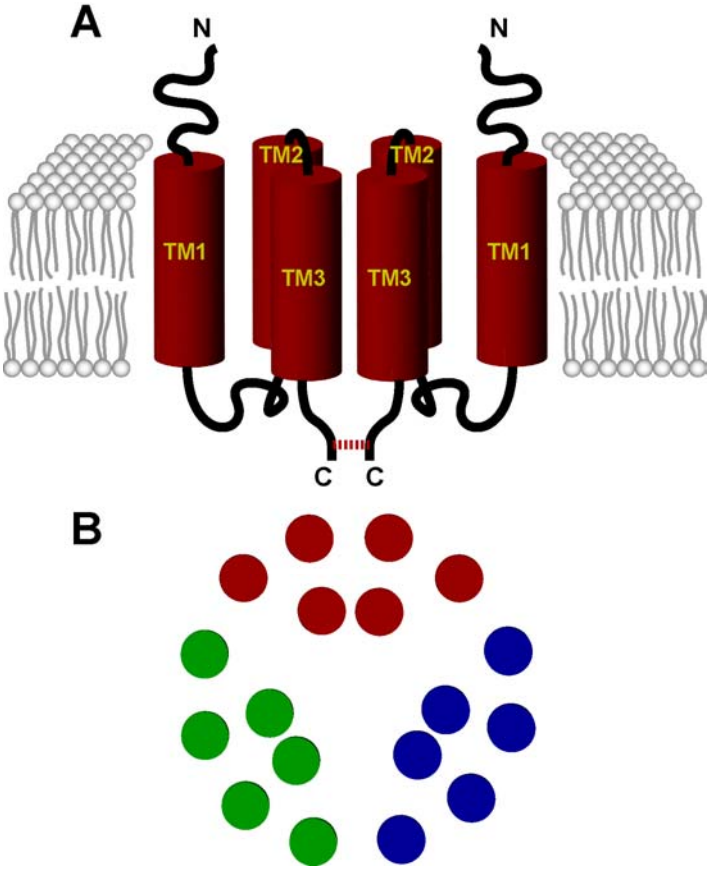


**Figure 5-6.** A proposed model of hCTR1 based on the current results. (A) The possible dimeric arrangement of hCTR1, with the dotted line representing the intermolecular disulfide bond between $Cys^{189}$ residues that seems to be necessary for the correct folding. The third transmembrane domain helix was arranged to be located inside of the core as suggested by previous report (Aller and Unger, 2006). (B) Top view of hypothetical 'trimer of dimer' arrangement of hCTR1.

In summary, the significance of the work presented in this chapter is the following: (1) Cloning, expression and purification of the transmembrane domain of hCTR1, $hCTR1_{TM}$, was successfully performed in *E. coli*, and the level of expression was sufficient for structural studies. (2) The first high-quality NMR spectrum of $hCTR1_{TM}$, consisting of three transmembrane domains, was obtained. (3) Residue $Cys^{189}$ was identified to be responsible for the oligomerization of $hCTR1_{TM}$ through

intermolecular disulfide bonds. (4) The quality of NMR spectra changes according to the oligomerization state of $hCTR1_{TM}$, proving the oligomerization, thus the presence of residue $Cys^{189}$, is required for correct folding.

# 6 Conclusion

The research described in this dissertation explores the experimental and practical aspects of structure determination of membrane proteins using solution NMR spectroscopy. Measurements of RDCs for samples weakly aligned in polyacrylamide gels were optimized for membrane proteins in micelles and applied to several model systems, including fd coat protein and Pf1 coat protein. Polyacrylamide gels were able to create anisotropic interactions large enough to be measured with simple procedures. These measurements were then used as an input for the subsequent Dipolar Wave analysis which determines the relative orientations between the helices. The simplicity of this approach motivated its application to the structural genomics on membrane proteins, using Rv1761c from *M. tuberculosis* as an example, with the aid of the accelerated resonance assignment process using data acquisition incorporating non-linear sampling schedules followed by maximum entropy reconstruction. Orientational degeneracy problems inherent in the dipolar coupling measurements were complemented by long-range distance information from PRE measurements using site-directed spin labeling techniques. The backbone structure of Rv1761c was successfully determined using orientational information from RDC measurements and translational information from PRE measurements, suggesting the possibility of rapid structure determination of membrane proteins using minimal experimental information from solution NMR spectroscopy. This approach will be applied to other selected structural genomics targets. However, generalization of the developed methods

requires extreme caution because of the potential disruption of structural elements when a mutation is involved. This was illustrated with cysteine mutants for $hCTR1_{TM}$, where the quality of NMR spectra was directly associated with the oligomeric states.

The success of the current methods relies heavily upon a single factor: sample preparation. It is important to make a homogeneous isotropic sample to obtain NMR spectra with high resolution, good signal-to-noise ratio, and the correct number of signals; however, the results presented in this dissertation suggest that the necessary orientational information cannot be obtained if a high-quality weakly aligned sample is not available and such sample can only be acquired through careful optimization. Unfortunately, this limits the number of environments that membrane proteins can be studied in. For example, although it is of great interest to study membrane proteins in isotropic bicelles using solution NMR spectroscopy, their inability to diffuse into polyacrylamide gels prevent the acquisition of structural information using the current approach. Therefore, the search for a better alignment medium should be continued. Sample preparation is also critical in terms of obtaining PRE measurements. The requirement of a rigid paramagnetic center limits the kinds of spin probes that can be used, as well as the methods to incorporate such probes. The necessary presence of a single cysteine residue for labeling with the spin labeling agent creates the possible interference within the structural elements, suggesting a careful design in mutational strategy. Removal of structural elements crucial for proper folding prohibits further study, highlighting the need for the developments of paramagnetic centers.

The structure determination of membrane proteins is an important aspect of contemporary structural biology. These proteins require the presence of lipids to maintain their structural and functional integrity. However, the lipids complicate the preparation of samples for experimental measurements. As a result, it is of interest to make comparisons among various types of lipid assemblies, including those that assemble into micelles, bicelles, and bilayers. The flexibility in sample conditions is a valuable feature of NMR approaches to structure determination, and the Dipolar Waves provides the general approach to interpret the experimental data, enabling the direct comparison between the structures from different environments. The method of utilizing dipolar couplings can be combined with PRE measurements to determine the backbone structure of membrane proteins in a short period of time.

The number of structures determined by the data based on the anisotropy and paramagnetism is expected to grow, generalizing and improving the method presented in this dissertation.

# Appendices

**A.** $^{1}$HN and $^{15}$N chemical shifts (ppm) for Pf1 coat protein in SDS micelles at 313K, pH 4.0.

| Residue | $^{1}$HN Shift | $^{15}$N Shift |
|---|---|---|
| G1 | | |
| V2 | 8.203 | 122.154 |
| I3 | 7.836 | 117.603 |
| D4 | 8.519 | 118.253 |
| T5 | 8.349 | 115.577 |
| S6 | 8.397 | 116.981 |
| A7 | 7.879 | 124.592 |
| V8 | 7.705 | 117.477 |
| E9 | 8.310 | 117.537 |
| S10 | 8.197 | 114.269 |
| A11 | 7.866 | 123.523 |
| I12 | 7.947 | 114.667 |
| T13 | 7.997 | 111.334 |
| D14 | 8.227 | 119.572 |
| G15 | 8.167 | 108.239 |
| Q16 | 8.176 | 118.437 |
| G17 | 8.515 | 109.077 |
| D18 | 8.342 | 118.371 |
| M19 | 8.406 | 120.074 |
| K20 | 8.100 | 118.165 |
| A21 | 7.523 | 119.750 |
| I22 | 7.770 | 114.170 |
| G23 | 8.448 | 106.233 |
| G24 | 8.319 | 107.657 |
| Y25 | 8.032 | 118.843 |
| I26 | 8.005 | 117.223 |
| V27 | 8.337 | 117.660 |
| G28 | 8.576 | 105.165 |
| A29 | 8.386 | 121.882 |
| L30 | 8.206 | 116.192 |
| V31 | 8.362 | 117.566 |
| I32 | 8.128 | 118.379 |
| L33 | 8.463 | 118.788 |
| A34 | 8.578 | 120.594 |
| V35 | 8.590 | 117.204 |
| A36 | 8.906 | 121.570 |
| G37 | 9.059 | 105.123 |
| L38 | 8.818 | 124.746 |
| I39 | 8.801 | 120.115 |
| Y40 | 9.076 | 119.733 |
| S41 | 8.551 | 113.137 |
| M42 | 8.261 | 120.102 |
| L43 | 8.421 | 119.167 |
| R44 | 8.026 | 117.063 |
| K45 | 7.684 | 118.699 |
| A46 | 7.863 | 124.925 |

**B.** N-H residual dipolar couplings (Hz) measured for Pf1 coat protein weakly aligned using an axially compressed 7% polyacrylamide gel.

| Residue | N-H RDC |
|---|---|
| G1 | |
| V2 | -1.23 |
| I3 | -1.96 |
| D4 | -0.65 |
| T5 | -1.93 |
| S6 | 0.05 |
| A7 | -3.33 |
| V8 | -2.49 |
| E9 | 0.19 |
| S10 | -0.16 |
| A11 | -2.43 |
| I12 | -1.98 |
| T13 | 0.69 |
| D14 | 1.52 |
| G15 | -3.44 |
| Q16 | -0.18 |
| G17 | -1.28 |
| D18 | 2.43 |
| M19 | 2.29 |
| K20 | -4.67 |
| A21 | -0.72 |
| I22 | 1.4 |
| G23 | -4.95 |
| G24 | -7.11 |
| Y25 | -6.62 |
| I26 | -0.93 |
| V27 | -8.1 |
| G28 | -6.07 |
| A29 | -1.68 |
| L30 | -5.84 |
| V31 | -8.36 |
| I32 | -5.19 |
| L33 | -1.8 |
| A34 | -6.35 |
| V35 | -7.34 |
| A36 | -2.86 |
| G37 | -3.16 |
| L38 | -9.14 |
| I39 | -7.07 |
| Y40 | -2.67 |
| S41 | -5.41 |
| M42 | -7 |
| L43 | -3.95 |
| R44 | -0.93 |
| K45 | -5.77 |
| A46 | -3.51 |

**C.** N-H residual dipolar couplings (Hz) measured for fd coat protein (Y21M) weakly aligned using an axially compressed 7% polyacrylamide gel.

| Residue | N-H RDC |
|---|---|
| A1 | |
| E2 | -0.75 |
| G3 | -0.87 |
| D4 | -2.75 |
| D5 | -3.69 |
| P6 | |
| A7 | -3.6 |
| K8 | -0.41 |
| A9 | -4.86 |
| A10 | -8.24 |
| F11 | -3.5 |
| D12 | -1.35 |
| S13 | -8.11 |
| L14 | -4.22 |
| Q15 | 1.13 |
| A16 | -0.98 |
| S17 | -4.85 |

| Residue | N-H RDC |
|---|---|
| A18 | 3.64 |
| T19 | 3.09 |
| E20 | 3.56 |
| M21 | 3.77 |
| I22 | -0.22 |
| G23 | -1.37 |
| Y24 | 2.46 |
| A25 | -6.67 |
| W26 | -15.53 |
| A27 | -9.3 |
| M28 | -3.85 |
| V29 | -8.05 |
| V30 | -11.2 |
| V31 | -14.49 |
| I32 | -6.87 |
| V33 | -14.71 |
| G34 | -10.3 |

| Residue | N-H RDC |
|---|---|
| A35 | -3.39 |
| T36 | -10.63 |
| I37 | -16.07 |
| G38 | -7.91 |
| I39 | -0.11 |
| K40 | -10.08 |
| L41 | -13.57 |
| F42 | -2.4 |
| K43 | -0.56 |
| K44 | -13.73 |
| F45 | -8.8 |
| T46 | 2.83 |
| S47 | -3.83 |
| K48 | -8.96 |
| A49 | 0.8 |
| S50 | -0.42 |

**D.** N-H residual dipolar couplings (Hz) measured for $Vpu_{28-81}$ weakly aligned using an axially compressed 6% polyacrylamide gel ($RDC_{compressed}$) and a radially compressed 6% polyacrylamide gel ($RDC_{stretched}$)

| Residue | $RDC_{compressed}$ | $RDC_{stretched}$ |
|---|---|---|
| E28 | | |
| Y29 | 3.1 | -3.42 |
| R30 | -3.95 | 4.75 |
| K31 | -0.06 | |
| I32 | 1.1 | -2.98 |
| L33 | -1.16 | 1.21 |
| R34 | -2.98 | 2.98 |
| Q35 | | 3.51 |
| R36 | 1.7 | 3.24 |
| K37 | | |
| I38 | | |
| D39 | 1.27 | -1.04 |
| R40 | 0.31 | -0.73 |
| L41 | -4.81 | 4.37 |
| I42 | -2.08 | -0.8 |
| D43 | 1.52 | -3.08 |
| R44 | | |
| L45 | -1.82 | -0.86 |
| I46 | 0.06 | 0.14 |
| E47 | 1.03 | -2.31 |
| R48 | -3.1 | 2.62 |
| A49 | -2.68 | 2.31 |
| E50 | 2.07 | -2.43 |
| D51 | 0.3 | -0.85 |
| S52 | -4.93 | 4.32 |
| G53 | 0.42 | -0.37 |
| N54 | | |

| Residue | $RDC_{compressed}$ | $RDC_{stretched}$ |
|---|---|---|
| E55 | -0.73 | 0.66 |
| S56 | 0.92 | -0.85 |
| E57 | -3.47 | 3.1 |
| G58 | -4.98 | 4.31 |
| E59 | -5.05 | 4.07 |
| I60 | -3.83 | 3.53 |
| S61 | -7.17 | 6.39 |
| A62 | -7.78 | 7.36 |
| L63 | -4.13 | 3.41 |
| V64 | | |
| E65 | | |
| L66 | -7.17 | 6.33 |
| G67 | -4.14 | 3.34 |
| V68 | -6.69 | 6.33 |
| E69 | -8.21 | 7.29 |
| L70 | -4.86 | 4.74 |
| G71 | -2.73 | 2.19 |
| H72 | -3.96 | 3.64 |
| H73 | -4.86 | 4.02 |
| A74 | -1.04 | 1.09 |
| P75 | | |
| W76 | -0.55 | 0.31 |
| D77 | 0.18 | 0 |
| V78 | -0.67 | 0.55 |
| D79 | -0.37 | 0.36 |
| D80 | 0.11 | -0.06 |
| L81 | 0.66 | -0.85 |

**E.** $^{1}$HN, $^{15}$N, $^{13}$Cα, $^{13}$Cβ, and $^{13}$C' chemical shifts (ppm) for Rv1761c in 100 mM dodecylphosphocholine at 323K, pH 4.08.

| Residue | $^{1}$HN Shift | $^{15}$N Shift | $^{13}$Cα Shift | $^{13}$Cβ Shift | $^{13}$C' Shift |
|---|---|---|---|---|---|
| M1 | 8.24 | 116.01 | 56.21 | 32.41 | 177.58 |
| S2 | 8.05 | 114.34 | 59.61 | 63.58 | 175.59 |
| D3 | 8.29 | 119.82 | 53.96 | 39.26 | 176.14 |
| F4 | 8.02 | 119.38 | 58.10 | 39.55 | 175.48 |
| D5 | 8.12 | 121.13 | 53.30 | 40.02 | 177.14 |
| T6 | 8.47 | 116.23 | 65.22 | 68.46 | 177.52 |
| E7 | 8.57 | 122.71 | 58.96 | 28.14 | 178.58 |
| R8 | 8.12 | 119.09 | 58.83 | 30.35 | 179.80 |
| V9 | 8.02 | 118.15 | 65.87 | 31.50 | 177.75 |
| S10 | 8.29 | 115.79 | 61.88 | | |
| R11 | 8.30 | 120.66 | 59.54 | | 179.23 |
| A12 | 7.90 | 122.02 | 54.99 | 18.40 | 181.14 |
| V13 | 8.29 | 117.77 | 66.35 | 31.35 | 178.23 |
| A14 | 8.21 | 120.58 | 55.52 | 18.19 | 180.04 |
| A15 | 8.08 | 117.43 | 54.48 | 18.15 | 180.61 |
| A16 | 7.80 | 119.10 | 53.90 | 18.65 | 179.68 |
| L17 | 7.90 | 115.75 | 56.66 | 42.88 | 177.69 |
| V18 | 7.64 | 111.23 | 62.11 | 31.89 | 177.18 |
| G19 | 7.89 | 109.65 | 44.84 | | |
| P20 | | | 64.72 | 31.59 | 179.90 |
| G21 | 8.94 | 108.48 | 46.33 | | 176.83 |
| G22 | 8.19 | 109.86 | 47.21 | | 176.12 |
| V23 | 8.21 | 120.58 | 66.50 | 31.53 | 177.97 |
| A24 | 8.07 | 120.47 | 54.99 | 18.01 | 181.10 |
| L25 | 7.78 | 118.22 | 58.00 | 41.81 | 179.10 |
| V26 | 7.99 | 117.91 | 66.98 | 31.34 | 178.60 |
| V27 | 8.27 | 117.83 | 66.93 | 31.37 | 178.61 |
| K28 | 7.84 | 119.02 | 59.41 | 32.36 | 180.18 |
| V29 | 8.19 | 118.27 | 66.16 | 31.31 | 178.57 |
| F30 | 8.34 | 117.53 | 60.68 | 38.77 | 178.01 |
| A31 | 8.25 | 119.42 | 54.21 | 18.39 | 179.68 |
| G32 | 8.08 | 104.28 | 45.19 | | 175.68 |
| L33 | 7.88 | 121.87 | 56.76 | 40.82 | |
| P34 | | | 65.28 | 31.11 | 178.26 |
| G35 | 8.21 | 104.54 | 46.15 | | 176.04 |
| V36 | 7.93 | 117.56 | 63.42 | 32.40 | 176.96 |
| I37 | 7.77 | 117.48 | 61.68 | 38.14 | 176.33 |
| H38 | 8.24 | 119.08 | 55.06 | 28.81 | 175.13 |
| T39 | 8.14 | 116.51 | 60.68 | 69.19 | |
| P40 | | | 63.72 | 31.66 | 177.63 |
| A41 | 8.34 | 122.64 | 52.84 | 18.99 | 178.85 |
| R42 | 8.21 | 118.87 | 56.48 | 30.51 | 177.16 |
| R43 | 8.25 | 120.36 | 56.39 | 30.59 | 177.55 |
| G44 | 8.38 | 108.58 | 45.54 | | 174.97 |
| F45 | 8.24 | 119.97 | 58.95 | 39.70 | 176.28 |
| F46 | 8.24 | 118.10 | 57.84 | 39.23 | 176.34 |
| R47 | 8.27 | 120.75 | 56.16 | 30.70 | |
| S48 | 8.27 | 116.12 | 58.29 | | 174.64 |
| N49 | 8.50 | 120.87 | 51.57 | 38.88 | |
| P50 | | | 63.71 | 31.79 | 177.70 |

| Residue | $^{1}$HN Shift | $^{15}$N Shift | $^{13}$Cα Shift | $^{13}$Cβ Shift | $^{13}$C' Shift |
|---|---|---|---|---|---|
| E51 | 8.43 | 119.37 | 56.49 | 29.11 | |
| R52 | | | 56.18 | | 176.91 |
| I53 | 8.11 | 120.75 | 61.16 | 38.64 | 176.35 |
| Q54 | 8.46 | 124.80 | 55.27 | 29.43 | 176.57 |
| I55 | 8.33 | 121.61 | 61.76 | 38.55 | 177.25 |
| G56 | 8.41 | 110.75 | 45.75 | | 175.21 |
| D57 | 8.32 | 120.90 | 53.92 | 40.16 | 177.97 |
| W58 | 8.41 | 123.30 | 59.38 | 28.51 | 178.17 |
| R59 | 7.83 | 120.80 | 59.64 | 29.41 | |
| Y60 | 7.77 | 118.83 | 60.88 | 38.13 | 178.85 |
| E61 | 8.21 | 119.36 | 58.75 | 28.63 | 179.22 |
| V62 | 8.43 | 117.93 | 66.30 | 31.68 | 177.85 |
| A63 | 7.90 | 120.80 | 54.74 | 18.52 | 179.69 |
| H64 | 8.34 | 114.94 | 58.18 | 28.25 | 176.98 |
| D65 | 8.61 | 118.51 | 55.15 | 38.44 | 178.20 |
| G66 | 8.72 | 109.19 | 47.79 | | 175.70 |
| R67 | 8.21 | 120.58 | 59.66 | 29.51 | 179.25 |
| L68 | 7.78 | 119.78 | 57.58 | 41.69 | 180.22 |
| L69 | 8.15 | 119.60 | 57.97 | 41.75 | 179.70 |
| A70 | 8.06 | 118.59 | 56.02 | | 178.31 |
| A71 | 8.30 | 122.84 | 54.89 | | |
| H72 | | | | | |
| M73 | | | 58.07 | | |
| V74 | 8.11 | 117.14 | 65.40 | | 180.03 |
| N75 | 8.09 | 117.82 | 55.18 | 38.42 | 177.99 |
| G76 | 8.40 | 107.05 | 46.89 | | 175.35 |
| I77 | 8.11 | 120.75 | 64.04 | 37.90 | 177.62 |
| V78 | 8.03 | 120.64 | 65.68 | 31.65 | 178.43 |
| I79 | 7.93 | 120.05 | 64.06 | 37.62 | 178.71 |
| A80 | 8.20 | 122.66 | 54.73 | 18.47 | 180.36 |
| E81 | 8.04 | 114.17 | 57.40 | | |
| D82 | 8.27 | 118.55 | 58.07 | 39.41 | 179.62 |
| A83 | 8.40 | 120.59 | 55.20 | 18.25 | |
| L84 | 8.14 | 118.14 | 55.21 | | |
| I85 | 8.12 | 118.17 | 64.39 | | 178.63 |
| A86 | 8.14 | 121.68 | 54.70 | 18.33 | 180.27 |
| E87 | 8.08 | 115.09 | 57.10 | 28.52 | 177.36 |
| A88 | 8.08 | 120.18 | 53.60 | 18.99 | 178.89 |
| C89 | 8.04 | 113.24 | 63.37 | | 177.63 |
| G90 | 8.41 | 109.60 | 47.23 | | |
| P91 | | | 65.26 | | 179.60 |
| H92 | 8.18 | 115.28 | 57.79 | 28.23 | 177.86 |
| L93 | 8.45 | 121.20 | 57.91 | 41.68 | 178.92 |
| A94 | 8.61 | 120.46 | 55.76 | 18.06 | 180.63 |
| R95 | 7.98 | 117.02 | 59.09 | 29.86 | 179.84 |
| A96 | 8.06 | 122.08 | 55.07 | 18.27 | 180.86 |
| L97 | 8.57 | 117.30 | 57.81 | 41.55 | 180.01 |
| G98 | 8.53 | 106.15 | 47.64 | | 177.18 |
| Q99 | 8.11 | 120.75 | 58.80 | 28.44 | 179.58 |
| I100 | 8.09 | 120.10 | 65.05 | 38.24 | 178.88 |

| Residue | $^{1}$HN Shift | $^{15}$N Shift | $^{13}$Cα Shift | $^{13}$Cβ Shift | $^{13}$C' Shift |
|---|---|---|---|---|---|
| V101 | 8.49 | 118.78 | 66.79 | 31.21 | 179.14 |
| S102 | 8.21 | 114.62 | 61.65 | 62.97 | 177.04 |
| T103 | 7.88 | 119.66 | 58.27 | 30.39 | 178.88 |
| Y104 | 8.17 | 115.45 | 59.15 | 39.54 | 178.17 |
| G105 | 8.54 | 109.28 | 47.50 |  | 175.18 |
| A106 | 8.41 | 121.22 | 54.10 | 18.67 | 179.59 |
| T107 | 7.94 | 109.44 | 63.60 | 69.26 | 176.24 |
| V108 | 7.89 | 118.27 | 64.21 | 32.68 | 177.04 |
| I109 | 8.15 | 120.21 | 63.71 | 36.33 |  |
| P110 |  |  | 65.07 | 31.24 | 178.71 |
| N111 | 7.79 | 114.69 | 54.84 | 39.10 | 177.40 |
| I112 | 8.36 | 121.29 | 64.33 | 37.61 | 177.77 |
| N113 | 8.50 | 119.04 | 56.45 | 37.92 | 178.31 |
| A114 | 8.01 | 121.38 | 54.63 | 18.31 | 180.32 |
| A115 | 7.90 | 119.65 | 54.82 | 18.30 | 179.80 |
| I116 | 8.18 | 116.07 | 64.54 | 37.62 | 179.04 |
| E117 | 8.01 | 119.00 | 58.70 | 28.14 | 179.42 |
| V118 | 7.85 | 117.34 | 64.84 | 31.76 | 178.42 |
| L119 | 8.07 | 119.80 | 56.68 | 42.10 | 179.01 |
| G120 | 8.21 | 105.23 | 45.99 |  | 175.54 |
| T121 | 7.96 | 112.24 | 62.83 | 69.67 | 176.49 |
| G122 | 8.52 | 110.45 | 45.85 |  | 175.11 |
| T123 | 8.00 | 112.12 | 61.98 | 69.59 | 175.06 |
| D124 | 8.43 | 120.86 | 53.74 | 39.49 | 176.03 |
| Y125 | 8.05 | 119.90 | 58.21 | 38.51 | 175.69 |
| R126 | 7.94 | 121.20 | 55.45 | 30.89 | 175.84 |
| F127 | 7.93 | 122.10 | 57.85 | 39.79 |  |

**F.** N-H residual dipolar couplings (Hz) measured for Rv1761c weakly aligned using a 5.5% polyacrylamide gel ($RDC_{neutral}$) and a 5% polyacrylamide gel with 1% 2-acrylamido-2-methyl-1-propanesulfonic acid ($RDC_{amps}$).

| Residue | $RDC_{neutral}$ | $RDC_{amps}$ |
|---|---|---|
| M1 | | 0.24 |
| S2 | 0.61 | 0.53 |
| D3 | 0.54 | 0.4 |
| F4 | -0.14 | 0.32 |
| D5 | -1.53 | 0.42 |
| T6 | -2.78 | 0.37 |
| E7 | -1.19 | -0.11 |
| R8 | -0.44 | 1.13 |
| V9 | -3.02 | 0.76 |
| S10 | -3.7 | -0.88 |
| R11 | -0.69 | 0.41 |
| A12 | -1.57 | 1.88 |
| V13 | -3.14 | 0.16 |
| A14 | -1.85 | |
| A15 | 0.97 | 1.62 |
| A16 | -3.59 | 1.27 |
| L17 | -3.44 | -0.57 |
| V18 | 0.65 | 0.06 |
| G19 | 0.74 | 1.43 |
| P20 | | |
| G21 | 3.76 | -0.71 |
| G22 | 4.6 | -0.77 |
| V23 | 2.72 | |
| A24 | 8.01 | -1.71 |
| L25 | 5.78 | -4.03 |
| V26 | 5.76 | -0.9 |
| V27 | 2.17 | -1.21 |
| K28 | 7.84 | -2.66 |
| V29 | 7.17 | -2.03 |
| F30 | 5.14 | -1.03 |
| A31 | 5.03 | -0.71 |
| G32 | 6.21 | -3.08 |
| L33 | 8.17 | 0.74 |
| P34 | | |
| G35 | 4.15 | -0.07 |
| V36 | 5.84 | 1.1 |
| I37 | 5.48 | 2.11 |
| H38 | -1.91 | 1.16 |
| T39 | -3.67 | -1.01 |
| P40 | | |
| A41 | -2.54 | 1.38 |
| R42 | -3.95 | -0.14 |
| R43 | -5.31 | -0.59 |
| G44 | -2.7 | 0.74 |
| F45 | -4.87 | 0.74 |
| F46 | -10.61 | -1.15 |
| R47 | -8.48 | -0.46 |
| S48 | -5.8 | -1.21 |
| N49 | -7.28 | -1.56 |
| P50 | | |

| Residue | $RDC_{neutral}$ | $RDC_{amps}$ |
|---|---|---|
| E51 | -5.67 | -1.37 |
| R52 | | |
| I53 | -7.71 | |
| Q54 | -8.58 | -0.55 |
| I55 | -3.99 | -0.27 |
| G56 | 1.29 | 1.22 |
| D57 | -4.01 | -0.31 |
| W58 | -3.43 | 2.54 |
| R59 | 10.84 | 4.68 |
| Y60 | | |
| E61 | 0.71 | 3.44 |
| V62 | 2.17 | 3.04 |
| A63 | | 0.17 |
| H64 | 5.85 | 3.42 |
| D65 | 1.7 | 2.94 |
| G66 | 5.56 | 1.39 |
| R67 | | 0.47 |
| L68 | 4.7 | 7.49 |
| L69 | 5.76 | 4.44 |
| A70 | -1.31 | 0.95 |
| A71 | 2.65 | -3.07 |
| H72 | -1.79 | |
| M73 | | |
| V74 | 5.33 | -1.75 |
| N75 | 5.59 | -1.65 |
| G76 | 6.39 | 1.47 |
| I77 | 1.51 | |
| V78 | 2.88 | -1.05 |
| I79 | 6.17 | -1.66 |
| A80 | 1.88 | |
| E81 | 0.57 | 3.07 |
| D82 | 4.08 | 0.21 |
| A83 | 6.69 | 0.74 |
| L84 | -1.93 | -3.07 |
| I85 | -3.08 | |
| A86 | 2.51 | |
| E87 | 1.81 | -1.67 |
| A88 | | |
| C89 | -1.96 | -4.36 |
| G90 | 2.13 | 3.7 |
| P91 | | |
| H92 | | 3.28 |
| L93 | -3.93 | 3.53 |
| A94 | 0.96 | 5.83 |
| R95 | 2.45 | 6.19 |
| A96 | -2.42 | 3.37 |
| L97 | -1.14 | 4.56 |
| G98 | 2.44 | 6.18 |
| Q99 | | |
| I100 | -1.54 | 4.40 |

| Residue | $RDC_{neutral}$ | $RDC_{amps}$ |
|---|---|---|
| V101 | 1.15 | 5.63 |
| S102 | 3.13 | 5.89 |
| T103 | 0.2 | |
| Y104 | -1.56 | 4.41 |
| G105 | 3.48 | 0.42 |
| A106 | 1.24 | |
| T107 | 2.9 | 3.31 |
| V108 | 2.79 | -0.07 |
| I109 | -0.12 | -4.38 |
| P110 | | |
| N111 | 2.77 | -0.53 |
| I112 | 1.27 | -3.54 |
| N113 | -2.33 | -4.89 |
| A114 | 0.12 | -2.8 |

| Residue | $RDC_{neutral}$ | $RDC_{amps}$ |
|---|---|---|
| A115 | 2.01 | -1.46 |
| I116 | -1.32 | -4.37 |
| E117 | -2.95 | -4.2 |
| V118 | 1.49 | -1.17 |
| L119 | 0.59 | -3.01 |
| G120 | -3.27 | -3.9 |
| T121 | -0.27 | -0.88 |
| G122 | 0.16 | -0.1 |
| T123 | -1 | -0.76 |
| D124 | 0.58 | -0.32 |
| Y125 | -0.38 | -0.48 |
| R126 | -0.11 | -0.09 |
| F127 | -0.19 | -0.2 |

# Bibliography


Al-Hashimi,H.M., Valafar,H., Terrell,M., Zartler,E.R., Eidsness,M.K. and Prestegard,J.H. (2000) Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings. *J.Magn.Reson.,* **143**, 402-406.

Aller,S.G., Eng,E.T., De Feo,C.J. and Unger,V.M. (2004) Eukaryotic CTR copper uptake transporters require two faces of the third transmembrane domain for helix packing, oligomerization, and function. *J.Biol.Chem.,* **279**, 53435-53441.

Aller,S.G. and Unger,V.M. (2006) Projection structure of the human copper transporter CTR1 at 6-A resolution reveals a compact trimer with a novel channel-like architecture. *Proc.Natl.Acad.Sci.U.S.A.,* **103**, 3627-3632.

Almeida,F.C. and Opella,S.J. (1997) fd coat protein structure in membrane environments: structural dynamics of the loop between the hydrophobic trans-membrane helix and the amphipathic in-plane helix. *J.Mol.Biol.,* **270**, 481-495.

Arora,A., Abildgaard,F., Bushweller,J.H. and Tamm,L.K. (2001) Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. *Nat.Struct.Biol.,* **8**, 334-338.

Barrientos,L.G., Dolan,C. and Gronenborn,A.M. (2000) Characterization of surfactant liquid crystal phases suitable for molecular alignment and measurement of dipolar couplings. *J.Biomol.NMR,* **16**, 329-337.

Battiste,J.L. and Wagner,G. (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry,* **39**, 5355-5365.

Bax,A. (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics. *Protein Sci.,* **12**, 1-16.

Bax,A. and Grishaev,A. (2005) Weak alignment NMR: a hawk-eyed view of biomolecular structure. *Curr.Opin.Struct.Biol.,* **15**, 563-570.

Bax,A., Kontaxis,G. and Tjandra,N. (2001) Dipolar couplings in macromolecular structure determination. *Methods Enzymol.,* **339**, 127-174.

Bax,A. and Tjandra,N. (1997) High-resolution heteronuclear NMR of human ubiquitin in an aqueous liquid crystalline medium. *J.Biomol.NMR,* **10**, 289-292.

Berthold,D.A., Stenmark,P. and Nordlund,P. (2003) Screening for functional expression and overexpression of a family of diiron-containing interfacial membrane proteins using the univector recombination system. *Protein Sci.,* **12**, 124-134.

Bogusky,M.J., Leo,G.C. and Opella,S.J. (1988) Comparison of the dynamics of the membrane-bound form of fd coat protein in micelles and in bilayers by solution and solid-state nitrogen-15 nuclear magnetic resonance spectroscopy. *Proteins,* **4**, 123-130.

Bogusky,M.J., Schiksnis,R.A., Leo,G.C. and Opella,S.J. (1987) Protein backbone dynamics by solid state and solution 15N NMR spectroscopy. *J.Magn.Reson.*, **72**, 186-190.

Bour,S., Schubert,U. and Strebel,K. (1995) The human immunodeficiency virus type 1 Vpu protein specifically binds to the cytoplasmic domain of CD4: implications for the mechanism of degradation. *J.Virol.,* **69**, 1510-1520.

Bour,S. and Strebel,K. (2000) HIV accessory proteins: multifunctional components of a complex system. *Adv.Pharmacol.,* **48**, 75-120.

Brey,W.W., Edison,A.S., Nast,R.E., Rocca,J.R., Saha,S. and Withers,R.S. (2006) Design, construction, and validation of a 1-mm triple-resonance high-temperature-superconducting probe for NMR. *J.Magn.Reson.,* **179**, 290-293.

Bruschweiler,R. and Zhang,F. (2004) Covariance nuclear magnetic resonance spectroscopy. *J.Chem.Phys.,* **120**, 5253-5260.

Chamberlain,A.K., Faham,S., Yohannan,S. and Bowie,J.U. (2003) Construction of helix-bundle membrane proteins. *Adv.ProteinChem.*, **63**, 19–46.

Chou,J.J., Gaemers,S., Howder,B., Louis,J.M. and Bax,A. (2001) A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles. *J.Biomol.NMR,* **21**, 377-382.

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

Cierpicki,T. and Bushweller,J.H. (2004) Charged gels as orienting media for measurement of residual dipolar couplings in soluble and integral membrane proteins. *J.Am.Chem.Soc.,* **126**, 16259-16266.

Cierpicki,T., Liang,B., Tamm,L.K. and Bushweller,J.H. (2006) Increasing the accuracy of solution NMR structures of membrane proteins by application of

residual dipolar couplings. High-resolution structure of outer membrane protein A. *J.Am.Chem.Soc.,* **128**, 6947-6951.

Clore,G.M., Gronenborn,A.M. and Bax,A. (1998) A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *J.Magn.Reson.,* **133**, 216-221.

Clore,G.M., Starich,M.R. and Gronenborn,A.M. (1998) Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. *J.Am.Chem.Soc.*, **120**, 10571-10572.

Cohen,E.A., Terwilliger,E.F., Sodroski,J.G. and Haseltine,W.A. (1988) Identification of a protein encoded by the vpu gene of HIV-1. *Nature,* **334**, 532-534.

Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C., Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry,C.E.,3rd, Tekaia,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T., Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S., Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J., Moule,S., Murphy,L., Oliver,K., Osborne,J., Quail,M.A., Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,J., Squares,R., Squares,S., Sulston,J.E., Taylor,K., Whitehead,S. and Barrell,B.G. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature,* **393**, 537-544.

Cross,T.A. and Opella,S.J. (1980) Structural properties of fd coat protein in sodium dodecyl sulfate micelles. *Biochem.Biophys.Res.Commun.,* **92**, 478-484.

Crowell,K.J., Franzin,C.M., Koltay,A., Lee,S., Lucchese,A.M., Snyder,B.C. and Marassi,F.M. (2003) Expression and characterization of the FXYD ion transport regulators for NMR structural studies in lipid micelles and lipid bilayers. *Biochim.Biophys.Acta,* **1645**, 15-21.

Dancis,A., Yuan,D.S., Haile,D., Askwith,C., Eide,D., Moehle,C., Kaplan,J. and Klausner,R.D. (1994) Molecular characterization of a copper transport protein in S. cerevisiae: an unexpected role for copper in iron transport. *Cell,* **76**, 393-402.

De Angelis,A.A., Howell,S.C., Nevzorov,A.A. and Opella,S.J. (2006) Structure determination of a membrane protein with two trans-membrane helices in aligned phospholipid bicelles by solid-state NMR spectroscopy. *J.Am.Chem.Soc.,* **128**, 12256-12267.

De Angelis,A.A., Jones,D.H., Grant,C.V., Park,S.H., Mesleh,M.F. and Opella,S.J. (2005) NMR experiments on aligned samples of membrane proteins. *Methods Enzymol.,* **394**, 350-382.

De Angelis,A.A., Nevzorov,A.A., Park,S.H., Howell,S.C., Mrse,A.A. and Opella,S.J. (2004) High-resolution NMR spectroscopy of membrane proteins in aligned bicelles. *J.Am.Chem.Soc.,* **126**, 15340-15341.

De Feo,C.J., Aller,S.G. and Unger,V.M. (2007) A structural perspective on copper uptake in eukaryotes. *Biometals,* .

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

Ding,K. and Gronenborn,A.M. (2003) Sensitivity-enhanced 2D IPAP, TROSY-anti-TROSY, and E.COSY experiments: alternatives for measuring dipolar 15N-1HN couplings. *J.Magn.Reson.,* **163**, 208-214.

Donaldson,L.W., Skrynnikov,N.R., Choy,W.Y., Muhandiram,D.R., Sarkar,B., Forman-Kay,J.D. and Kay,L.E. (2001) Structural characterization of proteins with an attached ATCUN motif by paramagnetic relaxation enhancement NMR spectroscopy. *J.Am.Chem.Soc.,* **123**, 9843-9847.

Dutzler,R., Campbell,E.B. and MacKinnon,R. (2003) Gating the selectivity filter in ClC chloride channels. *Science,* **300**, 108-112.

Eisses,J.F., Chi,Y. and Kaplan,J.H. (2005) Stable plasma membrane levels of hCTR1 mediate cellular copper uptake. *J.Biol.Chem.,* **280**, 9635-9639.

Eisses,J.F. and Kaplan,J.H. (2005) The mechanism of copper uptake mediated by human CTR1: a mutational analysis. *J.Biol.Chem.,* **280**, 37159-37168.

Eisses,J.F. and Kaplan,J.H. (2002) Molecular characterization of hCTR1, the human copper uptake protein. *J.Biol.Chem.,* **277**, 29162-29171.

Faller,M., Niederweis,M. and Schulz,G.E. (2004) The structure of a mycobacterial outer-membrane channel. *Science,* **303**, 1189-1192.

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. and Kay,L.E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry,* **33**, 5984-6003.

Frankel,A.D. and Young,J.A. (1998) HIV-1: fifteen proteins and an RNA. *Annu.Rev.Biochem.,* **67**, 1-25.

Gardner,K.H. and Kay,L.E. (1998) The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. *Annu.Rev.Biophys.Biomol.Struct.,* **27**, 357-406.

Gardner,K.H., Rosen,M.K. and Kay,L.E. (1997) Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR. *Biochemistry,* **36**, 1389-1401.

Hansen,M.R., Mueller,L. and Pardi,A. (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat.Struct.Biol.,* **5**, 1065-1074.

Harris,E.D. (2000) Cellular copper transport and metabolism. *Annu.Rev.Nutr.,* **20**, 291-310.

Hoch,J.C. and Stern,A.S. (2001) Maximum entropy reconstruction, spectrum analysis and deconvolution in multidimensional nuclear magnetic resonance. *Methods Enzymol.,* **338**, 159-178.

Holmes,D.L. and Stellwagen,N.C. (1991) Estimation of polyacrylamide gel pore size from Ferguson plots of linear DNA fragments. II. Comparison of gels with different crosslinker concentrations, added agarose and added linear polyacrylamide. *Electrophoresis,* **12**, 612-619.

Holmes,D.L. and Stellwagen,N.C. (1991) Estimation of polyacrylamide gel pore size from Ferguson plots of normal and anomalously migrating DNA fragments. I. Gels containing 3% N,N'-methylenebisacrylamide. *Electrophoresis,* **12**, 253-263.

Howell,S.C., Mesleh,M.F. and Opella,S.J. (2005) NMR structure determination of a membrane protein with two transmembrane helices in micelles: MerF of the bacterial mercury detoxification system. *Biochemistry,* **44**, 5196-5206.

Hubbell,W.L., Gross,A., Langen,R. and Lietzow,M.A. (1998) Recent advances in site-directed spin labeling of proteins. *Curr.Opin.Struct.Biol.,* **8**, 649-656.

Huffman,D.L. and O'Halloran,T.V. (2001) Function, structure, and mechanism of intracellular copper trafficking proteins. *Annu.Rev.Biochem.,* **70**, 677-701.

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

Ishida,S., Lee,J., Thiele,D.J. and Herskowitz,I. (2002) Uptake of the anticancer drug cisplatin mediated by the copper transporter Ctr1 in yeast and mammals. *Proc.Natl.Acad.Sci.U.S.A.,* **99**, 14298-14302.

Ishii,Y., Markus,M.A. and Tycko,R. (2001) Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. *J.Biomol.NMR,* **21**, 141-151.

Iwahara,J., Anderson,D.E., Murphy,E.C. and Clore,G.M. (2003) EDTA-derivatized deoxythymidine as a tool for rapid determination of protein binding polarity to DNA by intermolecular paramagnetic relaxation enhancement. *J.Am.Chem.Soc.,* **125**, 6634-6635.

Iwahara,J., Schwieters,C.D. and Clore,G.M. (2004) Ensemble approach for NMR structure refinement against (1)H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *J.Am.Chem.Soc.,* **126**, 5879-5896.

Jiang,Y., Lee,A., Chen,J., Ruta,V., Cadene,M., Chait,B.T. and MacKinnon,R. (2003) X-ray structure of a voltage-dependent K+ channel. *Nature,* **423**, 33-41.

Jones,D.H. and Opella,S.J. (2006) Application of Maximum Entropy reconstruction to PISEMA spectra. *J.Magn.Reson.,* **179**, 105-113.

Jones,D.H. and Opella,S.J. (2004) Weak alignment of membrane proteins in stressed polyacrylamide gels. *J.Magn.Reson.,* **171**, 258-269.

Khademi,S., O'Connell,J.,3rd, Remis,J., Robles-Colmenares,Y., Miercke,L.J. and Stroud,R.M. (2004) Mechanism of ammonia transport by Amt/MEP/Rh: structure of AmtB at 1.35 A. *Science,* **305**, 1587-1594.

Kim,S. and Szyperski,T. (2003) GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information. *J.Am.Chem.Soc.,* **125**, 1385-1393.

Klomp,A.E., Juijn,J.A., van der Gun,L.T., van den Berg,I.E., Berger,R. and Klomp,L.W. (2003) The N-terminus of the human copper transporter 1 (hCTR1) is localized extracellularly, and interacts with itself. *Biochem.J.,* **370**, 881-889.

Knight,S.A., Labbe,S., Kwon,L.F., Kosman,D.J. and Thiele,D.J. (1996) A widespread transposable element masks expression of a yeast copper transport gene. *Genes Dev.,* **10**, 1917-1929.

Koenig,B.W., Hu,J.S., Ottiger,M., Bose,S., Hendler,R.W. and Bax,A. (1999) NMR measurement of dipolar couplings in proteins aligned by transient binding to purple membrane fragments. *J.Am.Chem.Soc.*, **121**, 1385-1386.

Koharudin,L.M., Bonvin,A.M., Kaptein,R. and Boelens,R. (2003) Use of very long-distance NOEs in a fully deuterated protein: an approach for rapid protein fold determination. *J.Magn.Reson.,* **163**, 228-235.

Korepanova,A., Gao,F.P., Hua,Y., Qin,H., Nakamoto,R.K. and Cross,T.A. (2005) Cloning and expression of multiple integral membrane proteins from Mycobacterium tuberculosis in Escherichia coli. *Protein Sci.,* **14**, 148-158.

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

Kovacs,H., Moskau,D. and Spraul,M. (2005) Cryogenically cooled probes - a leap in NMR technology. *Prog. NMR Spectrosc.*, **46**, 131–155.

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. and Girvin,M.E. (2004) An evaluation of detergents for NMR structural studies of membrane proteins. *J.Biomol.NMR,* **28**, 43-57.

Kuliopulos,A., Nelson,N.P., Yamada,M., Walsh,C.T., Furie,B., Furie,B.C. and Roth,D.A. (1994) Localization of the affinity peptide-substrate inactivator site on recombinant vitamin K-dependent carboxylase. *J.Biol.Chem.,* **269**, 21364-21370.

Kupce,E. and Freeman,R. (2004) Projection-reconstruction technique for speeding up multidimensional NMR spectroscopy. *J.Am.Chem.Soc.,* **126**, 6429-6440.

Langen,R., Oh,K.J., Cascio,D. and Hubbell,W.L. (2000) Crystal structures of spin labeled T4 lysozyme mutants: implications for the interpretation of EPR spectra in terms of structure. *Biochemistry,* **39**, 8396-8405.

Lee,J., Pena,M.M., Nose,Y. and Thiele,D.J. (2002) Biochemical characterization of the human copper transporter Ctr1. *J.Biol.Chem.,* **277**, 4380-4387.

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

Leo,G.C., Colnago,L.A., Valentine,K.G. and Opella,S.J. (1987) Dynamics of fd coat protein in lipid bilayers. *Biochemistry,* **26**, 854-862.

Li,Y., Logan,T.M., Edison,A.S. and Webb,A. (2003) Design of small volume HX and triple-resonance probes for improved limits of detection in protein NMR experiments. *J.Magn.Reson.,* **164**, 128-135.

Liang,B., Bushweller,J.H. and Tamm,L.K. (2006) Site-directed parallel spin-labeling and paramagnetic relaxation enhancement in structure determination of membrane proteins by solution NMR spectroscopy. *J.Am.Chem.Soc.,* **128**, 4389-4397.

Llanos,R.M. and Mercer,J.F. (2002) The molecular basis of copper homeostasis copper-related disorders. *DNA Cell Biol.,* **21**, 259-270.

Losonczi,J.A. and Prestegard,J.H. (1998) Improved dilute bicelle solutions for high-resolution NMR of biological macromolecules. *J.Biomol.NMR,* **12**, 447-451.

Luecke,H., Schobert,B., Richter,H.T., Cartailler,J.P. and Lanyi,J.K. (1999) Structure of bacteriorhodopsin at 1.55 A resolution. *J.Mol.Biol.,* **291**, 899-911.

Ma,C., Marassi,F.M., Jones,D.H., Straus,S.K., Bour,S., Strebel,K., Schubert,U., Oblatt-Montal,M., Montal,M. and Opella,S.J. (2002) Expression, purification, and activities of full-length and truncated versions of the integral membrane protein Vpu from HIV-1. *Protein Sci.,* **11**, 546-557.

Ma,C. and Opella,S.J. (2000) Lanthanide ions bind specifically to an added "EF-hand" and orient a membrane protein in micelles for solution NMR spectroscopy. *J.Magn.Reson.,* **146**, 381-384.

Mal,T.K., Matthews,S.J., Kovacs,H., Campbell,I.D. and Boyd,J. (1998) Some NMR experiments and a structure determination employing a [15N,2H] enriched protein. *J.Biomol.NMR,* **12**, 259-276.

Marassi,F.M., Ma,C., Gratkowski,H., Straus,S.K., Strebel,K., Oblatt-Montal,M., Montal,M. and Opella,S.J. (1999) Correlation of the structural and functional domains in the membrane protein Vpu from HIV-1. *Proc.Natl.Acad.Sci.U.S.A.,* **96**, 14336-14341.

Marassi,F.M. and Opella,S.J. (2003) Simultaneous assignment and structure determination of a membrane protein from NMR orientational restraints. *Protein Sci.,* **12**, 403-411.

Marassi,F.M. and Opella,S.J. (2000) A solid-state NMR index of helical membrane protein structure and topology. *J.Magn.Reson.,* **144**, 150-155.

Marassi,F.M., Ramamoorthy,A. and Opella,S.J. (1997) Complete resolution of the solid-state NMR spectrum of a uniformly 15N-labeled membrane protein in phospholipid bilayers. *Proc.Natl.Acad.Sci.U.S.A.,* **94**, 8551-8556.

Margottin,F., Bour,S.P., Durand,H., Selig,L., Benichou,S., Richard,V., Thomas,D., Strebel,K. and Benarous,R. (1998) A novel human WD protein, h-beta TrCp, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Mol.Cell,* **1**, 565-574.

McDonnell,P.A. and Opella,S.J. (1993) Effect of detergent concentration on multidimensional solution NMR spectra of membrane proteins in micelles. *J.Magn.Reson.B*, **102**, 120-125.

McDonnell,P.A., Shon,K., Kim,Y. and Opella,S.J. (1993) fd coat protein structure in membrane environments. *J.Mol.Biol.,* **233**, 447-463.

Meier,S., Haussinger,D. and Grzesiek,S. (2002) Charged acrylamide copolymer gels as media for weak alignment. *J.Biomol.NMR,* **24**, 351-356.

Mesleh,M.F., Lee,S., Veglia,G., Thiriot,D.S., Marassi,F.M. and Opella,S.J. (2003) Dipolar waves map the structure and topology of helices in membrane proteins. *J.Am.Chem.Soc.,* **125**, 8928-8935.

Mesleh,M.F. and Opella,S.J. (2003) Dipolar Waves as NMR maps of helices in proteins. *J.Magn.Reson.,* **163**, 288-299.

Mesleh,M.F., Veglia,G., DeSilva,T.M., Marassi,F.M. and Opella,S.J. (2002) Dipolar waves as NMR maps of protein structure. *J.Am.Chem.Soc.,* **124**, 4206-4207.

Miozzari,G.F. and Yanofsky,C. (1978) Translation of the leader region of the Escherichia coli tryptophan operon. *J.Bacteriol.,* **133**, 1457-1466.

Montal,M. (2003) Structure-function correlates of Vpu, a membrane protein of HIV-1. *FEBS Lett.,* **552**, 47-53.

Mori,S., Abeygunawardana,C., Johnson,M.O. and van Zijl,P.C. (1995) Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J.Magn.Reson.B,* **108**, 94-98.

Nevzorov,A.A., Mesleh,M.F. and Opella,S.J. (2004) Structure determination of aligned samples of membrane proteins by NMR spectroscopy. *Magn.Reson.Chem.,* **42**, 162-171.

Olson,D.L., Peck,T.L., Webb,A.G., Magin,R.L. and Sweedler,J.V. (1995) High resolution microcoil H-l-NMR for mass-limited, nanoliter-vol Samples. *Science*, **270**, 1967-1970.

Opella,S.J., Ma,C. and Marassi,F.M. (2001) Nuclear magnetic resonance of membrane-associated peptides and proteins. *Methods Enzymol.,* **339**, 285-313.

Opella,S.J., Park,S., Lee,S., Jones,D.H., Nevzerov,A., Mesleh,M.F., Mrse,A., Marassi,F.M., Oblatt-Montal,M., Montal,M., Strebel,K. and Bour,S. (2004) Structure and function of Vpu from HIV-1 *in Viral Membrane Proteins: Structure, Function, and Drug Design*, Fischer W. (Ed), Kluwer, Academic Publishers, New York.

Opella,S.J., Stewart,P.L. and Valentine,K.G. (1987) Protein structure by solid-state NMR spectroscopy. *Q.Rev.Biophys.,* **19**, 7-49.

Ottiger,M. and Bax,A. (1998) Characterization of magnetically oriented phospholipid micelles for measurement of dipolar couplings in macromolecules. *J.Biomol.NMR,* **12**, 361-372.

Ottiger,M., Delaglio,F. and Bax,A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J.Magn.Reson.,* **131**, 373-378.

Oxenoid,K. and Chou,J.J. (2005) The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proc.Natl.Acad.Sci.U.S.A.,* **102**, 10870-10875.

Oxenoid,K., Kim,H.J., Jacob,J., Sonnichsen,F.D. and Sanders,C.R. (2004) NMR assignments for a helical 40 kDa membrane protein. *J.Am.Chem.Soc.,* **126**, 5048-5049.

Page,R.C., Moore,J.D., Nguyen,H.B., Sharma,M., Chase,R., Gao,F.P., Mobley,C.K., Sanders,C.R., Ma,L., Sonnichsen,F.D., Lee,S., Howell,S.C., Opella,S.J. and Cross,T.A. (2006) Comprehensive evaluation of solution nuclear magnetic resonance spectroscopy sample preparation for helical integral membrane proteins. *J.Struct.Funct.Genomics,* **7**, 51-64.

Park,S.H., Prytulla,S., De Angelis,A.A., Brown,J.M., Kiefer,H. and Opella,S.J. (2006a) High-resolution NMR spectroscopy of a GPCR in aligned bicelles. *J.Am.Chem.Soc.,* **128**, 7402-7403.

Park,S.H., De Angelis,A.A., Nevzorov,A.A., Wu,C.H. and Opella,S.J. (2006b) Three-dimensional structure of the transmembrane domain of Vpu from HIV-1 in aligned phospholipid bicelles. *Biophys.J.,* **91**, 3032-3042.

Park,S.H. and Opella,S.J. (2005) Tilt angle of a trans-membrane helix is determined by hydrophobic mismatch. *J.Mol.Biol.,* **350**, 310-318.

Park,S.H., Mrse,A.A., Nevzorov,A.A., Mesleh,M.F., Oblatt-Montal,M., Montal,M. and Opella,S.J. (2003) Three-dimensional structure of the channel-forming trans-membrane domain of virus protein "u" (Vpu) from HIV-1. *J.Mol.Biol.,* **333**, 409-424.

Paul,M., Mazumder,S., Raja,N. and Jabbar,M.A. (1998) Mutational analysis of the human immunodeficiency virus type 1 Vpu transmembrane domain that promotes the enhanced release of virus-like particles from the plasma membrane of mammalian cells. *J.Virol.,* **72**, 1270-1279.

Permi,P., Rosevear,P.R. and Annila,A. (2000) A set of HNCO-based experiments for measurement of residual dipolar couplings in 15N, 13C, (2H)-labeled proteins. *J.Biomol.NMR,* **17**, 43-54.

Pervushin,K., Riek,R., Wider,G. and Wuthrich,K. (1997) Attenuated T2 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.U.S.A.,* **94**, 12366-12371.

Peti,W., Meiler,J., Bruschweiler,R. and Griesinger,C. (2002) Model-free analysis of protein backbone motion from residual dipolar couplings. *J.Am.Chem.Soc.,* **124**, 5822-5833.

Pintacuda,G., Kaikkonen,A. and Otting,G. (2004) Modulation of the distance dependence of paramagnetic relaxation enhancements by CSA x DSA cross-correlation. *J.Magn.Reson.,* **171**, 233-243.

Prestegard,J.H. (1998) New techniques in structural NMR--anisotropic interactions. *Nat.Struct.Biol.,* **5 Suppl**, 517-522.

Prestegard,J.H., al-Hashimi,H.M. and Tolman,J.R. (2000) NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Q.Rev.Biophys.,* **33**, 371-424.

Prosser,R.S., Losonczi,J.A. and Shiyanovskaya,I.V. (1998) Use of a novel aqueous liquid crystalline medium for high-resolution NMR of macromolecules in solution. *J.Am.Chem.Soc.*, **120**, 11010-11011.

Ramirez,B.E. and Bax.A. (1998) Modulation of the alignment tensor of macromolecules dissolved in a dilute liquid crystalline medium. *J.Am.Chem.Soc.*, **120**, 9106-9107.

Rosenzweig,A.C. (2002) Metallochaperones: bind and deliver. *Chem.Biol.*, **9**, 673-677.

Rovnyak,D., Frueh,D.P., Sastry,M., Sun,Z.Y., Stern,A.S., Hoch,J.C. and Wagner,G. (2004) Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction. *J.Magn.Reson.,* **170**, 15-21.

Ruckert,M. and Otting,G. (2000) Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J.Am.Chem.Soc.*, **122**, 7793-7797.

Safaei,R. and Howell,S.B. (2005) Copper transporters regulate the cellular pharmacology and sensitivity to Pt drugs. *Crit.Rev.Oncol.Hematol.,* **53**, 13-23.

Sanders,C.R.,2nd and Schwonek,J.P. (1992) Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR. *Biochemistry,* **31**, 8898-8905.

Sass,H.J., Musco,G., Stahl,S.J., Wingfield,P.T. and Grzesiek,S. (2000) Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. *J.Biomol.NMR,* **18**, 303-309.

Sass,J., Cordier,F., Hoffmann,A., Cousin,A., Omichinski,J.G., Lowen,H. and Grzesiek,S. (1999) Purple membrane induced alignment of biological macromolecules in the magnetic field. *J.Am.Chem.Soc.*, **121**, 2047-2055.

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

Schubert,U., Bour,S., Ferrer-Montiel,A.V., Montal,M., Maldarell,F. and Strebel,K. (1996a) The two biological activities of human immunodeficiency virus type 1 Vpu protein involve two separable structural domains. *J.Virol.,* **70**, 809-819.

Schubert,U., Ferrer-Montiel,A.V., Oblatt-Montal,M., Henklein,P., Strebel,K. and Montal,M. (1996b) Identification of an ion channel activity of the Vpu transmembrane domain and its involvement in the regulation of virus release from HIV-1-infected cells. *FEBS Lett.,* **398**, 12-18.

Schubert,U., Henklein,P., Boldyreff,B., Wingender,E., Strebel,K. and Porstmann,T. (1994) The human immunodeficiency virus type 1 encoded Vpu protein is phosphorylated by casein kinase-2 (CK-2) at positions Ser52 and Ser56 within a predicted alpha-helix-turn-alpha-helix-motif. *J.Mol.Biol.,* **236**, 16-25.

Schwarzinger,S., Kroon,G.J., Foss,T.R., Wright,P.E. and Dyson,H.J. (2000) Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView. *J.Biomol.NMR,* **18**, 43-48.

Schwieters,C.D., Kuszewski,J.J., Tjandra,N. and Clore,G.M. (2003) The Xplor-NIH NMR molecular structure determination package. *J.Magn.Reson.,* **160**, 65-73.

Shon,K.J., Kim,Y., Colnago,L.A. and Opella,S.J. (1991) NMR studies of the structure and dynamics of membrane-bound bacteriophage Pf1 coat protein. *Science,* **252**, 1303-1305.

Shon,K. and Opella,S.J. (1989) Detection of 1H homonuclear NOE between amide sites in proteins with 1H/15N heteronuclear correlation spectroscopy. *J.Magn.Reson.*, **82**, 193-197.

Skrynnikov,N.R., Goto,N.K., Yang,D., Choy,W.Y., Tolman,J.R., Mueller,G.A. and Kay,L.E. (2000) Orienting domains in proteins using dipolar couplings measured by

liquid-state NMR: differences in solution and crystal forms of maltodextrin binding protein loaded with beta-cyclodextrin. *J.Mol.Biol.,* **295**, 1265-1273.

Smith,D.B. and Johnson,K.S. (1988) Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene,* **67**, 31-40.

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

Sprangers,R. and Kay,L.E. (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature,* **445**, 618-622.

Staley,J.P. and Kim,P.S. (1994) Formation of a native-like subdomain in a partially folded intermediate of bovine pancreatic trypsin inhibitor. *Protein Sci.,* **3**, 1822-1832.

Strebel,K., Klimkait,T. and Martin,M.A. (1988) A novel gene of HIV-1, vpu, and its 16-kilodalton product. *Science,* **241**, 1221-1223.

Tamm,L.K., Abildgaard,F., Arora,A., Blad,H. and Bushweller,J.H. (2003) Structure, dynamics and function of the outer membrane protein A (OmpA) and influenza hemagglutinin fusion domain in detergent micelles by solution NMR. *FEBS Lett.,* **555**, 139-143.

Thai,K., Choi,J., Franzin,C.M. and Marassi,F.M. (2005) Bcl-XL as a fusion protein for the high-level expression of membrane-associated proteins. *Protein Sci.,* **14**, 948-955.

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

Tjandra,N., Omichinski,J.G., Gronenborn,A.M., Clore,G.M. and Bax,A. (1997) Use of dipolar 1H-15N and 1H-13C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat.Struct.Biol.,* **4**, 732-738.

Tobias,D.J., Gesell,J., Klein,M.L. and Opella,S.J. (1995) A simple protocol for identification of helical and mobile residues in membrane proteins. *J.Mol.Biol.,* **253**, 391-395.

Tolman,J.R., Flanagan,J.M., Kennedy,M.A. and Prestegard,J.H. (1995) Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. *Proc.Natl.Acad.Sci.U.S.A.,* **92**, 9279-9283.

Torres,J., Stevens,T.J. and Samso,M. (2003) Membrane proteins: the 'Wild West' of structural biology. *Trends Biochem.Sci.,* **28**, 137-144.

Tugarinov,V., Hwang,P.M. and Kay,L.E. (2004) Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu.Rev.Biochem.,* **73**, 107-146.

Tycko,R., Blanco,F.J. and Ishii,Y. (2000) Alignment of biopolymers in strained gels: a new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR. *J.Am.Chem.Soc.*, **122**, 9340-9341.

Ulmer,T.S., Ramirez,B.E., Delaglio,F. and Bax,A. (2003) Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy. *J.Am.Chem.Soc.,* **125**, 9179-9191.

Venters,R.A., Huang,C.C., Farmer,B.T.,2nd, Trolard,R., Spicer,L.D. and Fierke,C.A. (1995) High-level 2H/13C/15N labeling of proteins for NMR studies. *J.Biomol.NMR,* **5**, 339-344.

Vold,R.R., Prosser,R.S. and Deese,A.J. (1997) Isotropic solutions of phospholipid bicelles: a new membrane mimetic for high-resolution NMR studies of polypeptides. *J.Biomol.NMR,* **9**, 329-335.

Wallin,E. and von Heijne,G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.,* **7**, 1029-1038.

Wang,Y.X., Marquardt,J.L., Wingfield,P., Stahl,S.J., Lee-Huang,S., Torchia,D. and Bax,A. (1998) Simultaneous measurement of 1H-15N, 1H-13C' and 15N-13C' dipolar couplings in a perdeuterated 30 kDa protein dissolved in a dilute liquid crystalline phase. *J.Am.Chem.Soc.*, **120**, 7385-7386.

Wu,C.H., Ramamoorthy,A. and Opella,S.J. (1994) High-resolution heteronuclear dipolar solid-state NMR spectroscopy. *J.Magn.Reson.A*, **109**, 270-272.

Wüthrich,K. (1986) *NMR of proteins and nucleic acids.* Wiley, New York.

Yamashita,A., Singh,S.K., Kawate,T., Jin,Y. and Gouaux,E. (2005) Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature,* **437**, 215-223.

Yang,D., Tolman,J.R., Goto,N.K. and Kay,L.E. (1998) An HNCO-based pulse scheme for the measurement of 13Ca-1Ha one-bond dipolar couplings in 15N,13C labeled proteins. *J.Biomol.NMR,* **12**, 325-332.

Yu,X.G., Lichterfeld,M., Addo,M.M. and Altfeld,M. (2005) Regulatory and accessory HIV-1 proteins: potential targets for HIV-1 vaccines? *Curr.Med.Chem.,* **12**, 741-747.

Zeri,A.C., Mesleh,M.F., Nevzorov,A.A. and Opella,S.J. (2003) Structure of the coat protein in fd filamentous bacteriophage particles determined by solid-state NMR spectroscopy. *Proc.Natl.Acad.Sci.U.S.A.,* **100**, 6458-6463.

Zhou,B. and Gitschier,J. (1997) hCTR1: a human gene for copper uptake identified by complementation in yeast. *Proc.Natl.Acad.Sci.U.S.A.,* **94**, 7481-7486.

Zhou,Y., Morais-Cabral,J.H., Kaufman,A. and MacKinnon,R. (2001) Chemistry of ion coordination and hydration revealed by a K+ channel-Fab complex at 2.0 A resolution. *Nature,* **414**, 43-48.

Zweckstetter,M. and Bax,A. (2001) Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage. *J.Biomol.NMR,* **20**, 365-377.