Studying Transmembane Helix Interactions in SDS Micelles

by

Tabussom Qureshi

Thesis submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy

Department of Chemistry
Faculty of Science
University of Ottawa

Abstract

The importance of interactions between transmembrane domains of integral membrane proteins has been well-established in a range of essential cellular functions. Most integral membrane proteins also possess regions that lie on the exterior of the membrane that may influence the ability of these transmembrane domains to interact. We sought to test this hypothesis by quantifying the energetics of transmembrane helix self-association in the absence and presence of an amphipathic helix that can bind to the membrane surface. The model chosen for this study was the major coat protein (MCP) of M13 bacteriophage, which has an N-terminal amphipathic helix linked to its single transmembrane segment via a flexible linker. Dimerization of both full-length MCP and a peptide containing only the transmembrane domain (MCP_{TM}) was studied by solution NMR in SDS micelles. We found that there was an increase in the apparent dimerization affinity in the absence of the N-terminal helix. However, this increase in apparent affinity could be attributed to differences in detergent-binding properties of the two polypeptides in monomeric versus dimeric states when the empty micelle was considered to be a participant in the dimer dissociation.

Preliminary results from the integral membrane protein, p7 of the hepatitis C virus are also presented in this thesis. It has been demonstrated that p7 enhances viral infectivity and accumulation, and that this function may require oligomerization in the membrane. While we encountered limitations due to challenges in the generation of sufficient quantities of pure p7 samples, we were able to perform circular dichroism spectroscopy under conditions that may favor different oligomeric states. These studies suggest that there is a change in the degree of helicity upon oligomerization, and suggest

that SDS could be a suitable system to characterize the interactions of the p7 oligomer in the future.

Acknowledgements

I would first like to thank my supervisor, Dr. Natalie Goto, for all of her help and support throughout my graduate studies. For her guidance and expertise helping to navigate my project and understanding NMR, in particular. Also, for her time spent helping me prepare for presentations and write this thesis. I appreciate it all very much.

I would also like to say thank you to people who have assisted me with carrying out experiments. In particular, I would like to thank Dr. Simon Sauve and Dr. Yves Aubin from Health Canada for taking their time to run the three dimensional experiments required for backbone assignments. I would also like to thank Dr. Scaiano for allowing me access to his circular dichroism spectrometer.

Within the lab I would like to thank Jason, Alex and Saud for their help and support. Thank you to Aya Mansour for her help with the ^{15}N-Alanine diffusion exerpiments. I would also like to thank Nasrin, Daria and Ana for their help as friends and colleagues.

I would also like to take this opportunity to say thank you to friends who provided me with support throughout grad school and other important parts of my life through countless conversations and many happy memories along the way. Husna, Sara, Asma, Samira, Omar, Virja and Ambreen, you all provided me with constant motivation and it is something I will always appreciate.

And finally, I would like to express a very big thank you to my family. To my brother Danish, who has been there for me every step of the way, pushing me when it was difficult and always believing in me. My sister-in-law Marina for her support during

these past couple of years was very much appreciated. Thank you to my brother-in-law Rahan, his wife Jaline and kids Ryah and Zara. Thank you to my in-laws, Asif and Nasreen, who have shown me great understanding and care while I wrote my thesis. I would like to express my deepest thanks to my own parents, Saeed and Tasnim. My dad has been my biggest supporter through this, never doubting me no matter how I felt and helping me every step of the way. My mom has done more for me in every part of life than anyone and without her I know that I would not have achieved what I have. And finally, my husband Farhan who has been patient during our first year of marriage supporting me while I write this thesis. He has been accommodating and understanding, which is something I will always be grateful for.

Table of Contents

List of Abbreviations

3D	Three-dimensional
ATP	Adenosine Triphosphate
AUC	Analytical Ultracentrifugation
BCA	Bicinchoninic Acid
Bcl-2	B-Cell Lymphoma 2
BLM	Black Lipid Membranes
BNip3	Bcl-2-Nineteen-kDa Interacting Protein 3
CD	Circular Dichroism
CHO	Chinese Hamster Ovary
CMC	Critical Micelle Concentration
DAA	Direct Acting Antiviral Agents
DAGK	Diacyl Glycerol Kinase
DDM	Dodecylmaltoside
DHPC	1,2-dihexanoyl-sn-glycero-3-phosphocholine
DM	Decylmaltoside
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DNA	Deoxyribonucleic Acid
DOPC	Dioleoylphosphatidylcholine
DOPG	Dioleoylphosphatidylglycerol
DPC	Dodecylphosphocholine
EDTA	Ethylenediaminetetraacetic Acid
EGFR	Epidermal Growth Factor Receptor

ER	Endoplasmic Reticulum
ESR	Electron Spin Resonance
FGFR	Fibroblast Growth Factor Receptor
FRET	Förster Resonance Energy Transfer
GpA	Glycophorin A
GPMV	Giant Plasma Membrane Vesicles
GST	Glutathione-S-Transferase
HCV	Hepatitis C
HIS_6	Hexahistidine tag
HIV	Human Immunodeficiency Virus
HSQC	Heteronuclear Single Quantum Coherence
IFN	Interferon
IPTG	Isopropyl B-D-Thiogalactopyranoside
kDa	Kilodalton
MCP	Major Coat Protein
MWCO	Molecular Weight Cut Off
Ni-NTA	Nickel-Nitrilotriacetic acid
NMR	Nuclear Magnetic resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhuaser Effect Spectroscopy
NS	Non-Structural
PC	Palmitoyl-oleoyl Phosphatidylcholine

PCS	Pseudo Contact Shift
PDB	Protein Data Bank
PDC	Protein Detergent Complex
PE	Phosphatidylethanolamine
PEG	Polyethylene Glycol
PFG	Pulse Field Gradient
PG	Palmitoyleoyl Phophatidylglycerol
PPM	Parts Per Million
PS	Palmitoyl-oleoyl Phophatidylserine
RF	Radiofrequency
RNA	Ribonucleic Acid
RP-HPLC	Reverse-Phase High-Performance Liquid Chromatography
RTK	Receptor Tyrosine Kinase
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
TE	Tris-EDTA
TEM	Transmission Electron Microscopy
TFA	Trifluoroacetic Acid
TM	Transmembrane
TOCSY	Total Correlation Spectroscopy
TOXCAT	ToxR Transciption Activator, Chloramphenicol Acetyltransferase
Vpu	Viral Protein Unique

WT	Wild-Type
β-OG	Beta Octyl Glucoside

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Membrane Proteins

1.1.1 Biological importance of integral membrane proteins

Integral membrane proteins are responsible for many significant cellular functions that are vital to the survival of an organism. Some of these functions include the transportation of nutrients, electrolytes, signaling agents and toxins across the cellular membrane (Fromme & Grotjohann, 2008; Levchenko, 2003). Integral membrane proteins are also important in the communication of the cell with its extracellular environment. Alterations to the functions of membrane proteins can be detrimental to an organism and give rise to disease states, such as cystic fibrosis, Alzheimer's disease, and long QT-syndrome (Rowe et al., 2005; Sanders & Myers, 2004; Sanguinetti & Tristani-Firouzi, 2006; Suh & Checler, 2002), to name but a few. Along with being important for proper cellular function within an organism, membrane proteins also serve as targets for therapeutic treatments (Landry & Gies, 2008). This has lead to a wealth of biochemical, biophysical, and structural studies on membrane proteins that seek to gain understanding of how membrane proteins function, how disruptions give rise to disease states, and how membrane protein functions can be modulated.

1.1.2 Oligomerization of membrane proteins

Integral membrane proteins are defined as possessing at least one continuous stretch of hydrophobic amino acids that transverses the lipid membrane. The most common form of the transmembrane spanning region is α-helical, since all polar backbone atoms are engaged in hydrogen bonds, minimizing the energetic penalty of

being immersed in the hydrophobic environment of the lipid core (Helenius et al., 1979; Hemminga et al., 1992). Although a smaller subset of integral membrane proteins cross the membrane in a β-barrel configuration, this thesis is focused on membrane proteins that traverse the lipid bilayer with transmembrane (TM) α-helices.

Within the membrane, these TM regions can interact with TM domains from other subunits to form dimers and higher-order oligomers. These interactions are often necessary to carry out the biological functions of these proteins. In some cases, dimerization is required for function; for example, the dimeric state of the single-spanning BNip3 is required to carry out cellular apoptosis, and receptor tyrosine kinases undergo dimerization to transmit a signal across the membrane (Lemmon & Schlessinger, 2010; Li & Hristova, 2006; Ray et al., 2000; Sulistijo et al., 2003) Higher order oligomerization is also seen, for example in the Vpu1 protein, which forms an oligomeric ion channel that is necessary for the release of HIV particles (Marassi et al., 1999)

In some cases, aberrant oligomerization of proteins can lead to disruption of cellular function, resulting in disease, as seen in some mutations in the family of receptor tyrosine kinases (RTK) (Robertson et al., 2000). RTKs are cell surface receptors with a single TM helix that initiate downstream signaling cascades by dimerization that involves the TM helices (Bocharov et al., 2013; Fantl etl al., 1993; Lemmon & Schlessinger, 2010; Li & Hristova, 2006; Schlessinger, 1992; Schlessinger, 2000). For example, pathogenic mutations (G380R and A391E) within the TM domain of the fibroblast growth factor receptor 3 (FGFR3) are associated with disorders in skeletal development and cancer (Li & Hristova, 2006; Passos-Bueno et al., 1999). The exact effect of these mutations on the

pathology of the disease is not known, it is however proposed that these mutations alter the dimeric structure of the protein thereby altering its function (Bocharov et al., 2013).

Similarly toll-like receptors (TLR), which are a family of proteins that are involved in the function of innate and adaptive immune systems, carry out their functions by dimerization (Medzhitov, 2001). TLR proteins form homodimeric signaling complexes that interact with molecules associated with a pathogen or viral nucleic acids (Jin et al., 2007; Kang et al., 2009; Kawai & Akira, 2010; Leonard et al., 2008; Liu et al., 2009; B. S. Park et al., 2009). It has been shown that disregulation of TLRs can lead to autoimmune diseases, such as rheumatoid arthritis, with therapeutic techniques targeting TLRs being studied (Cook et al., 2004; Goh & Midwood, 2012; Krishnan et al., 2009; Montero Vega & de Andrés Martín, 2009; Neill et al., 2009). The structure of the dimer of the TM domain of TLR3 was solved, as well as a structure of a trimer. They were able to identify two possible surfaces for helix-helix interactions (Mineev et al., 2014) raising the possibility that interactions involving the TLR TM domains can play more than one role in TLR function.

In order to prevent the development of diseases linked to membrane protein oligomerization, and develop therapeutics to target them, it would be useful to obtain more detailed information on the structure and strength of these interactions. The study of these interactions, however, can be challenging due the complexity of the systems and their surroundings, which are discussed in this thesis.

1.1.2.1 Interactions of TM domains in lipid membranes

The oligomerization of membrane proteins is constrained by the membrane, which can be considered as a two dimensional system (Cymer & Schneider, 2012; Grasberger et al., 1986; Li et al., 2013). TM helix oligomerization is favored by the reduction of protein-lipid interactions, which are often energetically less favorable than protein-protein interactions. When membrane proteins are present in the membrane in a monomeric form, the annular layer of lipids surrounding the TM domain have restricted mobility, and as the protein self-associates, some of these lipid molecules are released, increasing the overall entropy of the entire system (Cymer & Schneider, 2012; Helms, 2002; Schneider et al., 2004). Also, oligomerization can shield regions of the TM domain from the lipid membrane, making it possible for segments containing polar or charged residues to be inserted into the membrane. Other lipid properties can also affect the structure and function of membrane proteins, including lateral pressure, lipid chemistry and thickness of the hydrophobic membrane core (Anbazhagan et al, 2010; Cymer et al., 2012; Marsh, 2008).

Side chains also make important contributions to interactions between transmembrane helices. These interactions may involve hydrogen bonding by polar residues along with van der Waals interactions between hydrophobic residues (Li et al., 2013). Both are seen in a very well-studied example of homodimerization by the single TM helix in the human erythrocyte sialoglycoprotein glycophorin A (GpA). High specificity in dimerization was demonstrated by mutagenesis studies carried out in sodium dodecyl sulfate (SDS) micelles (Lemmon et al., 1992). Structure determination by solid-state NMR in bilayers (Smith et al., 2001), and solution-state NMR in

dodecylphosphocholine (DPC) micelles (MacKenzie et al., 1997; Mineev et al., 2011), as well as biochemical studies of the dimer of the TM helix (Fleming et al., 1997; Lemmon et al., 1992; Lemmon et al., 1994; D. Schneider et al., 2004) showed that the helices associate in a right-hand coiled-coil through van der Waals interactions involving residues within the TM domain (Lys75, Val80 and Thr87), as well as by interactions in a GXXXG motif also in the TM domain (Fleming et al., 1997; Lemmon et al., 1992; Lemmon et al., 1994; MacKenzie et al., 1997; Schneider et al., 2004; Smith et al., 2001) (Figure 1.1). The interactions of the GXXXG motif have been attributed to the glycines that form a flat surface on one side of the helix that provides the potential for hydrogen bonding across the dimer interface between the C_α hydrogens and carbonyl oxygen atoms on the adjacent helix (Senes et al., 2001). Also the complementarity of residues flanking the glycines, have been shown to be important in the interaction of this domain, with hydrophobic residues being favoured at these sites (Russ & Engelman, 2000). This motif was first discovered in GpA and since has been found to be widespread in self-associating TM domains, although its impact on oligomerization in all of these systems is not established (Senes et al., 2000).

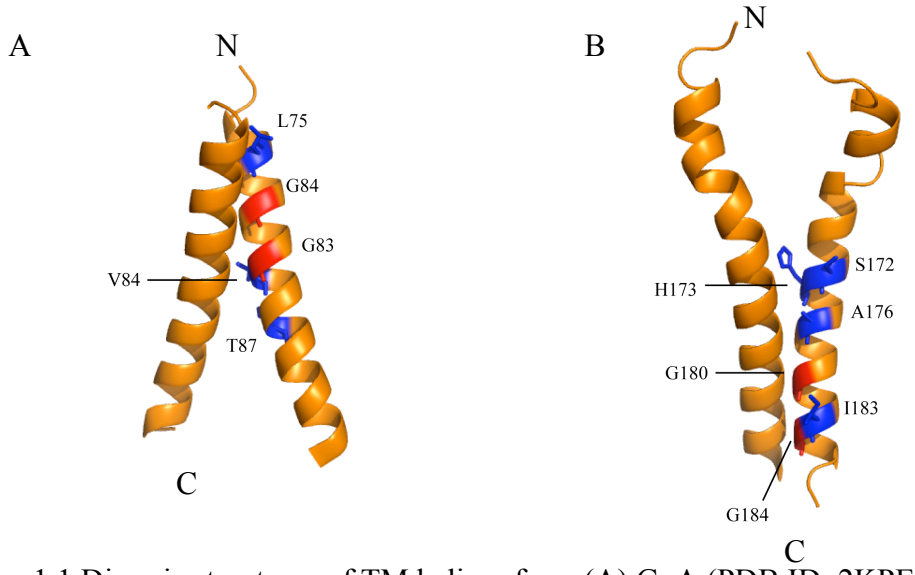

Figure 1.1 Dimeric structures of TM helices from (A) GpA (PDB ID: 2KPE) and (B) BNip3 (PDB ID: 2KA1). Residues participating in interactions of transmembrane domain are coloured and labeled on one subunit in each dimer. Red residues correspond to glycine residues that form the GXXXG motif in each dimer.

Another TM helix that associates via the GXXXG motif is found in BNip3, a member of the Bcl-2 family of apoptotic proteins that has been shown to play a role in programmed cell death by altering charge properties of the mitochondrial membrane (Chen et al., 1999; Shimizu & Tsujimoto, 2000). While their direct role is not completely understood, it has been well-established that the protein functions via the dimerization of the TM domain (Chen et al., 1997; Chen et al., 1999; Ohi et al., 1999; Ray et al., 2000; Yasuda et al., 1998). This interaction has been shown to be dependent on polar residues His173 and Ser172, the intersubunit interactions involving Ile183 and Ala176, and Gly180 and Gly184 that make up the GXXXG domain, which were identified by site-directed mutagenesis in SDS-PAGE and TOXCAT studies (Lawrie et al., 2010; Sulistijo et al., 2003; Sulistijo & MacKenzie, 2006). The structure of the TM dimer of BNip3 was solved by solution-state NMR in DPC micelles (Sulistijo & Mackenzie, 2009) and

DMPC/DHPC bicelles (Bocharov et al., 2007) (Figure 1.1). In these studies the interactions of the residues previously identified to be important for the interaction were corroborated by these structures. Based on the structure in lipid bicelles, they proposed that the dimer allows proton transfer across the membrane via a His-Ser node in the middle of the dimerization interface, which could be how the protein is altering the mitochondrial membrane potential that leads to cell apopotosis (Bocharov et al., 2007).

1.1.2.2 Impact of protein-lipid interactions on TM domain interactions

It is important to note that differences in lipid composition can affect the strength of interactions between TM helices. This is the result of differences in the physical properties of the membrane, such as membrane fluidity, a property that depends on both acyl chain and headgroup (Andersen & Koeppe, 2007) and on the packing of the acyl chains (Feigenson, 2009). Another important property that affects protein-protein interactions within the membrane is hydrophobic matching, where differences in the thickness of the bilayer and the hydrophobic region of the protein can lead to changes in function (Bloom et al., 1991; Mouritsen & Bloom, 1984).

The effects of different membrane properties on TM helix association was demonstrated for the GpA dimer by the Bowie group, who showed that PC membranes with negatively charged lipids such as palmitoyleoyl phophatidylglycerol (PG), or palmitoyloleoyl phophatidylserine (PS) destabilized the TM helix dimer, whereas PE lipids were stabilizing, likely by altering the phase of the membrane (Hong & Bowie, 2011). In separate studies, it was also found that the affinity of TM helix GpA dimerization was lower in bacterial and mammalian membranes, compared to the

interaction energies measured in membrane mimetic systems (Chen et al., 2010; Duong et al., 2007; Hong & Bowie, 2011).

1.1.3 Influence of Adjacent Amphipathic Domains on TM Helix Interactions

The majority of research on the oligomerization of membrane proteins has been focused on the elucidation of specific interactions between TM helices, with some work also being done to elucidate the effect of lipid contributions to these interactions. However, much less work has been done to study the impact of extramembranous sequences that interact with the membrane surface, and the potential effect they may have on the TM helix interactions. Nonetheless, many integral membrane proteins have amphipathic regions that lie at the C- or N-terminal side of a TM helix. Often these occur in the form of an amphipathic α-helix, where one side contains largely polar residues, and the other side being non-polar to create hydrophilic and hydrophobic faces. This distribution usually gives rise to a tendency for the amphipathic helix to interact with the surface of the membrane via its hydrophobic face (Figure 1.2) (Drin & Antonny, 2010; Segrest et al., 1990).

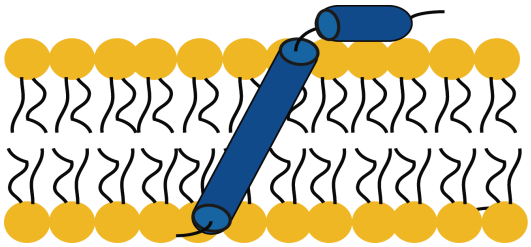

Figure 1.2 Schematic diagram of a single-spanning helical TM protein with a short amphipathic helix that lies on the surface of the lipid bilayer.

One physiologically important example of an integral membrane protein containing an amphipathic helix adjacent in sequence to a TM segment is phospholamban, a protein that regulates the contraction of cardiac muscles by maintaining calcium homeostasis in the sarcoplasmic reticulum (MacLennan and Kranias, 2003). This is carried out by the monomeric form of phospholamban that preferentially interacts with the Ca-ATPase transporter via the N-terminal amphipathic helix that must dissociate from the membrane (Reddy et al., 1999; Zamoon et al., 2008). The structure of monomeric phospholamban was solved by solution-state NMR in dodecylphosphocholine (DPC) micelles, and contains a single TM helix with an N-terminal amphipathic helix, connected by a flexible linker which had ambiguous NOEs and contributed to the high RMSD of the structure (Zamoon et al., 2008). The monomeric form of phospholamban has been found to be in equilibrium with a pentameric oligomer in the cell membrane (Hou et al., 2008). While it has been shown that the pentameric form of phospholamban can interact with the Ca-ATPase, this interaction leads to the disruption of the oligomer (Oxenoid & Chou, 2005; Reddy et al., 1999; Verardi et al., 2011). The interactions that hold the pentamer together are located within the TM domain, and are anchored by leucine zipper interactions, which is a common coiled-coil

motif found in helical protein dimers. This interaction was proposed by mutagenesis studies (Karim et al., 1998; Simmerman et al., 1996) and confirmed by solution-NMR studies on the structure of pentameric phospholamban in DPC micelles (Oxenoid & Chou, 2005).

Phosphorylation of the phospholamban amphipathic helix has been found to inhibit its ability to bind to Ca-ATPase (Hou et al., 2008). Phosphorylation also appears to stabilize the pentameric from, as shown by the effect of phosphomimetic mutations in the N-terminal helix that shifted the equilibrium towards the pentameric state (Hou et al., 2008) (Figure 1.3A). This may be due to changes in the interaction between the amphipathic helix and the surface of the lipid bilayer caused by phosphorylation, raising the possibility that these interactions could indirectly modulate the interactions of the TM domain (Hou et al., 2008).

Another well-characterized example of a self-associating TM helix protein with amphipathic helices that may modulate activity is the Vpu1 protein of HIV (Ma et al., 2002; Park et al., 2009) (Figure 1.3B). Vpu1 forms an ion channel in the cellular membrane that aids in the release of virus particles from the cell, and helps promote the intracellular degradation of the CD4 receptor (Ma et al., 2002). The C-terminal amphipathic helix has two highly conserved sites for phosphorylation (Ser52 and Ser56) that are essential for interactions with CD4 and its subsequent degradation. In the absence of the C-terminal amphipathic helix, dimerization and higher order oligomer formation properties are preserved (Marassi et al., 1999), however, there is a decrease in the probability of observing an open state of the channel, as demonstrated in channel conductance measurements (Ma et al., 2002). This suggests that the C-terminal

amphipathic helices play a role in modulating, or promoting, the interaction of the TM domain of Vpu1 and may stabilize the conductive state of the ion channel. Also the C-terminal amphiphathic helix in the monomeric form of Vpu has been shown by solid-state NMR to lie parallel to the membrane surface in oriented vesicles (Marassi et al., 1999), suggesting that this membrane interaction could play a role in modulating TM helix interactions.

More recently, the influence of amphipathic helices on TM helix interactions was demonstrated for the carboxyl-terminal domain of the amyloid precursor protein called C99, a protein that is responsible for the release of amyloid-β peptides that are associated with Alzheimer's disease (Kukar et al., 2011; Takami et al., 2009). C99 is made up of a single TM helix, and a small linker region joined to a short N-terminal amphipathic helix that interacts with the surface of the membrane (Barrett et al., 2012) (Figure 1.3C). In the absence of the amphipathic helix, the C99 TM helix undergoes weak dimerization that is essentially blocked by the amphipathic helix in the full-length construct (Barrett et al., 2012). This indicates that the amphipathic helix is required to prevent the oligomerization of C99.

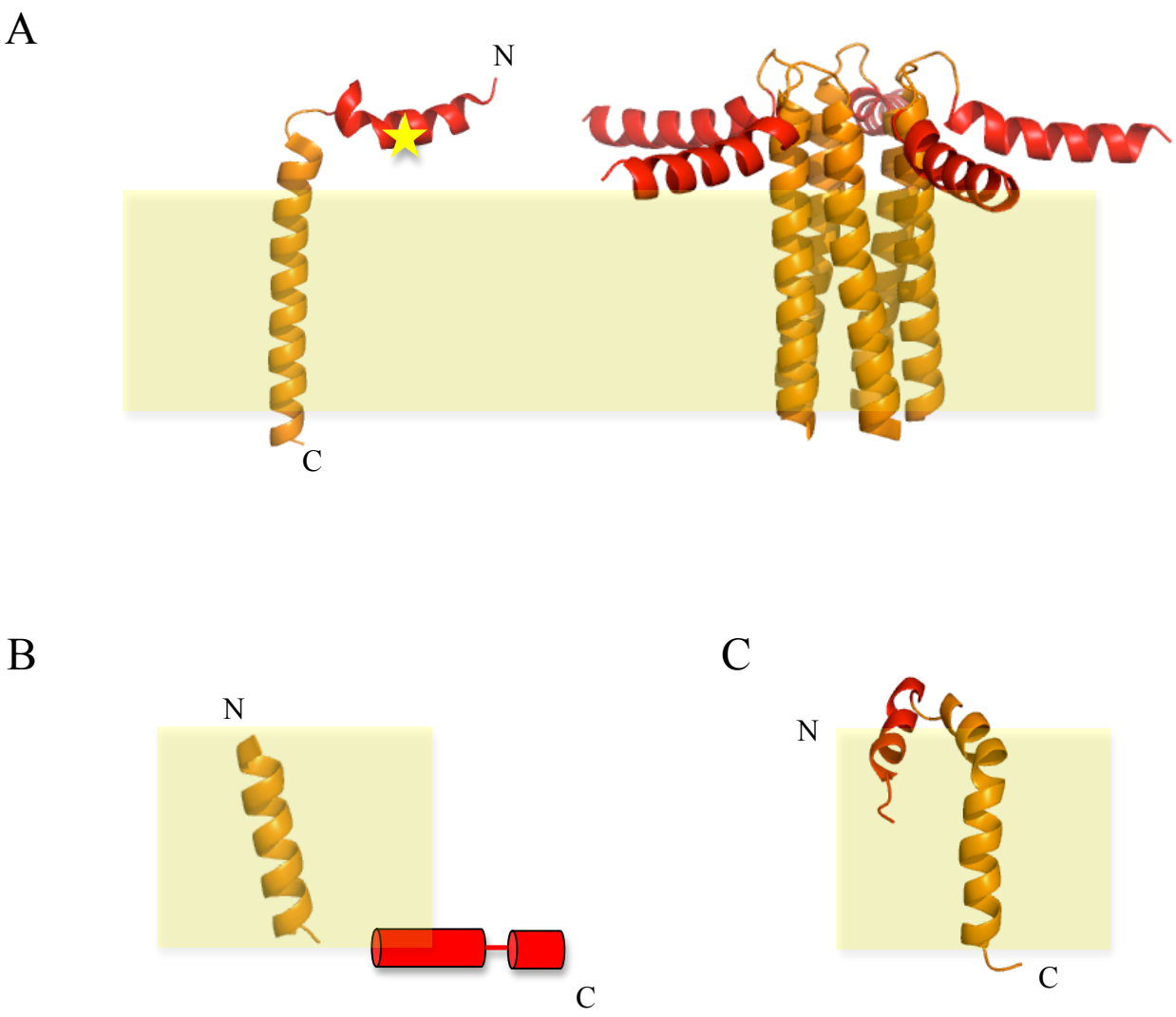

Figure 1.3 Ribbon diagram representation of membrane proteins containing a single TM helix with N-terminal or C-terminal amphipathic domains (shown in red) that alter the oligomeric properties of the protein, giving rise to changes in function. The approximate location of the lipid bilayer is indicated by the yellow shading. (A) Structures of phospholamban in monomeric (PDB ID: 1N7L) and oligomeric (PDB ID: 2KYV) forms. Point mutations targeted to the amphipathic helix at residues Ser16 and Thr17 (shown by star) have been shown to alter the monomer-oligomer equilibrium, even though these helices do not participate in direct interactions with each other in the pentamer. (B) Structure of the monomeric form of the TM domain of Vpu1 (PDB ID: 1PJE), with a schematic depiction of a C-terminal amphipathic helix that has been shown in truncation studies to affect its ion conductivity. (C) Monomeric structure of amyloid precursor protein C99 (PDB ID: 2LP1) with an N-terminal amphipathic helix that inhibits its dimerization.

While there has been a great deal of work undertaken to elucidate the direct interactions occurring between TM helices, these examples indicate that amphipathic helices can also play a role in modulating TM helix-helix interactions. Although interactions between the amphipathic helix and the membrane surface may be responsible for these effects, there is little research into the factors that give rise to this modulation. For this reason, this thesis focuses on evaluating the effect of an adjacent amphipathic helix on a TM helix interaction. I have chosen to use a simple and well-characterized model system for these studies, since this should facilitate a systematic approach to this investigation. The system that was used for these studies is the major coat protein (MCP) of M13 bacteriophage.

1.2 The M13 bacteriophage major coat protein as a model system

MCP has long been considered to be a model membrane protein system (Almeida & Opella, 1997; Marassi et al., 1997; Nazarov et al., 2006; Nazarov et al., 2007; Stopar et al., 2003; Vos et al., 2000). It is a relatively small protein, at 50 amino acid residues (6 kDa), with a simple structure consisting of an N-terminal amphipathic helix and a single α-helical TM segment (Figure 1.4).

Figure 1.4 Primary sequence of full-length MCP of M13 bacteriophage. The TM domain and N-terminal amphipathic helix are labeled in blue and green, respectively.

1.2.1 M13 bacteriophage replication cycle

M13 bacteriophage belongs to a family of filamentous phage (inovirus) that attack Gram-negative bacterial cells (Marvin, 1998). The M13 virion is rod-like in shape, with a circular single-stranded genome (Glucksmant & Bhattacharjee, 1992). Its DNA is encapsulated by a viral sheath that is made up of ~2700 copies of MCP (pVII) and a few copies of minor coat proteins (i.e. five copies of pVII and pIX each at one end, and five copies of pIII and PVI at the opposite end) (Glucksmant & Bhattacharjee, 1992; Kehoe & Kay, 2005; Marvin, 1998; Marvin et al., 1994).

The replication cycle of M13 is well characterized, and the role of MCP during this cycle has been extensively documented (Figure 1.5). The viral life cycle begins with infection of the bacterial cell by M13 bacteriophage via interactions on an F-pilus. After binding of the viral particle to the bacterial surface, the phage then transfers its DNA into the host cell, with the phage disassembling and the coat protein inserting into the bacterial membrane (Nakamura et al., 2003). Once inside the cell, the circular DNA forms a complex with minor protein pV, which enables DNA replication and packaging of progeny DNA into its circular form for incorporation into assembled phage. MCP is synthesized using host cell machinery, and inserted into the cellular membrane with the aid of a bacterial leader peptidase that cleaves MCP following its insertion into the membrane, where it is stored as a dimer until phage assembly (Marvin, 1998). Assembly occurs in a non-lytic manner, with incorporation of MCP into the bacteriophage coat as the virus particle is being extruded from the cell (Feng et al., 1997; Guan et al., 1995; Konings et al., 1995; Marciano et al., 1999; Marvin, 1998; Opalka et al., 2003).

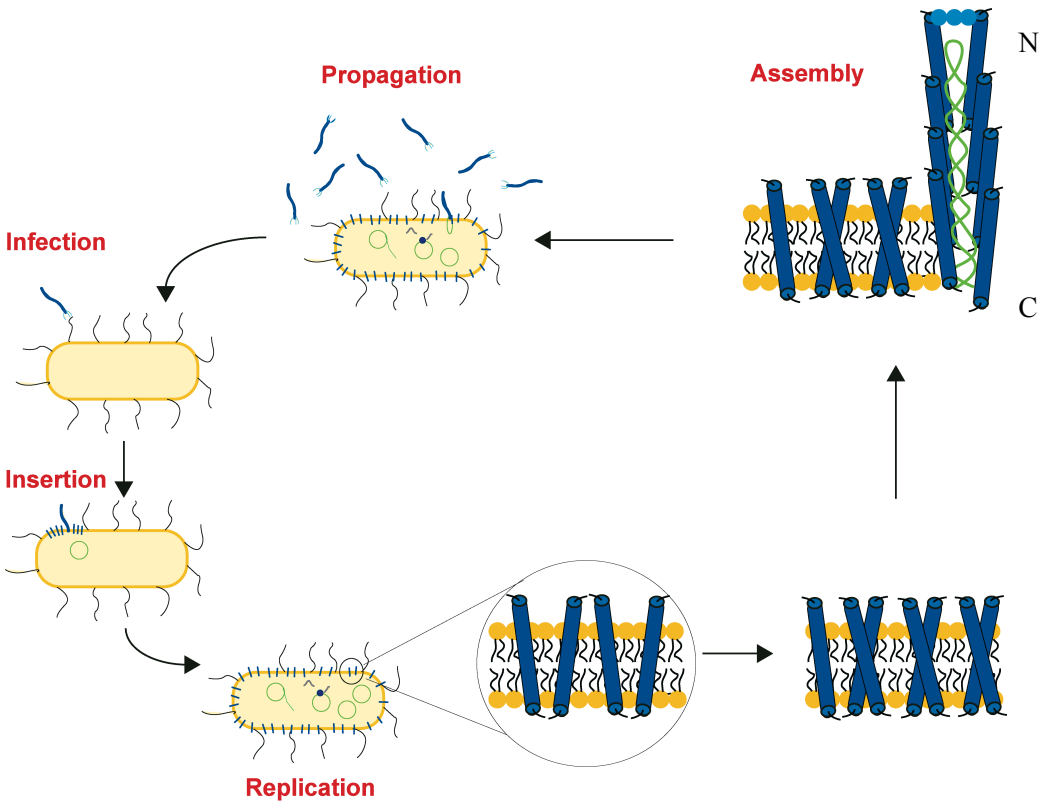

Figure 1.5 Schematic diagram of the viral life cycle of M13 bacteriophage in *E. coli* cells. Bacteriophage particles (blue) infect bacteria (yellow) via interactions with an F-pilus (grey), enabling transfer of genetic material to host. Utilizing host cell machinery, replication of ssDNA (green), synthesis of MCP (blue rods) and minor proteins are carried out. During replication of MCP, the protein is stored in the cellular membrane, with at least some part of the population forming a dimer. During assembly of the phage, MCP is incorporated into the viral coat with a staggered arrangement of subunits around the single-stranded phage genome. The assembled viral particle is extruded non-lytically from the bacterial cell.

1.2.2 MCP oligomerization

X-ray fiber diffraction and cryoelectron microscopy experiments conducted on the M13 bacteriophage have given rise to a model of the phage-bound MCP where the N-terminal amphipathic helix is a linear extension of the α-helical TM helix (Glucksmant & Bhattacharjee, 1992; Wang et al., 2006) (Figure 1.6). The structure of MCP within the bacteriophage was subsequently determined by solid-state NMR spectroscopy, providing

confirmation that MCP forms an almost perfectly continuous helix, a structure that is sometimes referred to in the literature as the "I" conformation of MCP (Zeri et al., 2003). In this state, the C-terminus of MCP interacts with phage DNA, and the N-terminus is exposed on the virus surface, with the helical axis angled away from the centre of the phage (Glucksmant & Bhattacharjee, 1992). There is a staggered arrangement of the MCP subunits, as determined by both x-ray diffraction and solid-state NMR. Mutagenesis studies identified residues Ile39, Leu41, Phe42, and Phe45 as residues that are critical for maintaining strong intermolecular interactions to form the bacteriophage coat (Glucksmant & Bhattacharjee, 1992; D. Marvin & Wachtel, 1976; Nagler et al., 2002; Williams et al., 1995).

During phage replication, the viral particle is disassembled, and the major coat protein is integrated into the bacterial inner membrane, along with newly synthesized coat protein. It has been shown by cross-linking studies that MCP in this membrane-integrated state undergoes oligomerization, with interactions that differ from those found in the viral particle (Haigh & Webster, 1998; Nagler et al., 2007), and will be discussed in further detail in Section 1.25. During extrusion of phage from the bacterial cell, this dimer is disrupted, and the MCP molecules are assembled around phage genetic material.

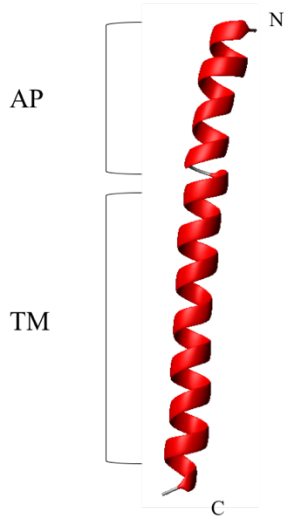

Figure 1.6 Structure of MCP within a bacteriophage particle, solved by solid-state NMR. The transmembrane helix (TM) and amphipathic helix (AP) are indicated.

1.2.3 Structure of the monomeric state of MCP

1.2.3.1 Monomeric MCP: "L" conformation

A structure of monomeric MCP solubilized in SDS micelles was solved by the Van de Ven group in 1998, and showed that the protein consists of two α-helices, the first running from residues 8-16 (the N-terminal amphipathic helix) and the second from residues 25-45 (the hydrophobic transmembrane helix), with the two helices being connected by a flexible hinge region (Papavoine et al., 1998; van de Ven et al., 1993). In this structure the N-terminal amphipathic helix shows considerable motional freedom on the ns-ps timesecale (Papavoine et al., 1998; Papavoine et al., 1994; Spruijt & Meijer, 2000; Vos et al., 2009). NMR experiments of MCP in SDS micelles doped with spin-labeled stearate were able to show that the N-terminal amphipathic helix interacts with the SDS micelle surface (Papavoine et al., 1994) (Figure 1.7). Similar results were also obtained on MCP in dodecylphosphocholine (DPC) micelles (Papavoine et al., 1995).

This secondary structure is also in line with results on MCP from closely related filamentous bacteriophages Pf1 and fd (McDonnell et al., 1993; Shon et al., 1991; Vos et al., 2009). The secondary structure of Pf1 MCP was determined by solution-state NMR and solid-state NMR, and it was found that the TM helix and the N-terminal amphipathic helix were in a perpendicular orientation with respect to each other, with the N-terminal amphipathic helix lying on the membrane surface (Shon et al., 1991). Similarly, MCP from fd phage has been shown to have the same topology, as determined by solution-state NMR in SDS micelles (Almeida & Opella, 1997).

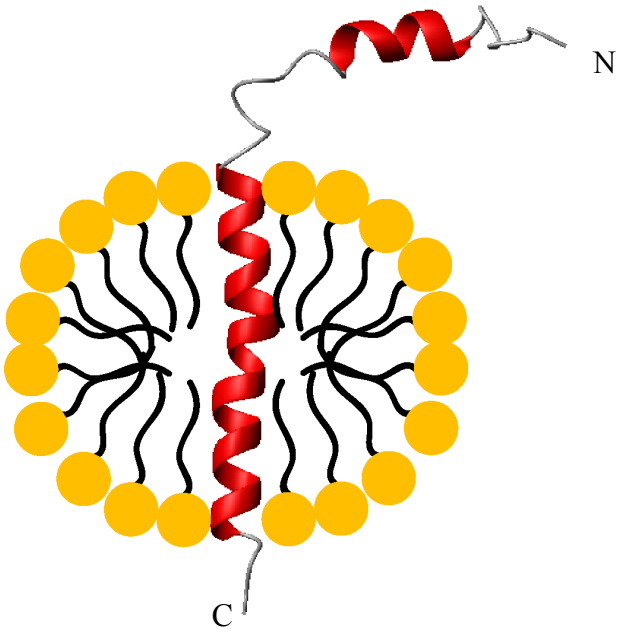

Figure 1.7 Lowest energy structure of monomeric MCP determined by solution state NMR in SDS micelles. The micelle surrounding the protein is illustrated schematically.

1.2.3.2 Monomeric MCP: "I" conformation

An alternate state that has been proposed in the literature is the "I" conformation, with the N-terminal amphipathic extending from the TM helix (Figure 1.6). A fluorescence experiment that was carried out on MCP in DOPC:DOPG bilayers that suggested a continuous helix structure (Koehorst et al., 2004). In this study they developed a new formalism using fluorescence and studied the orientation and tilt angles of MCP within phospholipid bilayers from which they were able to extract the orientation angle, tilt angle and the relative position of the transmembrane part of MCP. Using Stokes shift analysis of AEDEAN fluorescent-labeled cysteine mutants of MCP, and the assumption of a rigid α-helix, it was found that MCP is incorporated into the bilayer in a tilted orientation with respect to the bilayer normal, with the N-terminal helix forming an extension of the TM helix, with a kink in between (Koehorst et al., 2004).

In a follow up study done by the same group, the structure of MCP was examined by FRET (Nazarov et al., 2007), giving rise to similar estimates for TM helix tilt (18° with respect to the normal to the membrane). Also, residue 20 was found in the interfacial region between acyl chains and phopshpolipid head group phases, and introduced a distinct kink in the structure. However, the residues that made up the N-terminal helix were unresolved in the 100 structures calculated from FRET data, potentially reflecting mobility in this region. This indicates that while the full-length of the MCP structure may be mostly helical, the N-terminal portion is not always an extension of the TM domain and may spend some time bound to the exterior of the membrane.

It has been proposed that the difference between the L and I conformations is caused by differences in the membrane system being used (Vos et al., 2009). For example

the impact of micelle curvature is thought to induce strain on the embedded MCP molecule. However, oriented bilayers may force the protein into a bent conformation due to incomplete incorporation (Nazarov et al., 2007; Vos et al., 2009). Despite the differences in these environments, mobility in the N-terminal helix has been demonstrated in most membrane mimetic systems indicating that there is a possibility that it is interacting with the exterior of the membrane.

1.2.4 MCP dimerization

1.2.4.1 Evidence for dimerization

The dimerization of the major coat protein of M13 bacteriophage has been described as a weak association via a well-characterized GXXXG motif and surrounding residues (Lemmon et al., 1994; Marvin, 1998; Senes et al., 2000; Wu et al., 2007). This interface has been identified through *in vivo* TOXCAT assays and changes in SDS-PAGE gel migration rates for wild-type (WT) and mutant MCP sequences. In these studies, alterations to residues neighboring the ^{34}GXXXG38 domain, specifically, V29, V30, V31 and I32, were shown to significantly alter the affinity of MCP dimers (Dawson et al., 2003; Deber et al., 1993; Johnson et al., 2006; Johnson et al., 2006; Melnyk et al., 2004; Wang & Deber, 2000).

Some of the earliest solution NMR studies on MCP showed evidence that a homodimeric state could be obtained in SDS micelles (Henry & Sykes, 1992). Backbone chemical shift assignments on an ^{15}N-labeled sample allowed the identification of a subset of peaks with distinct chemical shifts that was proposed to correspond to a dimeric species. A high resolution structure of this dimer system has, however, not yet been determined, presumably due to difficulties in isolating MCP in a purely dimeric state.

1.2.4.2 Model of the transmembrane dimer of MCP

While there is currently no structure for the MCP dimer, a study on the TM segment of MCP was previously carried out that provided a model for the association of the TM domain within SDS micelles (Wu et al., 2007). In that study, a peptide was synthesized that spans residues 21-48 of MCP ($MCP_{TM\text{-}M28L/V31L}$), which primarily contains its TM helix. The peptide was selectively ^{15}N-labeled at 12 sites that spanned the length of the dimerization interface. Mutations were introduced at residues 28 and 31, where each leucine was substituted for Met28 and Val31, to increase affinity for dimerization as had previously been observed (Melnyk et al., 2002; Wu et al., 2007) (Figure 1.8A).

A

$MCP_{TM\text{-}M28L/V31L}$ KKKYIGYAWALVVLIVGATIGIKLFKKFTSK

B C

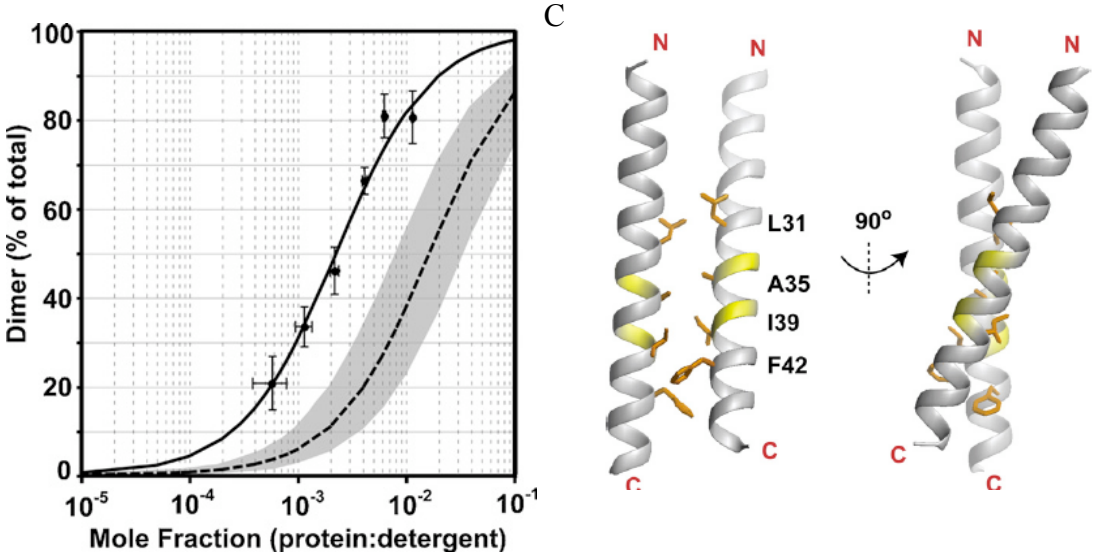

Figure 1.8 Structure and interaction of $MCP_{TM\text{-}M28L/V31L}$ (A) MCP_{TM} sequence, with sites that were ^{15}N-labeled shown in red, and mutated residues are underlined. (B) $MCP_{TM\text{-}M28L/V31L}$ association in SDS micelles. The fractions of the population (diamond) that were dimeric at different molar ratios of protein:detergent were determined by solution NMR. The data was then fit to a distribution (solid black curve) that describes a monomer-dimer equilibrium with a standard free energy of association of -3.8 kcal/mol. The dashed curve shows the distribution predicted for random distribution of $MCP_{TM\text{-}M28L/V31L}$ segments in micelles of SDS with 60 detergent molecules per micelle. The shaded region surrounding it depicts the range of distributions if the number of detergent molecules were increased or decreased by 50%. (C) Model of the $MCP_{TM\text{-}M28L/V31L}$ dimer in β-OG micelles determined by rigid body docking using ambiguously defined restraints from chemical shift differences between monomer and dimer, and PRE-based distance restraints from MTSL-labeled cysteine mutants (Wu et al., 2007).

Dimerization of the MCP_{TM} peptide was studied by solution NMR in SDS and β-octyl glucoside (β-OG) micelles (Wu et al., 2007). In SDS micelles ^{15}N-HSQC spectra showed the presence of two distinct sets of peaks for the ^{15}N-labelled residues. The intensity of these peaks was dependent on the molar ratio of detergent to protein, and as the ratio increased one set of peaks became less intense, while the other set increased in

intensity, as expected for an equilibrium system between monomeric and dimeric MCP_{TM}. The affinity of dimerization could be quantified by using peak intensities to measure the population of momomer and dimer species over a range of protein:detergent molar ratios (Figure 1.8B). These populations were found to favor the dimer over what would be expected based on random partitioning of two $MCP_{TM\text{-}M28L/V31L}$ peptides into micelles according to the Poisson distribution (Kobus & Fleming, 2005). These results showed that this variant of the TM domain of MCP undergoes specific self-association, with a standard free energy of -3.8 kcal/mol. This is a smaller energy of interaction than that of GpA, which has been shown to have a strong dimerization affinity of -5.5 kcal/mol in SDS micelles (Fisher et al., 1999). This lower affinity in the MCP peptide suggests that, while the interaction of $MCP_{TM\text{-}M28L/V31L}$ is specific, it is also relatively weak, allowing for significant populations of monomer to exist over a wide range of detergent:protein ratios (Wu et al., 2007).

A model for the MCP_{TM} dimer structure in β-OG micelles was also proposed using distance restraints from paramagnetic resonance enhancement (PRE) experiments on a MTSL-labeled cysteine mutant of the peptide, chemical shifts from ^{15}N-HSQC-TOCSY and ^{15}N-HSQC-NOESY spectra, and the docking program HADDOCK (Domingues et al., 2003; Wu et al., 2007). The model (Figure 1.8C) shows a right-handed coiled coil, as expected for a GXXXG motif interaction. Residues at the dimer interface included Ala35 and Leu31, which was consistent with the cross-linking studies carried out on full-length MCP in biological membranes (Haigh & Webster, 1998; Nagler et al., 2007).

1.2.5 I vs L Conformation of MCP

It is possible that the mobility of the N-terminal helix may play a role in modulating dimerization of MCP within the cellular membrane. This is supported by electron spin resonance (ESR) and fluorescence studies done on labeled MCP cysteine mutants in DOPC model membranes that suggest there is a dynamic equilibrium between different conformational states of the N-terminal amphipathic helix (Spruijt & Meijer, 2000). It was proposed that the N-terminal amphipathic helix fluctuates between a membrane-bound state where it lies on the surface of the lipid bilayer (the 'L' state), and an extended configuration to form a continuous helical structure with the transmembrane segment (the 'I' state) (Spruijt & Meijer, 2000; Vos et al., 2009). This raises the possibility that interaction between the amphipathic helix could be modulated by changes in the local lipid environment, which in turn may affect the strength of the interaction between TM helices in the membrane. Since the MCP dimer must be able to dissociate to form the new interactions found in the phage particle, this suggests that one of the roles of the N-terminal amphipathic helix could be to reduce TM helix-helix association thereby allowing phage packing interactions to occur.

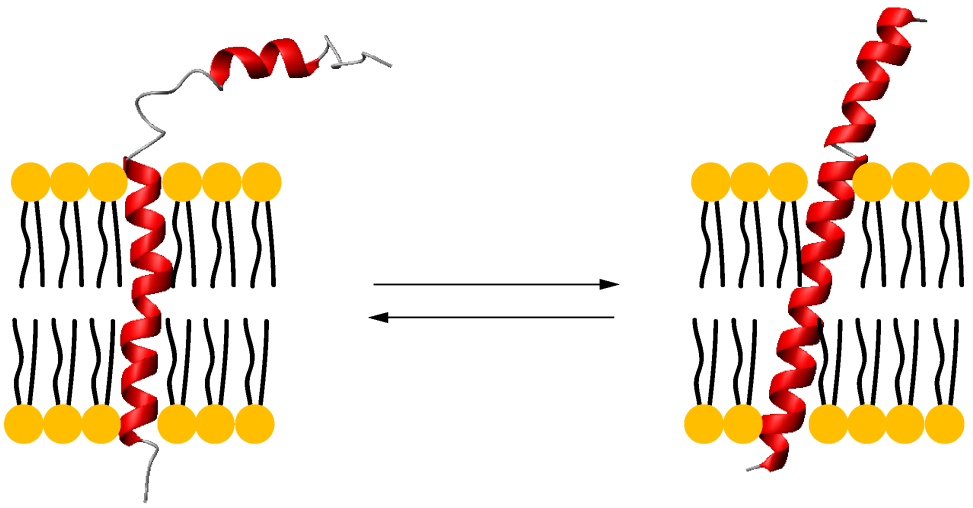

Figure 1.9 Proposed equilibrium between "L" and "I" conformations of MCP suggested by electron spin resonance studies (ESR) on cysteine mutants in DOPC model membranes.

1.2.6 Research focus and objectives

As discussed in Section 1.1, membrane proteins carry out many cellular functions via oligomerization of TM domains. While the interaction surface lies within the TM region in these proteins, in some cases other parts of the protein that are not directly involved in the interaction also appear to have an influence on oligomerization. The amphipathic helix has been shown to have an effect on the interaction of the transmembrane domains of Vpu1, phospholamban and C99 (Barrett et al., 2012; Hou et al., 2008; Ma et al., 2002). In all these examples there is evidence that these regions interact with the surface of the membrane, which may be important in modulating the impact of the amphipathic helix on TM helix interactions. However, little quantitative work has been done to measure the impact of these amphipathic helices on TM helix interactions. Therefore, one of the principle objectives in this thesis, and in particular the efforts described in Chapter 3, was to evaluate whether the amphipathic helix can change

TM helix interactions energies, to quantitate this change, and to try and understand the mechanism driving this difference. Although the examples of TM helices with adjacent amphipathic helices presented above are physiologically important systems, they are relatively difficult to work with and can be complicated by higher order oligomerization and/or difficulties in production of isotope-labeled samples required for NMR studies. Therefore, in order to examine the influence of an amphipathic helix on transmembrane helix self-association we chose to use the M13 bacteriophage MCP as a model system. The simplicity of the MCP structure, and the abundance of biophysical and biochemical data on this protein, makes it a good model system for this purpose, and was therefore the primary focus of this thesis.

Chapter 2: Theoretical Basis of Methods Utilized in Thesis

2.1 Solution-state nuclear magnetic spectroscopy (NMR)

2.1.1 Solution NMR principles

Nuclear magnetic resonance (NMR) spectroscopy is a technique that has been widely applied in the study of biological macromolecules such as proteins, often for the purpose of structure determination. Although X-ray crystallography has been a major source of atomic resolution protein structures, NMR has proven to be a complementary tool, allowing structures to be determined for smaller proteins when crystals were not available, with the unique capability to study conformational dynamics at high resolution under solution conditions (Snyder et al., 2005; Yee et al., 2005).

The physical basis of NMR spectroscopy relies on the intrinsic magnetic property of nuclei that have spin angular momentum (I) (Abragam, 1961). In protein NMR, spin-active isotopes that are commonly employed are ^{1}H, ^{15}N and ^{13}C, which all have an I value of ½. Under the influence of a homogenous magnetic field (B_0), as is provided by an NMR spectrometer, a dipole moment is induced in these nuclei. Each dipole adopts one of two orientations, parallel or anti-parallel with respect to the magnetic field, which are separated by a small difference in energy (ΔE) (Figure 2.1). In NMR spectroscopy, a radiofrequency (RF) pulse is applied to stimulate transitions between the low and high energy spin states, known as the Larmor frequency (ν). Absorption at this frequency stimulates transitions between the two states, and the frequency reflects influences from the local chemical environment surrounding the nucleus. The energy of this radiofrequency pulse is given by:

$$\Delta E = h\nu = \frac{h\gamma B_0}{2\pi} \qquad (1)$$

where ΔE is the energy difference between spin states oriented parallel and anti-parallel to the applied magnetic field, h is the Plank's constant and ν is the frequency of the pulse. B_0 is the applied magnetic field, which by convention is along the z-axis, and γ is the gyromagnetic ratio. The gyromagnetic ratio is a fundamental property of a spin and determines the size of the magnetic dipole moment induced by the interaction between the spin and the applied magnetic field. The frequency at which absorption occurs reflects contributions from the local chemical environment around the nucleus, and this is reported as the chemical shift in parts per million (ppm) from a reference frequency.

$$\delta = \frac{\upsilon_{sample} - \upsilon_{reference}}{\upsilon_{reference}} \qquad (2)$$

where ν_{sample} is the resonance frequency of the sample, and $\nu_{reference}$ is the resonance frequency of a standard reference compound.

The population of each energy state is determined by the Boltzmann distribution and the transitions are detected for the small excess population of nuclei in the lower energy state:

$$N_{high} / N_{low} = e^{-\Delta E / k_B T} \qquad (3)$$

Where N_{high} and N_{low} are the populations of nuclei in the high and low energy states, respectively, k is the Boltzmann constant and T is the temperature. The energy states are dependent on the orientation of the spin states, where for a nuclei with a positive gyromagnetic ratio, the low energy state is oriented parallel to the z-axis and high energy state oriented anti-parallel to this. The sensitivity of NMR is proportional to the difference in populations of the energy states. Since the difference is small, NMR is an insensitive technique. (Cavanagh et al., 1996; Claridge, 1999).

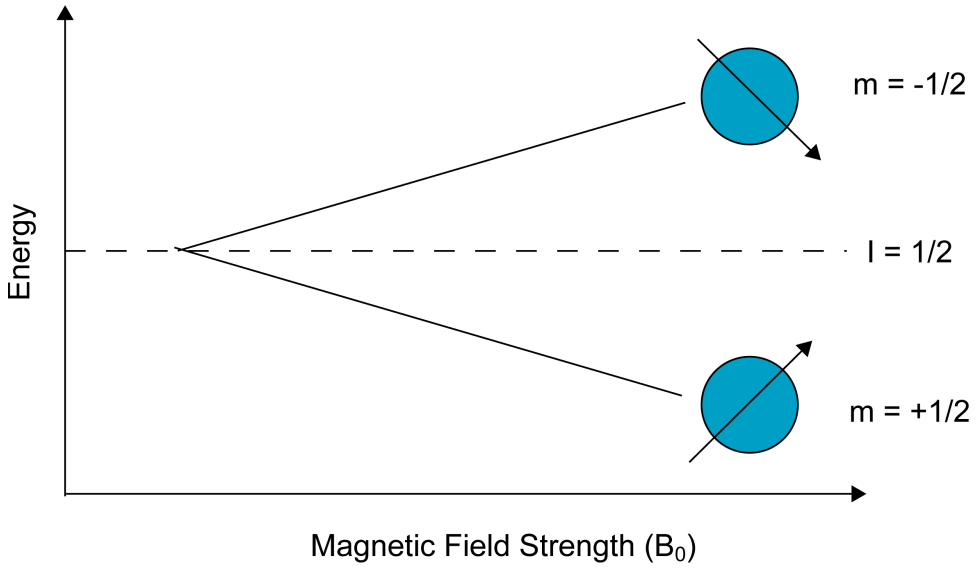

Figure 2.1 Splitting of energy levels of a spin 1/2 nuclei by the application of a magnetic field (B_0). Spins that are oriented parallel to B_0 are lower in energy, while those oriented anti-parallel to B_0 are higher in energy (for $\gamma>0$).

Molecular size also affects the sensitivity of NMR experiments, with a decrease in signal to noise with increasing molecular weight. This is due to deterioration of the signal in by transverse relaxation, which is the decay of the signal as a result of fluctuations in the local magnetic field. It is inversely proportional to the linewidth, therefore faster relaxation times result in the broadening of the peak in the NMR spectrum. Fluctuations in the local magnetic field impacts larger molecules because they undergo slower rotational diffusion, as measured by the rotational correlation time (τ_c), defined as the average time it takes for a molecule to rotate through one radian. This parameter can be related to the size of a spherical complex by the Debye equation:

$$\tau_c = \frac{4\pi\eta R_s^3}{3k_B T} \quad (4)$$

where k_B is the Boltzmann constant, T is temperature, η is the viscosity of the medium, V is the hydrodynamic volume of the molecule and R_s is the hydrodynamic radius of the molecule. According to this relationship, as the size of the molecule increases, the rate of tumbling decreases.

Molecular tumbling creates fluctuating fields due the dipoles from spins in the molecule. These fluctuating fields can induce transitions between states, leading to dephasing of the magnetization from the applied RF pulse. However, not all frequencies of field fluctuations cause transitions, with only those frequencies that can induce single, double or zero quantum transitions having this effect. Slower tumbling times give rise to higher densities of close to constant magnetic fields, increasing the frequency of zero quantum transitions for larger, slowly tumbling molecules. Relaxation-induced dephasing of magnetization results in a more rapid loss of signal, which gives rise to broader peaks in NMR spectra (larger linewidths). In addition, larger molecules have a larger number of dipoles that can give rise to a larger population of frequencies able to dephase the magnetization. Therefore, as the rate of tumbling decreases, the line-widths of peaks in the NMR spectra broaden, making it difficult to study larger molecules such as biological macromolecules like proteins.

To deal with the low sensitivity of NMR for proteins, samples are prepared with high concentrations of protein (~0.5 mM or more). Also uniform isotope labeling with spin 1/2 nuclei is utilized in NMR studies of proteins (Cavanagh et al., 2007). Isotopes that are most commonly used are ^{15}N and ^{13}C. Uniform labeling of proteins with these isotopes allows for recording of multidimensional NMR experiments that give rise to improved resolution of peaks by spectral editing and separation into additional

dimensions (Cavanagh et al., 2007). These experiments also provide correlations between spins that are connected by one or more bond, or are close to each other in space (Aue et al., 1976; Bodenhausen & Ruben, 1980; Gardner & Kay, 1998), critical information required for chemical shift assignment and structure determination.

2.1.2 The heteronuclear single quantum coherence (HSQC) experiment

The most commonly used experiment in the study of protein structure and function by NMR is the ^{1}H-^{15}N HSQC experiment. The ^{1}H-^{15}N HSQC experiment correlates the chemical shift of the amide proton with the chemical shift of its directly attached nitrogen atom. Therefore each residue in a protein, except for proline, should give rise to a backbone amide peak, giving rise to a "fingerprint" spectrum, with a unique pattern for every protein. Peaks can also arise from side chain N-H groups in this spectrum, namely from tryptophan, asparagine, glutamine, arginine and lysine, depending on the pH of the sample.

HSQC spectra can be useful diagnostics for the folded state of a protein, since a stably folded protein should give rise to peaks that span a wide range of proton chemical shifts. In contrast, unfolded proteins do not give rise to much dispersion in the proton chemical shift dimension, since most amino acids in an unfolded protein experience very similar local chemical environments.

The presence of oligomeric species can also be detected by ^{15}N-HSQC experiments, since the monomeric and oligomeric states will have different local chemical environments for residues that are involved in the interaction. However, the effect of oligomerization on the appearance of the spectrum will depend on the timescale

of exchange between species. There are three different exchange regimes that can be observed: (i) slow exchange (ms-s timescale) (ii) intermediate exchange (μs – ms) and (iii) fast exchange (ps-ns) (Figure 2.2). When there is slow exchange between oligomeric states, the spectrum will show two distinct peaks for residues that experience a change in local chemical environment upon oligomerization. If the exchange is on an intermediate timescale, peaks from residues involved in the interaction will be broadened relative to peaks from residues that do not participate in the interaction, with some change in chemical shift of the two species towards a population-weighted average. In the extreme case, these peaks can coalesce into one broad signal. At the other extreme are interactions that undergo fast exchange, which gives rise to a single peak at the population-weighted average chemical shift. In this thesis MCP in SDS micelles was examined, where the monomeric species was in slow exchange with the dimeric species.

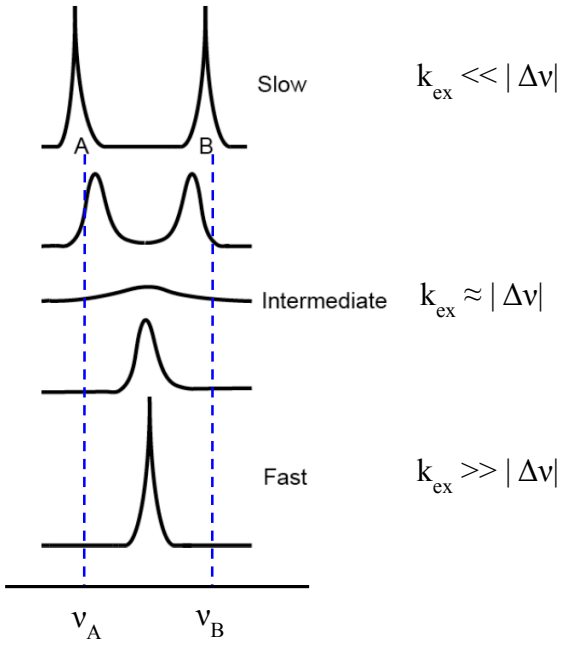

Figure 2.2 The effect of chemical exchange on 1D spectra under conditions of: (i) slow (ms-s timescale), (ii) intermediate (μs – ms) (iii) and fast exchange (ps-ns timescale). Exchange regimes are determined by the magnitude of the exchange constant (kex) and

the difference in absorbance frequency between the two states ($\Delta\nu$) in the relationships shown on the right.

2.1.3 Chemical shift assignment

In order to obtain atomic-resolution structure and dynamics data on proteins by solution NMR, it is necessary to assign chemical shifts using a series of three-dimensional experiments and ^{15}N, ^{13}C-labeled protein. One experiment that is key to backbone chemical shift assignment, is the CBCA(CO)NH (Grzesiek & Bax, 1992). In this experiment the amide nitrogen and proton chemical shifts of each residue (position *i*) are correlated with the alpha and beta carbon atoms of the preceding residue (*i*-1) (Figure 2.3). This 3D experiment can be visualized as a cube, with the end view of the ^{1}H-^{15}N dimension resembling a ^{1}H-^{15}N HSQC spectrum of the protein. The position of a peak along the z-axis will correspond to the chemical shift of the corresponding alpha and beta carbon of the preceding residues (Figure 2.4).

The CBCA(CO)NH experiment can be combined with the HNCACB experiment to identify peaks in the NMR spectrum that arise from adjacent residues. The HNCACB experiment correlates the nitrogen and amide proton chemical shifts of a residue (*i*) with the alpha and beta carbon chemical shifts of both the same residue (*i*) and the preceding residue (*i*-1) (Grzesiek & Bax, 1992) (Figure 2.3).

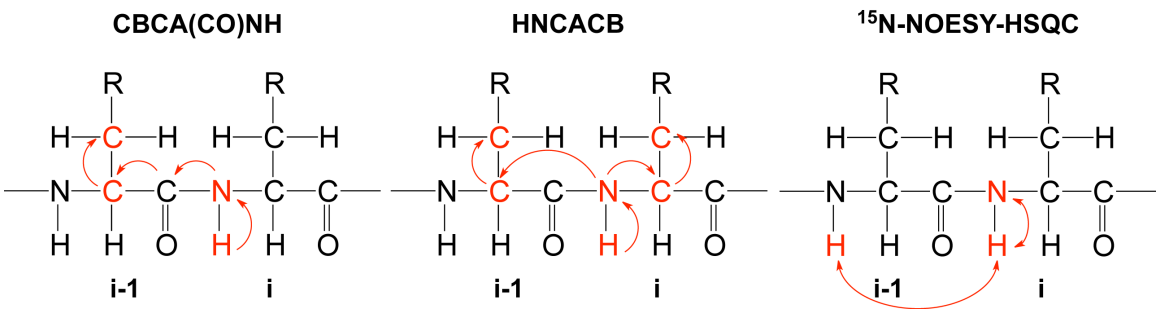

Figure 2.3 Correlations provided by 3D NMR experiments used for backbone assignment and secondary structure calculations in this thesis. The atoms highlighted in red indicate the nuclei that are correlated in each experiment. Magnetization transfer pathways are indicated with arrows. For the ^{15}N-NOESY-HSQC, only those that were assigned are shown.

The 3D experiments are analyzed as strips that correspond to the position in the ^{15}N dimension of the CBCA(CO)NH and HNCACB of the corresponding ^{15}N-HSQC where the peak is found. In the HNCACB experiment, each strip arising from a ^{1}H-^{15}N correlation will give rise to as many as four peaks along the ^{13}C dimension; two negative in intensity (red) and two positive in intensity (black). The two large peaks correspond to intra-residue Cα (black) and Cβ (red) correlations, and the smaller peaks correspond to inter-residue Cα (black) and Cβ (red) correlations with the preceding residue. By matching strips from the HNCACB experiment with those in the CBCA(CO)NH spectrum, it is possible to confirm the presence of inter-residue peaks (Cavanagh et al., 1996; Kanelis et al., 2001; Sattler, 1999) (Figure 2.4). The larger peaks in the HNCACB can be used to identify the strip in the HNCACB that comes from the subsequent residue in the sequence since this strip should have inter-residue correlations with carbon chemical shifts that match the intra-residue correlations in the preceding strip.

This process allows strip sequences to be determined, which can then be matched to protein amino acid sequence by comparison of carbon shift values with characteristic

shift values of each amino acid as derived from the chemical shift database ("Biological Magnetic Resonance Data Bank," ; Johnson, 2004). Some of the shifts for these atoms are distinct, as is the case for alanine, serine, threonine and glycine residues making them easier to identify and assign. For smaller proteins such as M13 MCP, determination of strip sequence along with comparison with chemical shift database values and the known amino acid sequence allows us the unambiguous assignment of backbone chemical shifts.

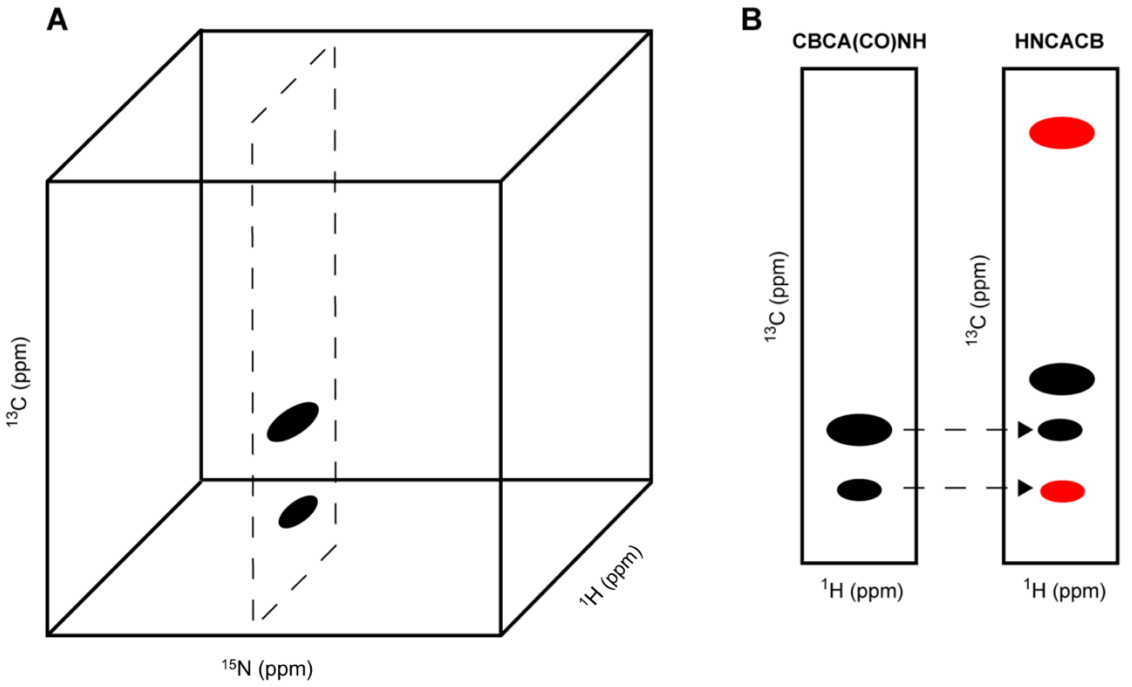

Figure 2.4 Overview of the triple resonance experiments used for backbone assignment. (A) Simplified schematic of peaks from a single residue pair in a CBCA(CO)NH spectrum. Correlations are obtained between the amide proton, the directly attached nitrogen and the Cα and Cβ of the preceding (i-1) residue. Resonances that correspond to the ^{1}H-^{15}N HSQC are separated in the carbon dimension to obtain the Cα and Cβ shifts of the preceding residue (i-1). (B) 3D experiments are analyzed as strips that correspond to the position in the ^{15}N dimension of the CBCA(CO)NH and HNCACB of the corresponding ^{15}N-HSQC where the peak is found. For the CBCA(CO)NH two peaks are detected for each residue, corresponding to the i-1 Cα (large peak) and Cβ (small peak). For the HNCACB four peaks are detected for each residue corresponding to the Cα of residue i (larger black peak), Cα of residue i-1 (smaller black peak), Cβ of residue i (larger red peak) and Cβ of residue of i-1 (smaller red peak). Red and black denote peaks of positive and negative intensities, respectively.

2.1.4 NMR spectroscopy of membrane proteins

The application of solution-state NMR to membrane proteins presents many challenges. These challenges include difficulties in the production of isotope-labeled membrane protein samples, in their solubilization in solvents that can support a folded state, and in the acquisition of NMR data on large protein-detergent complexes.

One of the most common methods used to generate isotope labeled protein samples is to use *E. coli* as an expression host, as was utilized in this thesis. It is the most cost effective method for the production of isotope labeled protein and has also proven useful for the expression of membrane proteins (Baneyx, 1999; Cunninghame & Deber, 2007; Drew et al. 2003; Laage & Langosch, 2001; Mao et al., 2009; Schneider et al., 2010; Schneider et al., 2009; Sørensen & Mortensen, 2005; Suzuk et al., 2005; Vaiphei et al., 2011). Membrane protein expression in *E. coli* often results in incorporation of the target protein into the cell membrane from which it can be extracted (Columbus et al., 2006; Duquesne & Sturgis, 2010; Gautier et al., 2010; Mancia & Love, 2010; Van Horn et al., 2009; Wang et al., 2003; Zhou et al., 2008). This form of expression, however, can lead to toxicity and cell death, making expression in *E. coli* impossible in some cases (Drew et al., 2003; Klepsch et al., 2011; Laage & Langosch, 2001; Meyenburg et al., 1985; Wagner et al., 2007; Wagner et al., 2006). An alternate strategy to incorporation in the membrane is to target expression into inclusion bodies. This can happen spontaneously, or through fusion tags that help to target the desired membrane protein into inclusion bodies (Binding & Kiefer, 1996; Park et al., 2003).

Following expression, membrane proteins must be isolated in an environment that maintains its fold and structure by mimicking the native lipid bilayer environment. This is usually accomplished by the use of detergents to solubilize membrane proteins (Garavito & Ferguson-Miller, 2001; Kim et al., 2009; Krueger-Koplin et al., 2004; Poget & Girvin, 2007; Qureshi & Goto, 2012; Sanders & Sönnichsen, 2006; Seddon et al., 2004). It is also essential that the size of the complex that forms with the protein and the solubilizing agent is not too large for the application of standard NMR techniques, preferably smaller

than ~40 kDa (Fernández & Wider, 2003; Kim et al., 2009; Sanders & Sönnichsen, 2006). For this reason, a commonly employed membrane-mimetic system for solution NMR studies of membrane proteins are detergents since they have a tendency to form relatively small, approximately spherical micelles (Poget & Girvin, 2007). In this assembly, the detergent head group localizes to the micelle surface, with the hydrocarbon tails being sequestered away from the water in the micelle interior. In the presence of membrane proteins, the detergent molecules aggregate around the hydrophobic region of the TM domain of the protein to reduce exposure to the aqueous environment, creating a soluble protein-detergent complex (PDC) that can be studied by NMR.

Developments in technology and methodology allowed the study of larger and more complex systemts. This can be seen in the increase in the number of membrane protein structures determined to date by solution-state NMR, at 93 for α-helical membrane proteins (including single and multi-spanning TM domains) and 36 for TM β-barrel proteins (as seen in "Membrane proteins of known structure determined by NMR http://www.drorlist.com/nmr/MPNMR.html#alpha,"). Although these numbers are small relative to the total number of structures in the database, the number of multi-spanning integral membrane protein structures solved by solution NMR has shown exponential growth since the structure of the GpA dimer was solved in 1997, the first structure of an integral membrane protein with more than one TM helix (Leong et al., 1997; Qureshi & Goto, 2012). Currently proteins up to the size of ~40-50 kDa can be studied by solution-state NMR (Kaptein & Wagner, 2015). One of the largest membrane proteins solved by solution NMR is the 7 TM-helix protein photosensitive rhodopsin II (pSRII), which was determined using NOEs obtained on a sample with selectively methyl protonated

isoleucine/leucine/valine (Gautier et al., 2010). This resulted in a high quality structure, however, NOE assignment for large proteins with broad signals can be very time- and resource-intensive. Other methods of structure determination that do not require assignment of side chain protons and NOEs can be used in these cases to determine protein global folds. For example, the pSRII structure was also solved using lanthanide-induced pseudo contact shift (PCSs) restraints, a method that only requires backbone chemical shift assignments and measurement of PCSs in ^{1}H-^{15}H HSQCs (Crick et al., 2015). The potential to apply this technique and other advances in methodology to the study of membrane proteins highlights the growing potential of solution NMR to contribute to the expanding database of membrane proteins (Kaptein & Wagner, 2015).

2.1.4.1 Thermodynamics of TM helix self-association using the continuum model

Solution-state NMR can be utilized to study the interactions of TM domains. Membrane mimetic systems that are used as a solvent to maintain the protein fold contribute to the apparent energy of interaction of these domains. This is due to contributions to the measured energy from interactions between protein and detergent, detergent and detergent, along with the protein-protein interaction of interest. However, it is possible to standardize measured free energies in a particular solvent using a model that is commonly implemented in the analysis of TM helix-helix interactions, known as the continuous solvent model. This model assumes the micellar phase behaves like an ideal solvent, with rapid exchange of solute between micelles such that it can be treated like a continuous phase, separate from the aqueous region. The state that defines the standard energy of interaction refers to conditions where the concentration of micellar

detergent is 1 M. Measurement of standard free energies of interaction according to this model allows values to be compared for different interacting systems, so long as the same micellar solvent is used (Fleming, 2002).

To calculate the standard free energy of association (ΔG_x^0) for a protein confined to a micellar phase of an SDS solution, the following equation is used:

$$\Delta G_{app} = \Delta G_x^{\ 0} + RT \ln([\text{micellarSDS}]\gamma_{SDS}) \qquad (5)$$

where ΔG_{app} is the apparent free energy of association, [micellarSDS] is the total concentration of SDS minus the critical micelle concentration and γ_{SDS} is the activity coefficient for deviations from thermodynamically ideal solution conditions (K Fleming, 2002). The apparent free energy of association, of the helix-helix interaction, can be determined from the association equilibrium constant (K_x):

$$K_{app} = \frac{[D]}{[M]^2} \qquad (6)$$

Solution-state NMR can be used to solve for K_{app} by extracting the fraction of the protein in dimeric form (f_d) based on the peak volumes in ^{1}H-^{15}N HSQC spectra:

$$f_d = \frac{[D]}{[D]+[M]} \qquad (7)$$

By substituting equation 6 into equation 7, we obtain:

$$f_d = \frac{[M]}{K_{app}^{-1}+[M]} \qquad (8)$$

From K_{app} we can calculate ΔG_{app} by the following relationship:

$$\Delta G_{app} = -RT \ln K_{app} \qquad (9)$$

By applying these relationships and assuming that γ_{SDS} is 1, the dimer populations determined from NMR spectra (or another method, such as analytical ultracentrifugation, or FRET (Dutta et al., 2014; Fleming, 2002; Fleming & Engelman, 2001; Kobus & Fleming, 2005)), can be utilized to solve for the apparent free energy of association, which can then be used to solve for the standard free energy of assocation (Equation 8).

2.1.5 Diffusion coefficient measurements by solution NMR

Dimerization of the TM domain within a micelle could affect the size of the PDC, not only by the addition of a second subunit in the micelle, but also by altering the number of detergent molecules that are associated with the PDC. A commonly employed method to study this potential change in size is measurement of the translational diffusion coefficient (D_s), since it is directly related to the hydrodynamic radius (Altieri & Byrd, 1995; Dingley et al., 1995; Gounarides et al., 1999). If the complex is approximately spherical, then it is possible to use the Stokes-Einstein equation to relate the diffusion coefficient to the hydrodynamic size of the complex:

$$D_s = \frac{k_B T}{6\pi\eta R_s} \quad (10)$$

This relationship shows that the hydrodynamic radius (R_s) is inversely proportional to the translational diffusion constant; in other words, larger species diffuse more slowly.

It is possible to measure translational diffusion constants for protein-detergent complexes using pulse field gradient (PFG) experiments. While there are other techniques available to measure diffusion coefficients (e.g. light scattering, ultracentrifugation), many of these techniques can be difficult to optimize for detergent solutions (Cavanagh et al., 1996). Meanwhile PFG-NMR sample conditions can be identical to those used to study the structure and self-association properties of the membrane protein sample, making it an ideal technique for the study of protein-detergent complex size in this thesis.

Pulsed field gradients are magnetic field gradients that are applied to the sample in an NMR spectrometer, disturbing the homogeneity of external magnetic field in a position dependent manner. As a result of this inhomogoneity, the Larmor frequency, and therefore precessional frequencies of nuclei will depend on position in the NMR sample (Stejskal & Tanner, 1965). This altered Larmor frequency, ω_Z, which is the frequency of a nucleus at a position z along the z-axis is described by the following equation:

$$\omega_z = \gamma B_o + \gamma G_z z \quad (11)$$

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

In a PFG-spin echo experiment, a magnetic field gradient introduces inhomogeneity in the magnetic field, but a second gradient is subsequently applied at an equal but opposite magnitude relative to the first, which essentially reverses the action of the first gradient, to refocus the magnetization. In the PFG-spin echo experiment, a delay is included between these two gradients to allow diffusion to occur. If the spins diffuse to a different part of the tube during this time, then the effect of the second gradient pulse will not perfectly reverse the effect of the gradient pulse, giving rise to an incompletely refocused bulk magnetization vector. The more translational diffusion that occurs, the poorer the refocusing, which is reflected by a decrease in intensity in the 1D NMR spectrum. The decay as a function of gradient strength can be fit to a monoexponential function that depends on the parameters of the gradients, the diffusion time and the diffusion coefficient, as described in Section 3.2.4.4.

2.2 Circular dichroism spectroscopy of proteins

To study the impact of dimerization and oligomerization on the structure of protein systems that were analyzed in this thesis, circular dichroism spectroscopy was implemented, since this provides an estimate of secondary structure content. Circular dichroism (CD) is the difference in absorption of left-handed and right-handed circularly polarized light, normally seen in molecules with assymetric chromophores (Greenfield, 1996; Hou et al., 2008). CD is widely applied in the field of protein research, since there is asymmetry that arises from the restriction in rotation about the amide bonds that comprises the polypeptide backbone. The CD signal ($\Delta\varepsilon(\lambda)$), is by:

$$\Delta\varepsilon(\lambda) = \varepsilon_L(\lambda) - \varepsilon_R(\lambda) \qquad (12)$$

where $\Delta\varepsilon_L$ and $\Delta\varepsilon_R$ are the molar extinction coefficients of the left- and right- components of circularly polarized light at a particular wavelength (λ). For proteins the CD signal is usually reported in units of mean residual ellipticity (MRE), which is defined by:

$$[\theta]_{MRE} = \frac{100 \times \theta}{n \times c \times l} \qquad (13)$$

where **n** is the number of backbone peptide bonds in the protein, **c** is the protein concentration, **l** is the path length and θ=3298x$\Delta\varepsilon$.

CD spectroscopy of proteins can be utilized to obtain an estimate of protein secondary structure content. α-helices, β-sheets and random coil secondary structures can be detected by CD based on the $\Delta\varepsilon$ in the far-UV range (190 nm – 250 nm). CD spectra of α-helices have characteristic minima at ~222 nm and ~208 nm, and a maximum at 192 nm. For β-sheets the spectra typically gives rise to a minimum at ~215 nm and a maximum at ~198 nm. In the case of a random coil, there is a weakly positive signal at ~226 nm and a strong negative signal at ~205 nm (Greenfield, 1996; Hou et al., 2008; Sreerama & Woody, 1994; Sreerama & Woody, 2004). Changes in the secondary structure content of protein systems in this thesis, based on changes in signals at the described UV regions, were utilized to study the impact of dimerization and oligomerization on the protein.

Chapter 3: Modulation of TM helix dimerization of major coat protein of M13 bacteriophage by the N-terminal amphipathic helix

3.1 Introduction

Using the major coat protein (MCP) of M13 bacteriophage we studied the thermodynamic properties of the interaction of the TM domain. The following objectives were established to gain insight into the modulation of TM helix-helix interaction by the N-terminal amphipathic helix of MCP:

(1) Measure the effect of the N-terminal amphipathic helix on the affinity of MCP TM helix self association

(2) Determine the effect of micellar detergent systems on MCP TM helix interactions

3.2 Materials & Methods

3.2.1 M13 bacteriophage propagation

3.2.1.1 M13 bacteriophage propagation of unlabeled phage

To propagate M13 filamentous bacteriophage, the K38A *E. coli* strain (HfC+, T2R, relA1, pit-10, spoT1, ton A22, ompF627, phoA4, T2Rλ1) (Lyons LB et al., 1972) were plated from a stock stored in glycerol onto a non-selective LB agar plate and incubated overnight at 37°C (Henry et al., 1986; Papavoine et al., 1994; van de Ven et al., 1993). A single colony from the overnight growth was used to inoculate 50 mL of M9 minimal media (0.05 M Na_2HPO_4, 0.02 M KH_2PO_4, 0.009 M NaCl, 0.1% [w/v] NH_4Cl, 0.3% [w/v] glucose, 1% [v/v] Gibco MEM vitamins, 0.1 mM $CaCl_2$, 0.1 mM $MgSO_4$), in a 250 mL Erlenmeyer flask, which was stirred overnight at 220 rotations per minute (rpm) at 37°C. When the optical density (OD) at 600 nm reached ~0.7-0.8, the culture

was infected with M13 bacteriophage that was prepared in an LB solution (100 μl of phage at a concentration of 4 mg/mL in 200 mL of LB) before addition. Following 10 minutes of incubation, 25 mL of the infected culture was transferred to 750 mL of 1X minimal media in a 2.5 L Fernbach flask and incubated with shaking at 200 rpm overnight at 37°C.

When isotope-labeled samples were required, ^{15}N-labeled ammonium chloride and/or ^{13}C-labeled glucose were used in place of the corresponding natural abundance nutrients in the M9 minimal medium throughout the growth. For selectively ^{15}N-alanine and ^{15}N-valine labeled bacteriophage, the G11a1 strain of *E. coli* bacteria was used in place of K38A (Yale Coli Genetic Stock Center). Growth were carried out in M9 minimal medium supplemented with 100 mg/L of each amino acid (except for valine or alanine), 100 mg/L adenine and guanosine, and 50 mg/L of cytosine, thymine, and uracil. ^{15}N-labeled alanine (or valine) was added to a final concentration of 100 mg/L.

The following day, the culture was centrifuged at 4500 g for 20 minutes at 4°C. The phage was precipitated from the supernatant with 20 g/L NaCl and 60 g/L polyethylene glycol (PEG, MW 8000 g/mol, Bioshop). The solution was stirred for 15 minutes to completely dissolve the PEG, and centrifuged at 4500 g for 20 minutes at 4°C. The pellet was then solubilized in 10 mL of double distilled water (ddH_2O) and centrifuged at 23500 g, for 30 minutes at 4°C to remove any remaining cell debris. The supernatant was transferred to dialysis tubing (12000-14000 Da MW cut off) and dialyzed against 4 L of 1 mM EDTA, 10 mM Tris pH 8.0 (TE buffer) overnight at 4°C. The dialysis was repeated with 4 L of freshly prepared TE buffer, overnight. The sample was then filter sterilized (0.22 um Millex-GS Sterile Filter Unit [Millipore]).

The concentration of M13 bacteriophage obtained after purification was determined by absorbance at 280 nm in a quartz cuvette, using an extinction coefficient (ε) of 2.25 $L \cdot mol^{-1}cm^{-1}$ (calculated by ProtParam EXPASY). Phage solutions were stored at 4°C until isolation and purification of MCP.

3.2.2 Isolation and purification of MCP from M13 bacteriophage

To isolate and purify MCP from M13, a modified phenol extraction technique was implemented (optimization described in the Results) (Cohn & Conant, 1926; Knippers & Hoffmann-Berling, 1966; Konings et al., 1970). A stock solution of borate-saturated phenol was prepared by adding solid phenol (Bioshop) to 10 mM borate buffer (pH 8.0). The stock solution was then vortexed for 5 minutes and centrifuged at 4500g for 10 minutes at room temperature. The aqueous layer was removed and the organic layer was added in a 1:1 (v/v) ratio to 4 mL of M13 phage (4 mg/mL) in TE buffer in a 15 mL Falcon tube (Fisher). The phage-phenol solution was vigorously vortexed for ~20 minutes to ensure disruption of phage coat. The solution was then centrifuged at 4500 g for 10 minutes at room temperature. After centrifugation, the aqueous (top) layer was removed and discarded. The phenolic layer (bottom) was transferred to a new 15 mL Falcon tube and diluted, in a 1:1 (v/v) ratio, with methanol.

Reversed-phase HPLC (Waters, controlled by Empower) was performed on the phenol-extracted bacteriophage protein using an acetonitrile gradient. HPLC or Optima grade solvents (Fisher, Bioshop) were used, and solvents were filtered prior to use. A silica 300SB-C3 semi-prep column (Waters) was first rinsed with 100% acetonitrile, 0.1% trifluoroacetic acid (TFA), and then equilibrated in 30% acetonitrile, 0.1% TFA.

To prepare the sample for HPLC purification, 1 mL of MCP in phenol/methanol was transferred to a 1.5 mL microfuge tube (Fisher). The sample was then centrifuged at 16,000 g for 2 minutes and the supernatant drawn up, via syringe (Fisher). The 1 mL samples were then loaded onto the HPLC column through the injection port or via the injection valve and coil.

The column was then washed for 5 minutes, at a flow rate of 4 mL/min, in 30% aqueous acetonitrile (0.1% [v/v] TFA), and then the concentration of acetonitrile (MeCN) was increased to 100% (0.1% TFA) using a linear gradient over 50 minutes, and elution monitored by absorbance at 215 and 280 nm. MCP eluted in a peak that appears at approximately 60% MeCN. The column was then washed for 10 minutes with 100% MeCN, 0.1% TFA, and re-equilibrated in starting conditions for 10 minutes to set the column up for the next run. After all runs were complete, the column was extensively washed with 100% MeCN (no TFA) for storage.

Fractions containing MCP were combined in a single 50 mL Falcon tube, stored at -80°C overnight and lyophilized the following day. Lyophilized MCP was then stored at -20°C.

3.2.3 Peptide and protein concentration determination

Protein and peptide concentrations were determined using the bicinchoninic acid (BCA) protein assay kit (Thermo Scientific), following the manufacturer's instructions. Bovine serum albumin (BSA) standards in a concentration range from 25 mg/mL to 2000 mg/mL, were prepared by serial dilutions of a 2000 mg/mL stock solution of BSA (Thermo Scientific). A total 50 μL of each standard solution was typically used to

construct the standard curve, and 5 μL of protein/peptide solution was diluted in 45 μL of ddH_2O for protein concentration measurements.

3.2.4 NMR spectroscopy

3.2.4.1 Sample preparation

NMR samples were made by solubilizing the lyophilized MCP to a final concentration of ~0.8 mM – 1.2 mM in SDS at concentrations ranging from 50 mM – 500 mM. In a few cases where decylmaltoside and dodecylphosphocholine detergents were used in place of SDS, lower concentrations were used, as specified in the Results. Detergent solutions were prepared by solubilizing detergent in ddH_2O, adding 10% D_2O and adjusting the pH prior to addition of lyophilized protein. Once protein was solubilized the pH was adjusted to 5.5. Samples volumes were 500 μl and experiments were run in a standard 5 mm NMR tube.

The $^{15}N/^{13}C$ –labeled sample used for backbone chemical shift assignments contained 1.2 mM MCP in 100 mM deuterated SDS (d_{25}-SDS), 10% D_2O pH 5.5 in a 300 μl volume loaded into a Shigemi tube.

NMR samples that were prepared for diffusion analysis of selectively ^{15}N-alanine or ^{15}N-valine labeled MCP were prepared by solubilizing the lyophilized protein in 500 μL of 100 mM SDS and 10% D_2O, pH 5.5.

3.2.4.2 3D HNCACB, CBCA(CO)NH and ^{15}N-NOESY-HSQC for backbone assignment

NMR experiments for backbone chemical shift assignments were carried out at the Health Canada NMR Facility on a Bruker 600 MHz spectrometer equipped with a triple resonance cryoprobe, by Dr. Simon Sauve and Dr. Yves Aubin. All other NMR

experiments were conducted at the University of Ottawa NMR Facility on a Varian Inova 500 spectrometer equipped with a triple-resonance probe and pulsed field gradient.

Backbone assignment of the MCP monomer and dimer was done with HNCACB (Grzesiek & Bax, 1992) , CBCA(CO)NH (Grzeskie & Bax, 1992), and ^{15}N-HSQC-NOESY (Marion et al., 1989) spectra processed with nmrPipe (Delaglio et al. 1995), and NMRView was used to analyze the data (Johnson & Blevins, 1994; Johnson, 2004) A mixing time of 100 ms was used for the ^{15}N-HSQC-NOESY.

3.2.4.3 ^{1}H-^{15}N HSQC for dimer populations and secondary amide shift analysis

2D ^{1}H-^{15}N HSQC (Bodenhausen and Ruben, 1980) at 37°C were acquired to monitor sample quality during protein purification optimization, and for measurement of monomer/dimer populations. Experiments were typically run using 32 transients, 128 increments, a ^{1}H spectral width of 7017.4 Hz, ^{15}N spectral width of 1267.2 Hz and interscan delays of 4 s (for equilibrium population measurements). NMRPipe was used to process the all NMR data and, peak volumes were measured using the nlinLS module of nmrPipe (Delaglio et al., 1995).

Average backbone amide chemical shift differences (Δδ) between the monomeric and dimeric forms of MCP were calculated using $\Delta\delta=((\Delta\delta_{HN})^2+(\Delta\delta_N/5)^2)^{0.5}$, where $\Delta\delta_{HN}$ and $\Delta\delta_N$ are the chemical shift differences between the monomeric and dimeric species for the amide proton and nitrogen atoms, respectively.

3.2.4.4 Diffusion coefficient experiments

To measure the diffusion coefficient (D_s) values of selectively ^{15}N-alanine and ^{15}N-valine labeled full-length MCP, the bipolar gradient pulse pair longitudinal-eddy-current-delay (RAW-BPP-LED) sequence was utilized (Chou et al., 2004). The length of

the position-encoding gradients (δ) was 1.2 ms and the duration of the diffusion delay (T) was 100 ms. 1024 transients were accumulated at each gradient strength, with an inter-scan time delay (d1) of 2 s. The gradient strength (G) was increased by 2% increments from 2% to 90% of the maximum gradient strength, and the series of gradient strengths were queued to run in a randomized sequence to reduce systematic error. Standard error of measurements was determined from duplicate experiments run for gradient strengths of 2, 10, 20, 30, 40, 50, 60, 70, 80 and 90%.

Spectra were analyzed using VNMR (Agilent) and the intensities (I) of resolved peaks determined by integration using Varian software. These intensities were then converted to ratios relative to the reference spectrum peak intensity (I_0) at 2% maximal gradient strength. These intensity ratios were plotted against the fraction gradient strength, which gives a monoexponential decay that can be fit to (Stejskal & Tanner, 1965):

$$I = I_0 \exp[-(\gamma\delta G)^2(\Delta - \delta/3)D_s] \qquad (1)$$

3.2.4.5 Rate constants for dimerization of MCP in SDS micelles

1D ^{1}H spectra were otabined for selectively ^{15}N-alanine labeled MCP using ^{1}H-^{15}N HSQC with 256 transients, 1 increment, 1 second interscan delay and a spectral width of 7017.5 Hz. Using iNMR (Nucleomatica), spectra were integrated to determine peak volumes. Peak volumes were used to determine the population of protein in the dimeric and monomeric states. iNMR was then used to perform deconvolution, followed by smoothing, to determine peak linewidths. The linewidth measured for the monomer in a sample that is primarily monomeric was subtracted from the linewidth measured for the

monomer from a samples containing monomer and dimer, and multiplied by π to give the generalized dimerization rate constant (k_{dim}) (Bocharov, 2012; Rule, 2006).

3.2.5 Circular dichroism (CD) spectroscopy

Far-UV circular dichroism (CD) spectra were recorded on a Jasco-810 instrument with 0.1 mM MCP in SDS, pH 5.5, under detergent:protein molar ratios that favor either the monomer or dimer. Eight scans were performed from 250 to 200 nm with a step resolution of 0.2 nm, a speed of 20 nm/min, a bandwidth of 1.0 nm, and response time of 2 s using a 0.1 mm path-length quartz cuvette. Blank spectra for different 50 mM SDS and 15 mM conditions, without protein, were run and subtracted from spectra of protein sample.

3.3 Results and Discussion

3.3.1 Purification of MCP from M13 Bacteriophage

In order to utilize solution-state NMR to study the self-interaction of MCP, samples were prepared from M13 bacteriophage propagated in K38A *E. coli* grown in isotope-labeled media (^{15}N or $^{15}N/^{13}C$). Following isolation of MCP by the phenol extraction method, samples were lyophilized and solubilized in SDS, since the structural and dynamic properties of MCP are well characterized in this detergent system (Henry et al., 1986; C H Papavoine et al., 1994; van de Ven et al., 1993).

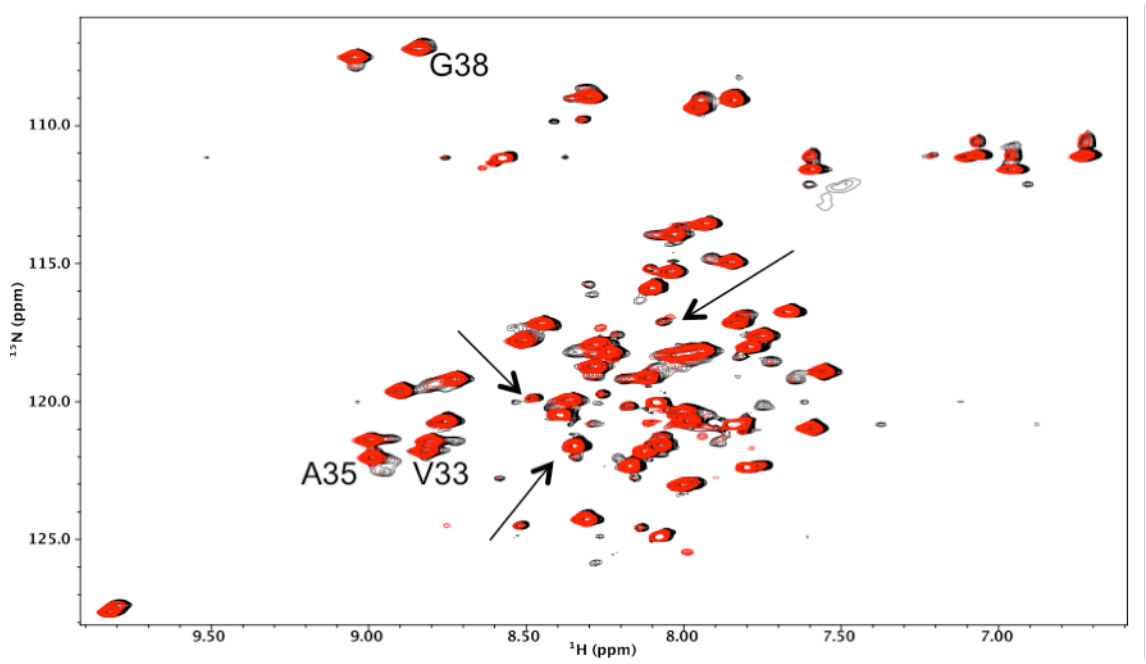

Figure 3.1 ^{1}H-^{15}N HSQC spectra of full-length MCP purified by phenol extraction followed by dialysis acquired in 500 mM SDS, 10% D_2O, pH 5.5 with 1.2 mM (black) or 0.6 mM (red) MCP. Arrows highlight some of the peaks that were not observed in previously published spectra of MCP. Residues for which a peak from the dimeric species was difficult to resolve are indicated with residue labels.

As shown in Figure 3.1, when the sample was solubilized in a high detergent:protein ratio, the NMR spectrum was similar to that which had been previously published for MCP in SDS micelles (van de Ven et al., 1993). However, relative to previously published spectra, there was poor resolution of monomeric and dimeric peaks for some residues, making it difficult to determine whether there was dimer present in the sample (e.g. Val33, Ala35 and Gly38, highlighted in Figure 3.1). Also, there were extra peaks in this spectrum that were not apparent in previously published spectra (indicated by arrows in Figure 3.1).

To investigate the source of these extra peaks, SDS-PAGE analysis was performed on the purified sample (Figure 3.2). As described in the introduction, MCP is expected to run as band between 6 and 12 kDa, however this was not observed in the gel. Instead, protein appeared to be distributed at a range of molecular weights, with the most prominent band appearing at ~ 30 kDa. Since MCP should comprise ~90 % of total bacteriophage protein, these higher molecular weight species may have been aggregates formed during the purification that were not disrupted by SDS. To determine whether this contaminant could be disrupted by higher concentrations of SDS, the sample showing extra peaks in the HSQC spectrum was diluted in SDS. However, no change in the intensity of contaminant peaks relative to the main species was observed (Figure 3.1, red spectrum), indicating that the contaminant was either another phage coat protein present in low abundance, or an irreversible aggregate of MCP. Previous CD studies done on MCP that examined the secondary structure of the protein directly following phenol extraction found that most of the protein forms a large aggregate composed of ~70% β-sheet (Peleen et al.,1992, Sprujit et al., 1989) indicating that we were likely seeing this higher order oligomer.

Considering that irreversible aggregation may have given rise to these extra peaks, a number of experiments were carried out to prevent its formation during the purification. One possible source of this aggregate could have been a precipitate that was observed during the phenol extraction that was found between the aqueous and organic layer. Although different centrifugation conditions, phenol suppliers and solution preparation methods were tested, in all cases a white precipitate was observed at the

phenol-aqueous interface. It is possible that this precipitate may have been carried though the purification, in spite of best efforts to avoid its transfer into the sample.

Another part of the phenol extraction protocol that may have caused irreversible aggregation was the dialysis from organic to aqueous solution, particularly given the obvious appearance of an aggregate during the dialysis. Although this had been reported to be a common occurrence in the MCP purification (Spruijt et al., 1989), we sought to minimize this aggregation by varying the number, content, temperature and duration of the dialysis steps. However, regardless of the protocol used, it was not possible to prevent the formation of precipitate during the dialysis, with most of MCP partitioning into this aggregate.

We then developed a modified purification technique for the separation of the major coat protein of M13 bacteriophage from the genetic material encased within the protein shell. The most commonly implemented technique for MCP purification in the literature is the phenol extraction. However, in a number of cases modifications were made to the phenol extraction protocol, including some that used size exclusion chromatography as a final purification step. When samples were analyzed by solution NMR, these different protocols were found to be equivalent, suggesting that the phenol extraction method was apparently sufficient for the preparation of MCP samples for NMR structural studies of the monomeric state (Henry & Sykes, 1992; Henry et al., 1986; Papavoine et al., 1994; van de Ven et al., 1993; Wickner, 1975).

We decided to develop an alternate method to purify MCP after phenol extraction, implementing reversed-phase high-performance liquid chromatography (RP-HPLC). This

chromatographic technique is commonly implemented in the separation of peptides based on their polarity, and has been used to purify hydrophobic polypeptides that encompass one or more TM segments (Carr, 2002; Clarke et al., 2006; Fassina et al., 1994; Lee et al., 1996; Oxenoid & Chou, 2005; Sato et al., 2002; Welling et al., 1987). A Zorbax 300SB-C3 (Agilent) semi-prep column was chosen for the separation, since the short carbon chain length of the stationary phase decreases its hydrophobicity and helps prevent non-specific aggregation on the column. To increase the quantity of protein that could be purified per injection, a semi-preparative scale column was chosen and a standard acetonitrile gradient was implemented. An example of the chromatography profile obtained for the purification of phenol-extracted MCP by RP-HPLC is shown in Figure 3.2.

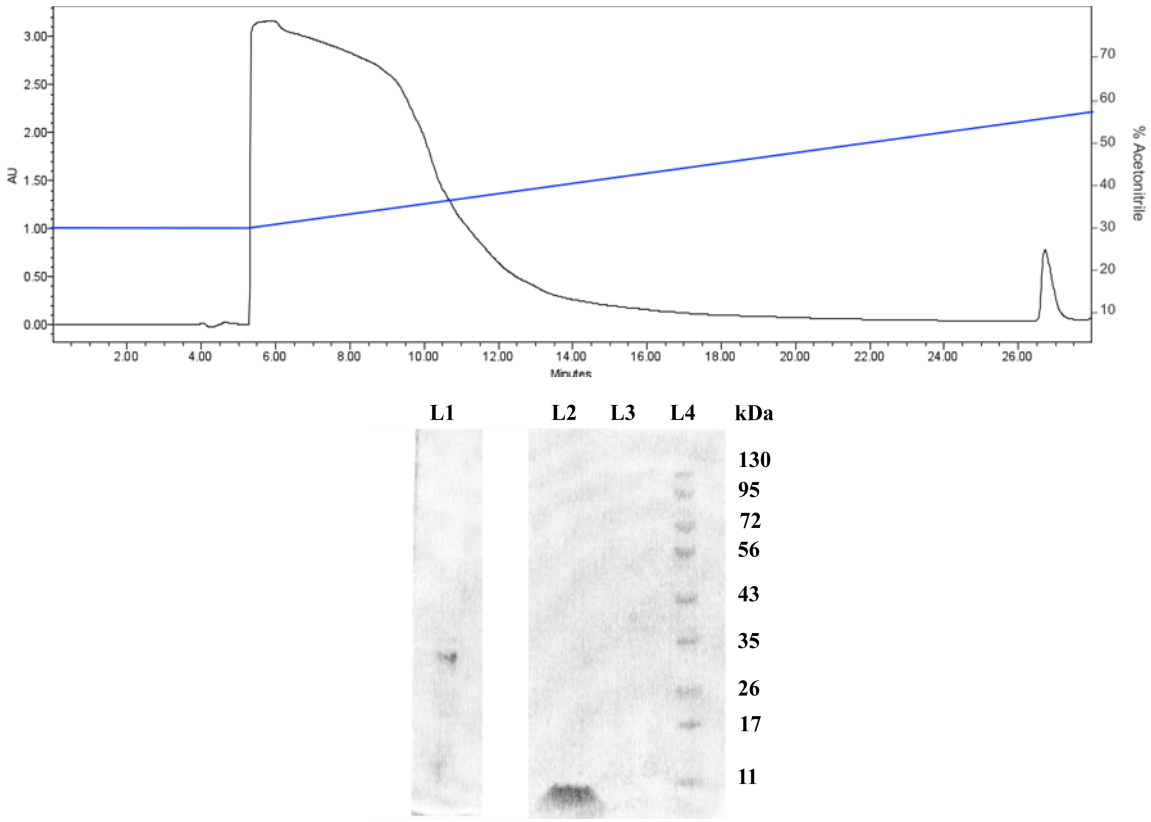

Figure 3.2 Purification of MCP by phenol extraction and RP-HPLC (A) RP-HPLC profile of a 1 mL sample of MCP in borate-saturated phenol solution purified on a Zorbax 300SB-C3 column, monitored at 280 nm. Blue line shows gradient of acetonitrile/ddH_2O (0.01% TFA) (B) Coomassie-stained SDS-PAGE gel of fractions collected during MCP purification. L1 – NMR sample prepared without phenol extraction (1 mM MCP, 300 mM SDS, 10% D_2O, pH 5.5) L2- fraction collected from RP-HPLC at ~ 58 % acetonitrile, L3 – fraction collected at 6-8 minutes at the beginning of the acetonitrile gradient, L4 – protein molecular weight ladder.

The RP-HPLC profile shows the elution of two major species with significant absorbance at 280 nm (Figure 3.2). The first species eluted over a ~10 minute time interval, and contained phenol with no detectable protein. A smaller peak eluting at approximately 58% acetonitrile was confirmed by SDS-PAGE gel to contain protein at the expected molecular weight at 6 to 12 kDa, corresponding to >95% pure MCP.

An NMR spectrum of ^{15}N-labeled MCP prepared using this method showed good peak dispersion and minimal overlap (Figure 3.3). Also, the number of peaks observed in this spectrum corresponds to the number expected for monomeric MCP, indicating that this HPLC-based method was effective for the generation of pure monomeric samples of MCP.

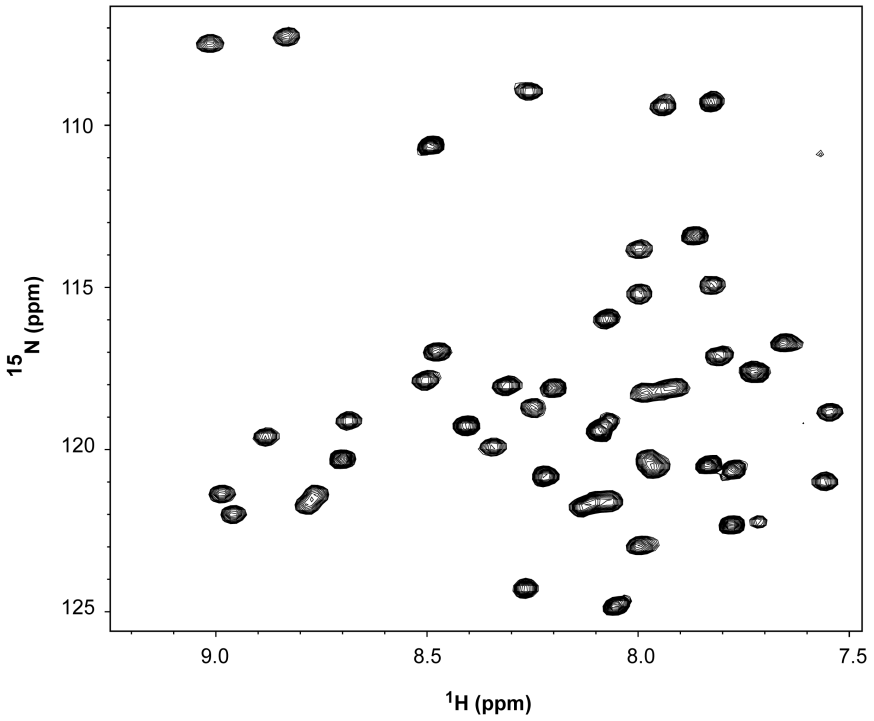

Figure 3.3 ^{1}H-^{15}N HSQC spectrum of monomeric MCP, containing 1 mM MCP in 300 mM SDS and 10% D_2O at a pH of 5.5

3.3.2 Secondary structure of full-length MCP in SDS micelles

Although we were interested in investigating the effects of regions outside of the transmembrane domain on MCP dimerization, it was first necessary to confirm that full-length MCP can adopt a dimeric state in detergent micelles. For this purpose,

an ^{1}H-^{15}N HSQC spectrum was acquired on a sample with an MCP:SDS molar ratio of 1:125, that would be expected to promote dimerization (Figure 3.4).

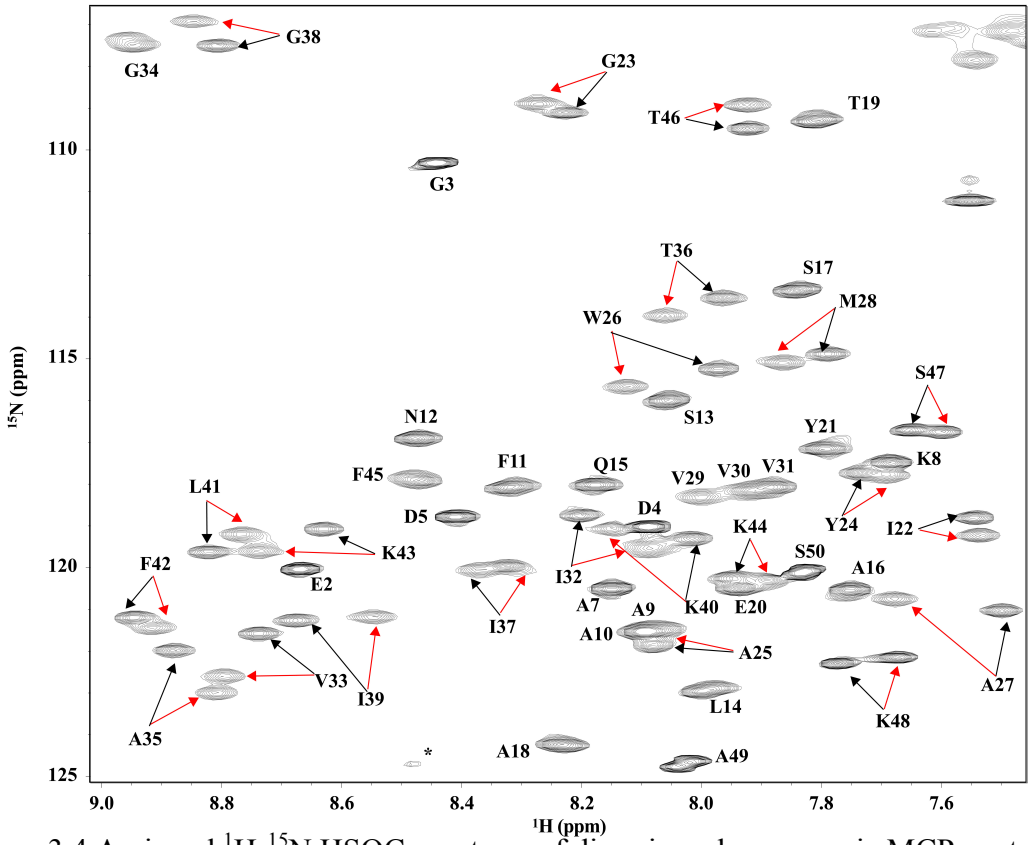

Figure 3.4 Assigned ^{1}H-^{15}N HSQC spectrum of dimeric and monomeric MCP containing 0.8 mM MCP, 100 mM SDS and 10% D_2O at a pH of 5.5. Black arrows highlight peaks from the monomeric state and red arrows show peaks of the dimeric state. Assignments from suite of standard 3D experiments.

The spectrum obtained was comparable to previously published spectra of an MCP sample containing a substantial population of dimeric species in SDS (Henry & Sykes, 1992). As expected, a subset of peaks resemble those of the spectrum for monomeric MCP and with the new set of peaks being similar to those previously been attributed to the MCP dimer.

Since no backbone carbon chemical shifts for the dimeric state are available and a complete assignment of dimeric species had not been previously undertaken, a standard

suite of 3D triple resonance experiments were performed on this sample for this purpose. As shown in representative strips taken from HNCACB and ^{15}N-NOESY-HSQC experiments (Fig. 3.5 and 3.6), it was possible to trace assignments for a subset of peaks that arose from the dimeric state that were distinct from peaks arising from the monomer.

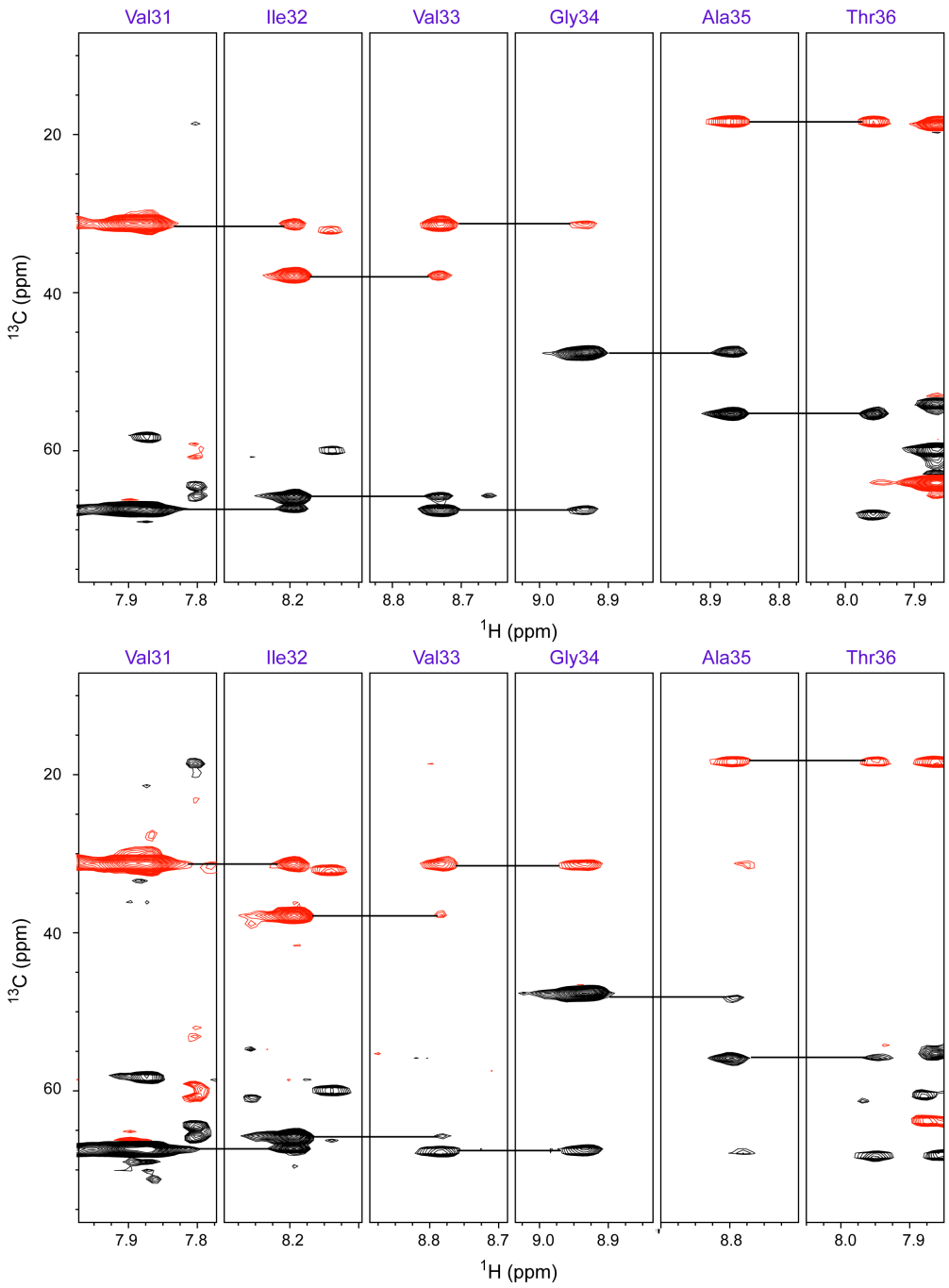

Figure 3.5 Representative strips from the HNCACB experiment showing sequential assignment of the amide backbone of monomeric (top) and dimeric (bottom) MCP. Lines have been drawn between peaks that arise from the same carbon atom correlated either to the intra-residue or inter-residue amide.

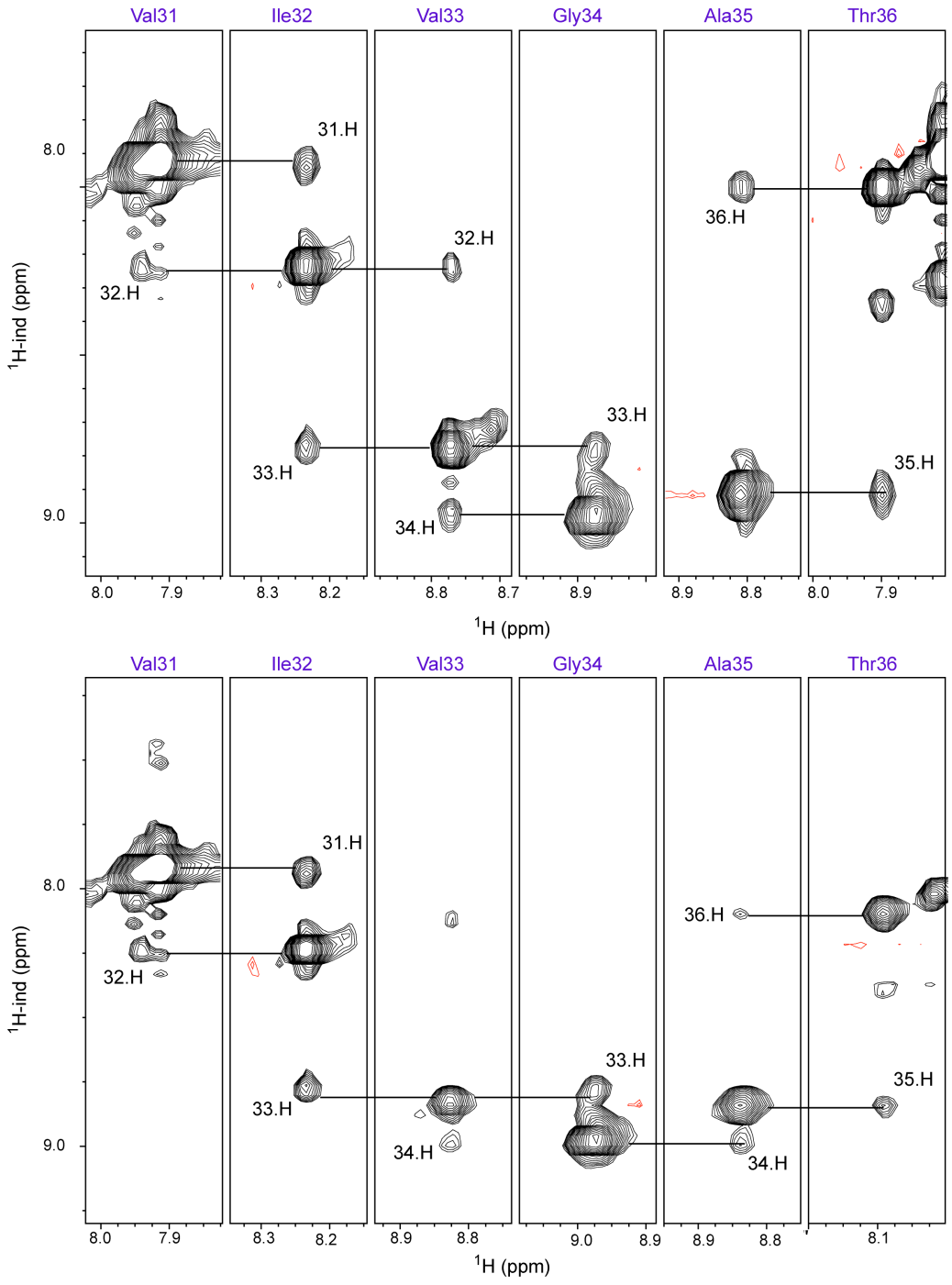

Figure 3.6 Representative strips from the amide region of the ^{15}N-HSQC-edited NOESY spectrum for regions from monomeric (top) and dimeric (bottom) MCP. Links shown between strips to indicate NOEs between adjacent amide protons, as expected for an α-helical structure.

Distinct assignments for monomer and dimer species reflect slow to intermediate timescale exchange between the two populations for residues that span the TM domain, except for residues Val29-Val31 and Phe45 which only showed one peak. Average amide shift differences between monomeric and dimeric species were most significant around Trp26 and Ala27, but were also significant for a number of residues throughout the TM segment (Fig. 3.7 A).

With these backbone assignments it was possible to assess MCP secondary structure based on secondary shift analysis (Wishart et al., 1991; Wishart & Sykes, 1994). As shown in Figure 3.7, there are strong positive secondary shifts for the Cα atom for most of the MCP, indicating a helical structure. In addition, the secondary structure between the monomeric and dimeric states were indistinguishable, with a helix predicted for residues 6-19 and 25-45.

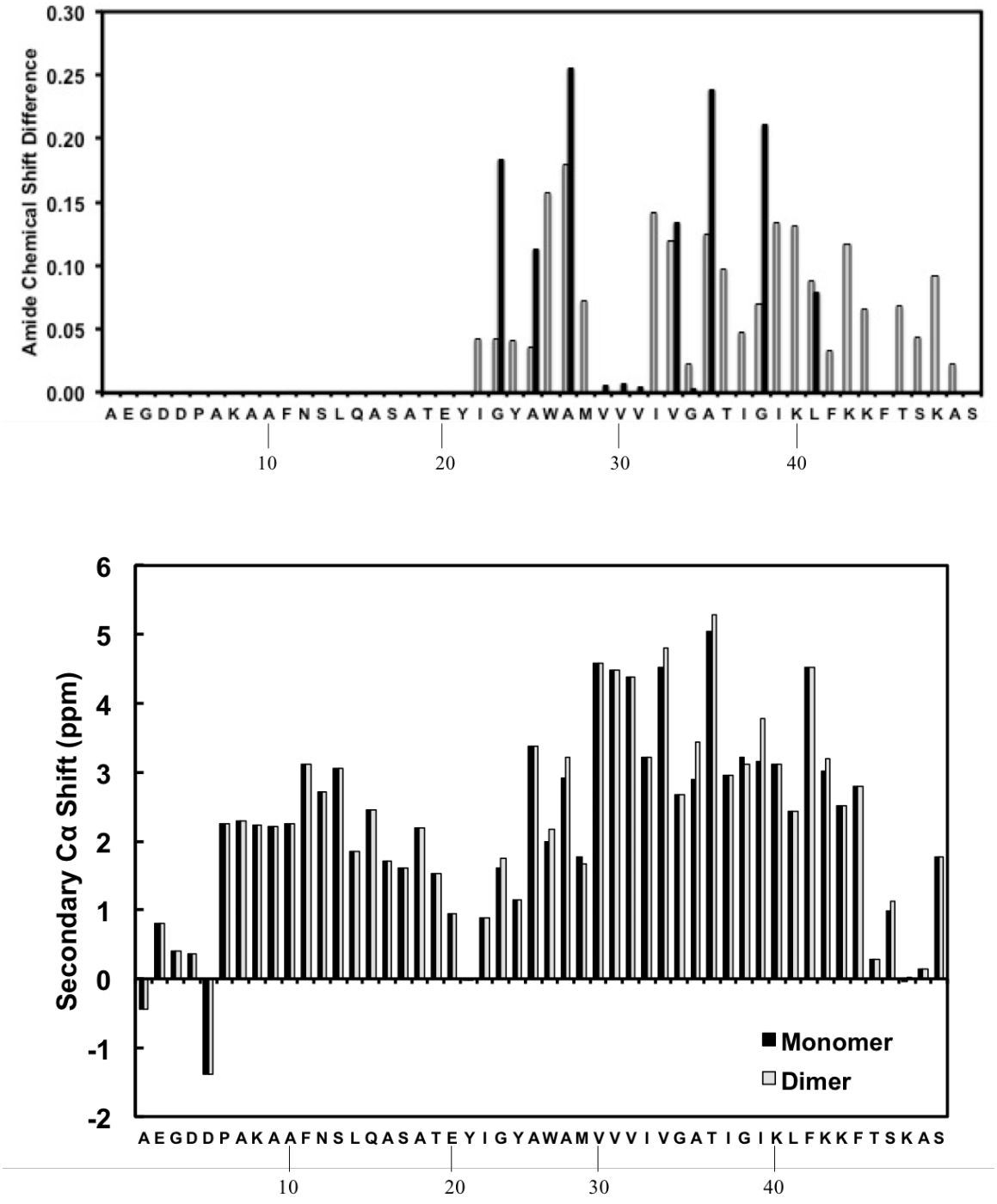

Figure 3.7 Top - Average amide chemical shift differences (Δδ) of the transmembrane peptide (gray) and full-length MCP (black); Bottom - Secondary chemical shift analysis for Cα atoms on full length MCP.

3.3.3 Detection of monomeric and dimeric MCP_{TM} species in SDS micelles

Since we were interested in investigating the effect of the N-terminal amphipathic helix on MCP dimerization, a peptide was synthesized consisting of the TM domain of MCP (Residues 21-40) with ^{15}N-labeled residues incorporated at 11 sites (Figure 3.8). This labeling scheme distributes ^{15}N-labeled residues throughout the dimerization interface, to facilitate the study of MCP_{TM} self-association by NMR, particularly since the spectrum of full length MCP showed significant chemical shift differences between monomer and dimer at these sites. At the N-terminus a lysine tag (consisting of three lysine residues) was introduced and the C-terminus was amidated to enhance peptide solubility (Melnyk et al., 2003).

MCP_{TM} KKKYI**G**Y**A**W**A**M**VVV**I**V**GA**TI**G**IK**L**FKKFTSK

Figure 3.8 Sequence of the synthesized transmembrane peptide with lysine tag at the N-terminus. The selectively ^{15}N-labeled residues are shown in blue.

An ^{1}H-^{15}N HSQC spectrum of an MCP_{TM} sample with a relatively high peptide:detergent ratio (1:100) was acquired and found to contain 11 peaks of strong intensity, and another set of peaks corresponding to a less abundant population. Based on similarities in chemical shifts between these peaks and those in the spectrum for full-length MCP, it was possible to assign the major species as the dimeric state and the minor species as the monomeric species (Figure 3.9). Average amide shift differences between monomer and dimer states followed a similar trend to that observed for the full-length

protein, although the magnitude of the shift difference was generally larger for the TM peptide (Fig. 3.7 A).

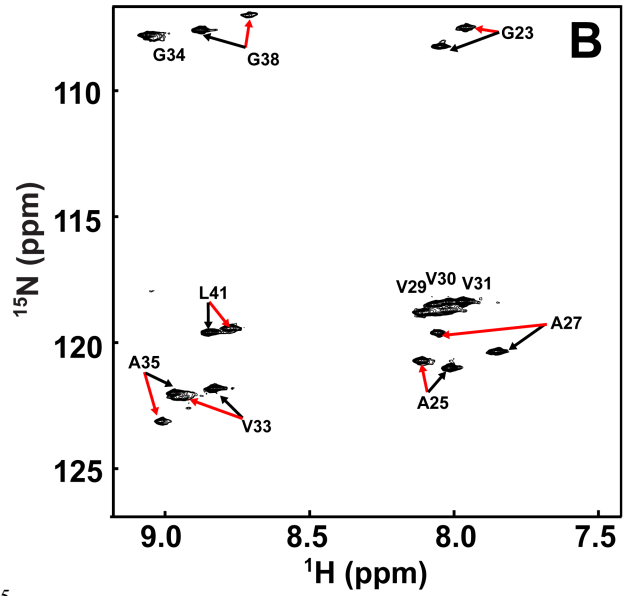

Figure 3.9 ^{1}H-^{15}N HSQC of 0.5 mM MCP_{TM} with 50 mM SDS in 10% D_2O, pH 5.5. Black arrows highlight peaks from the monomeric state and red arrows show peaks of the dimeric state.

3.3.4 Secondary structure of MCP_{TM} during dimerization

Since backbone chemical shift assignments were not available for the TM peptide, we used circular dichroism spectroscopy to determine the effect of dimerization on its secondary structure. As shown in Figure 3.10, when CD spectra were acquired for full-length MCP under conditions that favoured the monomer or dimeric state, the CD spectra were superimposable, confirming the preservation of helical structure that was shown in the secondary shift analysis. Similarly, CD spectra of the MCP_{TM} peptide in conditions favoring the monomeric and dimeric states where almost identical, with the minima at 208 and 222 nm that is characteristic of helical structure.

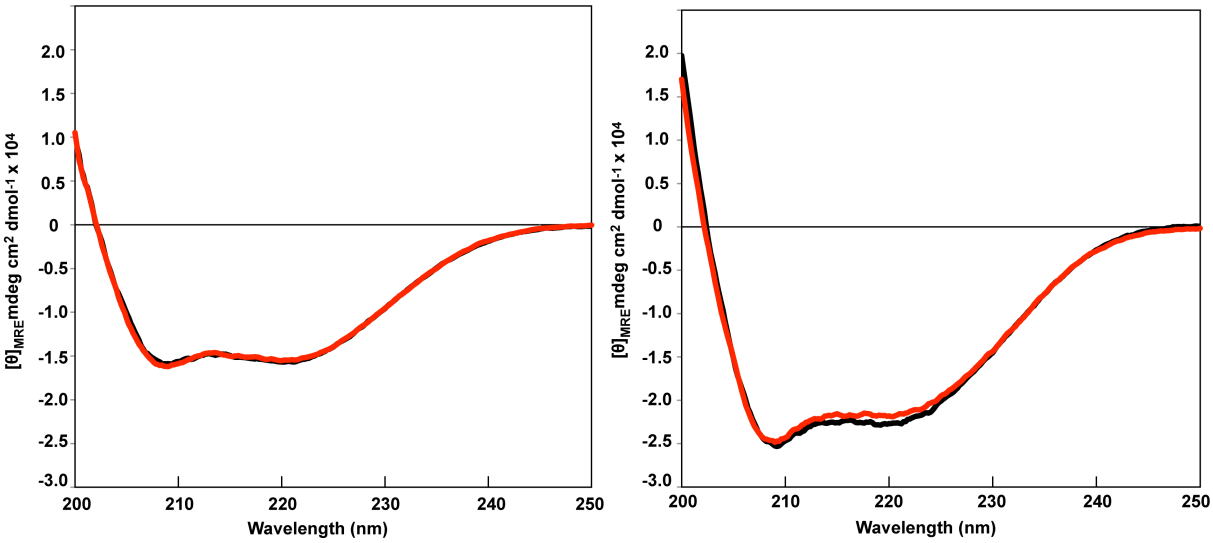

Figure 3.10 CD spectra acquired for monomeric (red) and dimeric (black) MCP (left) and MCP_{TM} (right).

Mean residue ellipticities for peptide spectra were more intense than the full-length protein, consistent with a peptide that is expected to contain an α-helix with a small number of unstructured residues flanking this helix. The full-length MCP would also be expected to have non-helical structure between the two helices that would decrease the proportion of helical content relative to MCP_{TM}, even though the TM helices are likely of the same length between the two MCP samples.

3.3.5 Thermodynamic properties of MCP self-association

To determine the impact of removal of the N-terminal helix on MCP TM domain interactions, we used ^{1}H-^{15}N HSQC experiments to monitor the effect of changing the protein:detergent molar ratio on population of monomeric and dimeric species for both full-length MCP and MCP_{TM}. These populations were used to calculate the standard free

energy of interaction based on the continuum model as described in the Chapter 1. For this calculation a subset of peaks with well separated monomer and dimer resonances was selected (Figure 3.11) and peak intensities were used to calculate relative populations over a range of protein:detergent ratios.

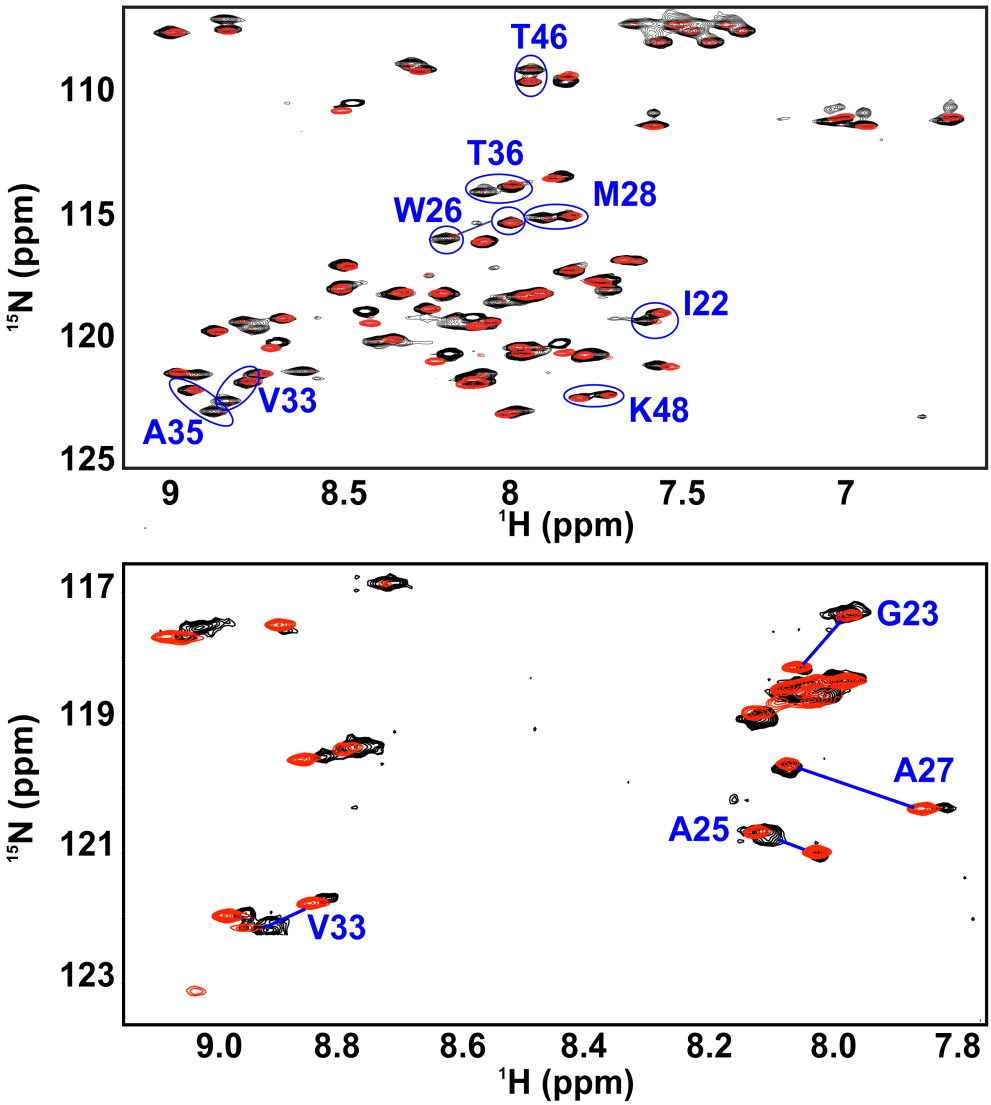

Figure 3.11 Representative ^{1}H-^{15}N HSQC of MCP (top) and MCP$_{TM}$ (bottom) in conditions favoring the monomeric (red spectrum) and dimeric (black spectrum) states. The blue labels indicate peaks whose intensities were measured to monitor the monomer-dimer equilibrium. Note that the spectrum in B was recorded with a reduced spectral width, with peaks from Gly being folded back into the spectrum.

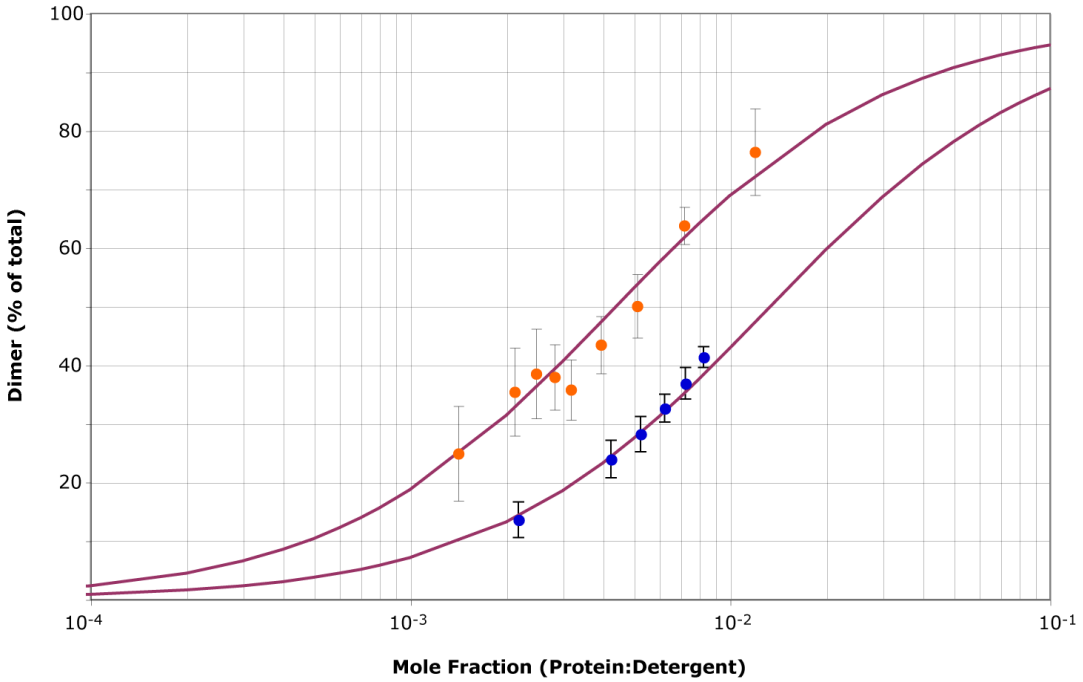

Figure 3.12 Association of the full length (blue) and TM peptide (orange) of MCP in SDS micelles. Each point represents the fraction of the population that is dimeric as a percentage of the total protein concentration plotted against the mole fraction of protein to micellar detergent. The purple line represents the fit of this data to a monomer-dimer equilibrium.

As shown in Figure 3.12, over a range of protein:detergent ratios, these populations could be fit to a free energy of self-interaction of -2.7 (± 0.1) kcal/mol with respect to a standard state of 1 M micellar SDS for the full-length MCP. In the case of MCP_{TM}, which is missing the N-terminal helix, the standard energy of association was found to be -3.4 (± 0.1) kcal/mol, reflecting a stronger interaction. This confirms that the N-terminal amphipathic helix of MCP weakens the affinity of its self-interaction. The standard energies obtained for MCP are consistent with values obtained for other weakly interacting TM domain peptides, such as receptor tyrosine kinase ErbB4 TM domain (-2.7 kcal/mol) (Bocharav et al., 2012) in DMPC/DHPC micelles. Another well studied protein system is glycophorin A (GpA), where the standard free energy of interaction in

SDS micelles was found to be -5.5 kcal/mol, which reflects a stronger interaction as would be expected from its observed migration as a dimer in SDS-PAGE analyses (Fisher et al., 1999). The interaction energies obtained in this thesis are also smaller than that of a mutant MCP TM peptide, which was found to be -3.8 kcal/mol in SDS (Wu et al., 2007). This is consistent with the expectation that these mutations in the peptide should enhance the affinity of dimerization, since a greater proportion of dimer was observed in SDS-PAGE analysis of full-length MCP harboring these same mutations (Melnyk et al., 2002).

Our study is the first in which the energy of interaction of full-length MCP was obtained and the experiments carried out in a manner that allow comparison of the values obtained for the standard free energy of interactions to those obtained for the peptides of MCP. It's also the first study that quantitates the magnitude of the reduction in TM helix self-association energy by the inclusion of an amphipathic helix adjacent to the TM segment, demonstrating that the equilibrium between the monomer and dimer is affected by the amphipathic helix. This may arise from interactions of the amphipathic helix with the micelle surface, an interaction may also be in effect in the lipid membrane, as was suggested by ESR and fluorescence for MCP reconstituted in lipid bilayers (Spruijt & Meijer, 2000). This interaction may reduce TM helix dimerization affinity, and ultimately may be required to facilitate formation of the oligomer that comprises the bacteriophage coat. (Glucksmant & Bhattacharjee, 1992; Marvin & Wachtel, 1976; Nagler et al., 2007; Roth et al., 2002; Williams et al., 1995).

3.3.6 Evaluation of the ideality of SDS as a solvent in measurement of TM helix interactions

The solvent model used to analyze the ^{1}H-^{15}N HSQC data for calculation of dimerization affinities was the continuum solvent model (described in the Introduction). According to this model the detergent phase can be approximated as a continuous solvent phase, with rapid exchange of polypeptides between micelles (Fleming, 2002). If the detergent phase behaves like an ideal solvent then the apparent affinity measured at each detergent concentration could be normalized for differences in detergent concentration by simply multiplying the apparent affinity constant by the concentration of detergent in the micellar phase. If the system does in fact follow the continuum model then this detergent concentration-normalized affinity constant should be constant over the range of detergent:protein ratios tested. To evaluate the ideality of SDS in our system, the detergent-normalized dissociation constant was plotted as a function of the detergent:protein ratio. As shown in Figure 3.13, while the detergent-normalized affinity was constant at high detergent:protein ratios, there was an increase in the apparent affinity for dimerization at lower detergent:protein ratios that could not be normalized by the detergent concentration. Similar behavior was observed for both the full-length and TM peptide of MCP.

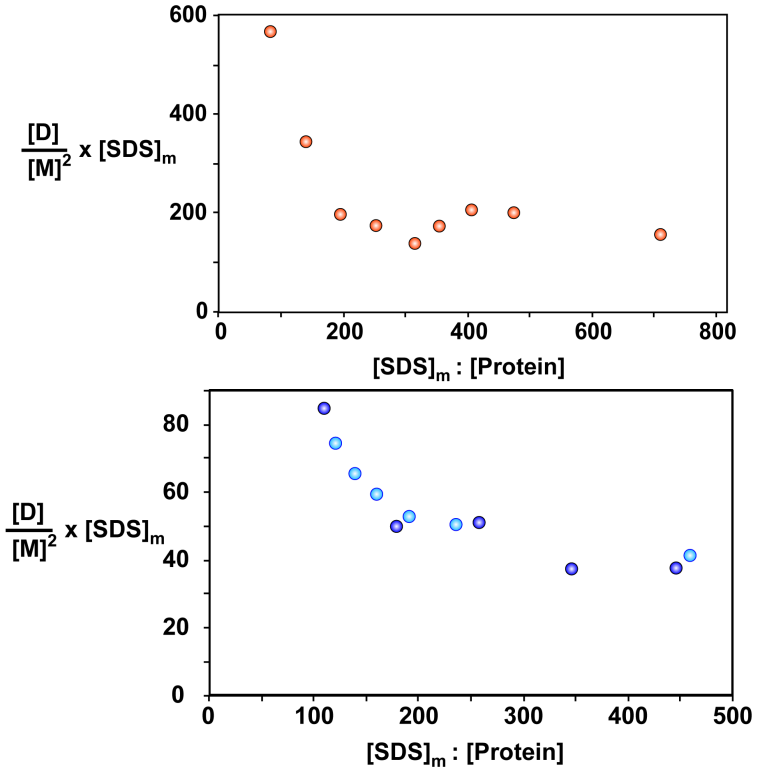

Figure 3.13 The detergent concentration-normalized dimerization affinity ([D]/[M]^2*[micellar SDS]) plotted as a function of micellar SDS to protein ratios for MCP_{TM} (top) and, full-length MCP (bottom). The two shades of blue for full-length MCP (dark and light blue) correspond to two different NMR samples.

The change in the detergent-normalized association constant with detergent:protein ratios indicates that the micelle solution is not behaving like an ideal solution under these conditions. The magnitude of the deviation from ideality can be quantified by an activity coefficient γ, that relates the apparent free energy to the standard free energy of interaction by:

$$\Delta G^0 = \Delta G^0{}_{\text{app}} - \gamma RT \ln[\text{SDS}]_{\text{micelle}} \qquad (2)$$

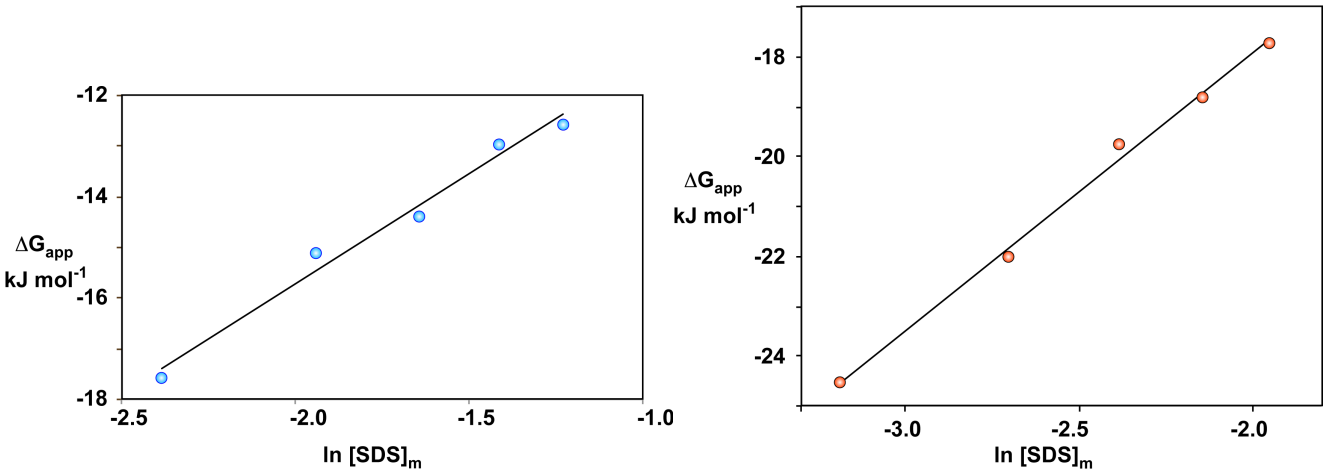

Figure 3.14 Calculated apparent free energy values for different micellar concentrations of SDS for (A) full-length MCP and (B) the TM peptide. From the slope of each fit γ was calculated by equation 2 and found to be 1.7 for the full-length and 2.2 for the transmembrane peptide.

According to this relationship, a plot of the apparent association constant versus the natural logarithm of the micellar SDS concentration ($[SDS]_m$) will be linear, with a slope that can be used to determine γ. As seen in Figure 3.14, this was the case for both MCP and MCP_{TM}, giving γ values of 1.7 and 2.2 respectively. This suggests that the continuum model does not hold under these conditions of lower SDS concentrations, and that these points should not be used in the standard free energy of association calculations. Omitting these points, the standard free energy of interaction was found to become -2.4 (± 0.1) kcal/mol for MCP and -3.2 (± 0.1) kcal/mol for MCP_{TM}, which was only slightly lower than the values obtained when all points were used. According to these calculations, under conditions where SDS behaves like an ideal solvent in the continuum model, the free energy of association between MCP decreases by 0.8 (± 0.1) kcal/mol when the amphipathic helix is removed.

Although detergent micelle systems such as SDS are often characterized as poor membrane mimetic systems, non-ideal solution behavior has also been observed for the

ErbB4 TM segment in lipid bicelles. The fact that similar behavior was seen for MCP in SDS indicates that while the shortcomings of the bicellar system are also seen in the micelle, it is still possible to find conditions where the monomer-dimer equilibrium of the TM helix interaction can be monitored and used to extract standard free energies. For the dimerization of MCP there has been debate about the impact of the membrane mimetic system on the structure of MCP, since micelles have a strong curvature that can effect the structure of the protein (Vos et al., 2009). This was the case for the HIV-1 Env peptide in DHPC micelles, where micelle stress induced curvature of the surface-bound peptide (Chou et al., 2011). Here, however, we show that the dimerization of MCP in SDS micelles is comparable to TM helix interactions characterized in bicelles. Therefore, by ensuring that SDS:protein ratios are high enough to preserve ideal solution behavior in SDS, it is possible to quantitate interaction energies in this detergent system, allowing comparison of free energy values for full-length versus TM peptide variants of MCP in our study.

3.3.7 Determination of MCP protein-detergent complex size

A potential cause for the non-ideal behavior observed for MCP in SDS micelles is a change in the size of the protein-detergent complex at the lower detergent:protein ratios. To determine whether the size of the protein-detergent complex (PDC) is the source of the non-ideality at lower protein:detergent ratios, we measured translational diffusion coefficients of the PDC by solution NMR. For these experiments it was necessary to simplify the spectra to allow peaks from a monomeric species to be distinguished from those of the dimer in a 1D spectrum. Inspection of the 2D spectra obtained with uniformly ^{15}N-labeled MCP indicated that Val33 and Ala35 were good potential

candidates to monitor in 1D spectra using selectively labeled samples, since the monomer and dimer species had large differences in ^{1}H chemical shifts (Figure 3.15).

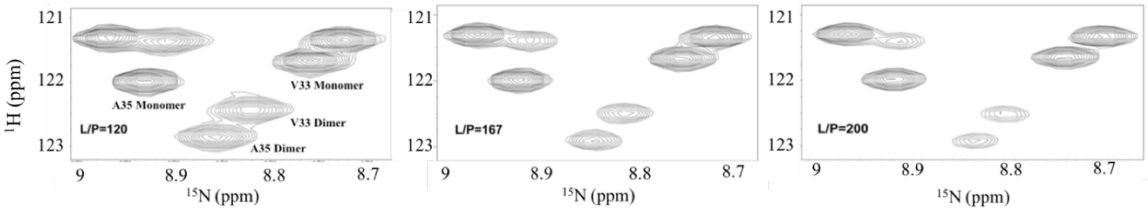

Figure 3.15 Expanded section of ^{1}H-^{15}N HSQC spectra of MCP acquired with increasing molar ratios of detergent to protein. The monomeric and dimeric peaks of Val33 and Ala35 are labeled in the first panel, and decline in intensity of the monomeric peaks can be seen with an increase in SDS:MCP ratios in the following panels.

For this purpose, selectively ^{15}N-alanine or ^{15}N-valine labeled M13 bacteriophage was propagated and MCP was isolated. As shown in 2D ^{1}H-^{15}N HSQC spectra on these samples run under conditions promoting monomeric MCP, only peaks corresponding to the labeled amino acid are observed, confirming the success of the selective labeling strategy (Figure 3.16).

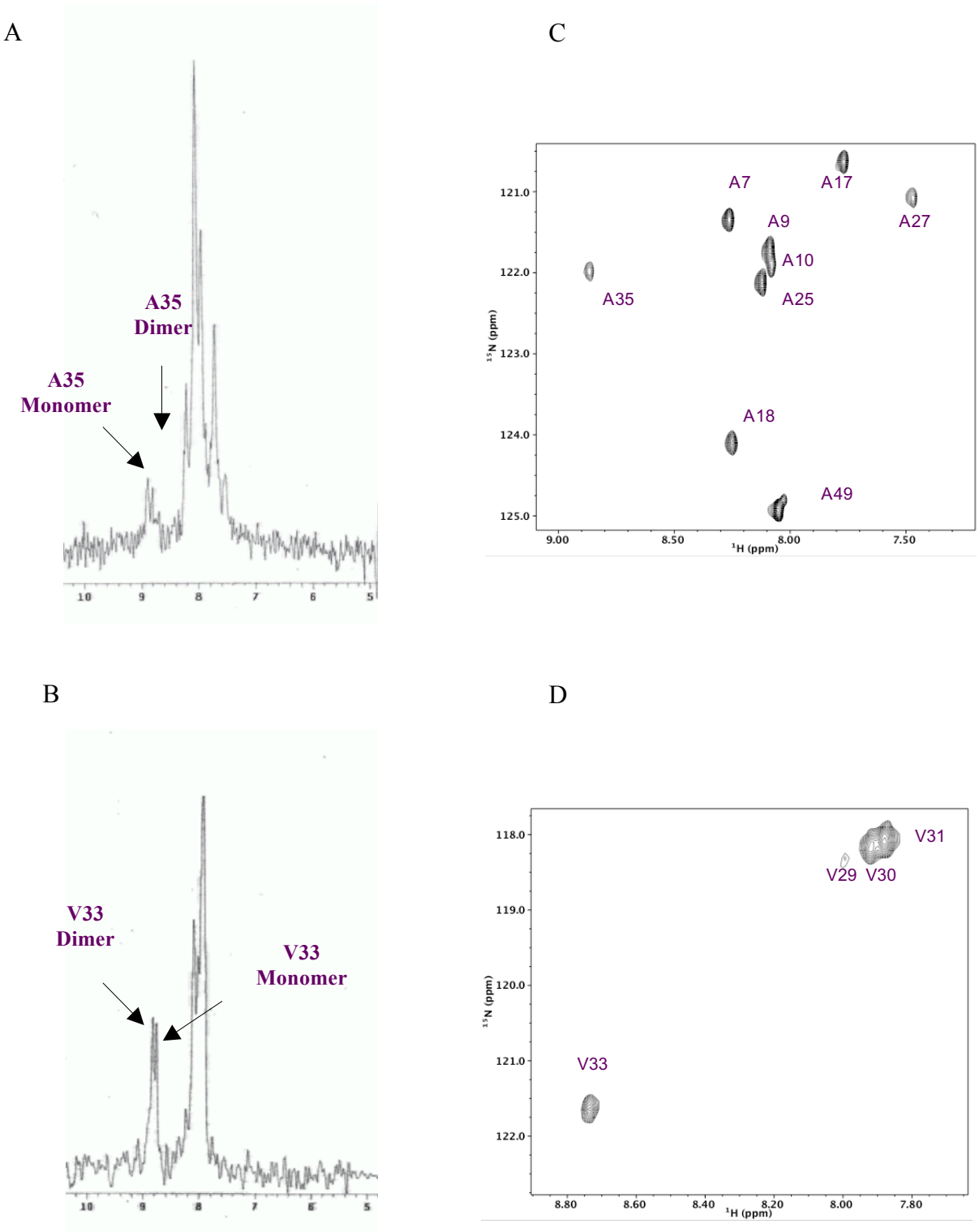

Figure 3.16 1D ^{1}H spectra of MCP using (A) 0.9 mM ^{15}N-alanine labeled MCP and (B) 0.7 mM ^{15}N-valine labeled MCP in 100 mM SDS, 10% D_2O at pH 5.5. 2D ^{1}H-^{15}N HSQC spectra showing the monomeric forms of MCP in (C) 0.2 mM ^{15}N-alanine labeled MCP and (D) 0.3 mM ^{15}N-valine labeled MCP in 100 mM SDS, 10% D_2O at pH 5.5.

Based on chemical shifts, and the absence of overlapping resonances from other amino acids, it was possible to assign the A35 and V33 peaks in the ^{15}N-edited 1D ^{1}H spectra for the monomeric and dimeric forms of full-length MCP (Figure 3.16). Samples were prepared with protein:detergent molar ratios that favoured monomeric or dimeric MCP, and ^{15}N-edited pulse field gradient echo experiments were performed to measure translational diffusion coefficients. Examples of the decay of the NMR signal with increasing gradient strength is shown for a sample prepared under conditions that favor the dimeric state in Figure 3.17.

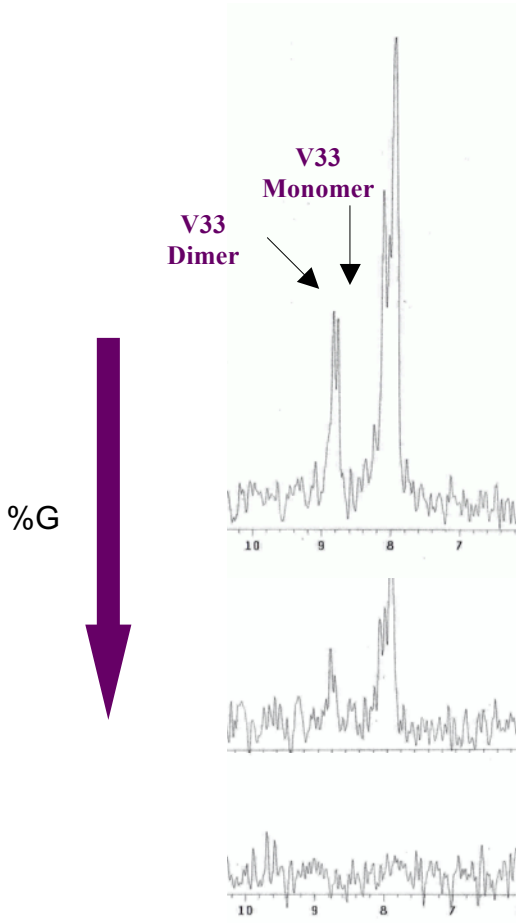

Figure 3.17 1D spectra from NMR diffusion experiments run on a 1 mM ^{15}N-valine labeled sample (100 mM SDS, 10% D_2O, pH 5.5). The spectra for increasing gradients strengths (%G) are shown with 20% (top), 40% (middle) and 90% (bottom).

The peak intensities were measured at a range of gradient field strengths and plotted with respect to the fraction gradient strength to obtain the value of the translation diffusion coefficient (D_s) (Figure 3.18).

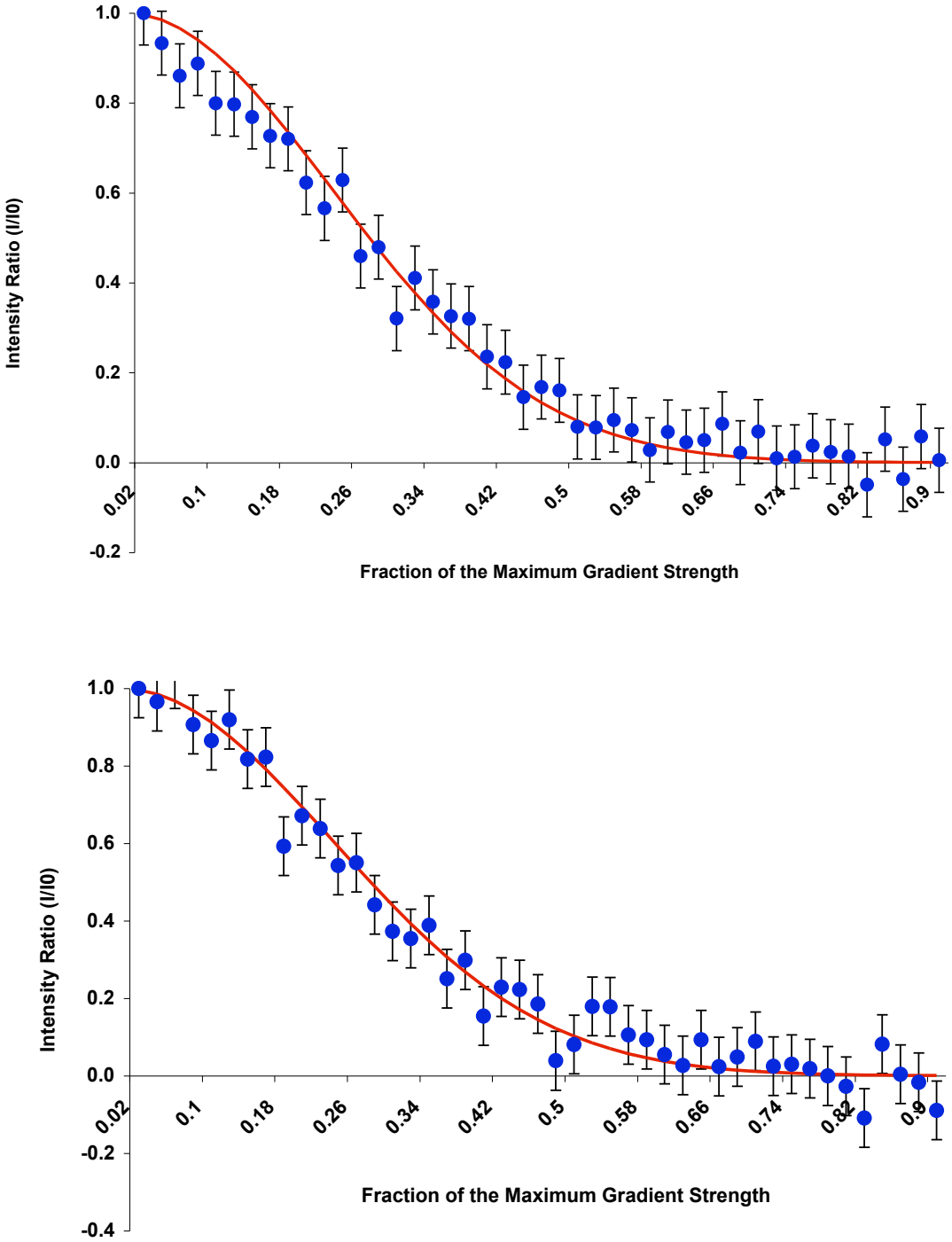

Figure 3.18 Normalized peak intensity versus the percent gradient strength for the dimeric (top) and monomeric (bottom) peak of Val33 in a 1 mM ^{15}N-valine labeled MCP, 100 mM SDS and pH 5.5, 37°C. The red line corresponds to the best fit to equation $I = I_0 \exp[-(\gamma\delta G)^2(\Delta - \delta/3)D_s]$ to solve for the translational diffusion coefficients. For the monomeric form D_s= 1.20x10^{-6} cm^2/s and for the dimeric form D_s=1.10x10^{-6} cm^2/s. Error bars represent the average standard of error of the mean of duplicate measurements on the same sample.

The results of the diffusion coefficient measurement for the monomeric (D_s= $1.20x10^{-6}$ cm^2/s) and dimeric (D_s=$1.15x10^{-6}$ cm^2/s) species in the ^{15}N-valine labeled samples were the same within experimental error, indicating that the size of the protein-detergent complex is similar between the two states. To corroborate these results the concentration of the sample was diluted down to 0.3 mM, with the concentration of SDS remaining constant at 100 mM, to give a protein:detergent ratio that favored the monomeric state. The value of the translational self-diffusion coefficient of the monomeric form was $1.15x10^{-6}$ cm^2/s which is comparable to that obtained in the sample that contained both states.

The results of the same analysis for ^{15}N-alanine labeled MCP, along with the results from ^{15}N-valine labeled MCP are summarized in Table 3.1.

^{15}N – Valine			**^{15}N – Alanine**		
[MCP] (mM)	Average D_s ± SD (cm^2/s)		[MCP] (mM)	Average D_s ± SD (cm^2/s)	
	Monomer	Dimer		Monomer	Dimer
0.7 mM	$1.20\pm0.09x10^{-6}$	$1.15\pm0.05x10^{-6}$	1 mM	$1.2\pm0.3x10^{-6}$	$0.98\pm0.05x10^{-6}$
0.3 mM	$1.2\pm(0.1)x10^{-6}$	N/A	0.2 mM	$1.29\pm0.05x10^{-6}$	N/A

Table 3.1 Summary of diffusion experiments on ^{15}N-alanine and ^{15}N-valine MCP samples. All experiments were carried out in 100 mM SDS, 10% D_2O at a pH of 5.5

These results show that for both the selectively ^{15}N-alanine and ^{15}N-valine labeled samples, there was a minimal effect on the translational diffusion of the complex in the sample. The rate of diffusion of the monomer PDC for selectively ^{15}N-alanine labeled MCP was in line with the rate of diffusion of the monomeric and dimeric forms of selectively ^{15}N-valine labeled MCP indicating there was no change in PDC size when MCP is in its monomeric or dimeric form.

3.3.8 Kinetic mechanism for MCP self-association in micelles

Since distinct peaks can be observed for monomeric and dimeric states of MCP in the NMR spectrum, the rate of exchange between these states is slow on the NMR time scale. This slow-exchange may be a factor that can help explain the non-ideal behavior observed for MCP self-association in SDS. Based on the smallest chemical shift difference measured between the two states, the rate constant for exchange can be estimated to be slower than 25 s^{-1}, with some coalescence being observed for peaks that were separated by ~10 Hz. In addition, when 1D spectra of ^{15}N-labeled Ala MCP were recorded at 45°C to 50°C, peak broadening was observed, suggesting that exchange rates approach shift differences at the higher temperatures (Figure 3.19). Comparison of the line width of the Ala35 monomer peak in spectra acquired in conditions favoring the dimeric state versus those favoring the monomer also suggest a dimerization rate constant (k_{dim}) of ~10 s^{-1}, in line with estimates from the amide shift differences.

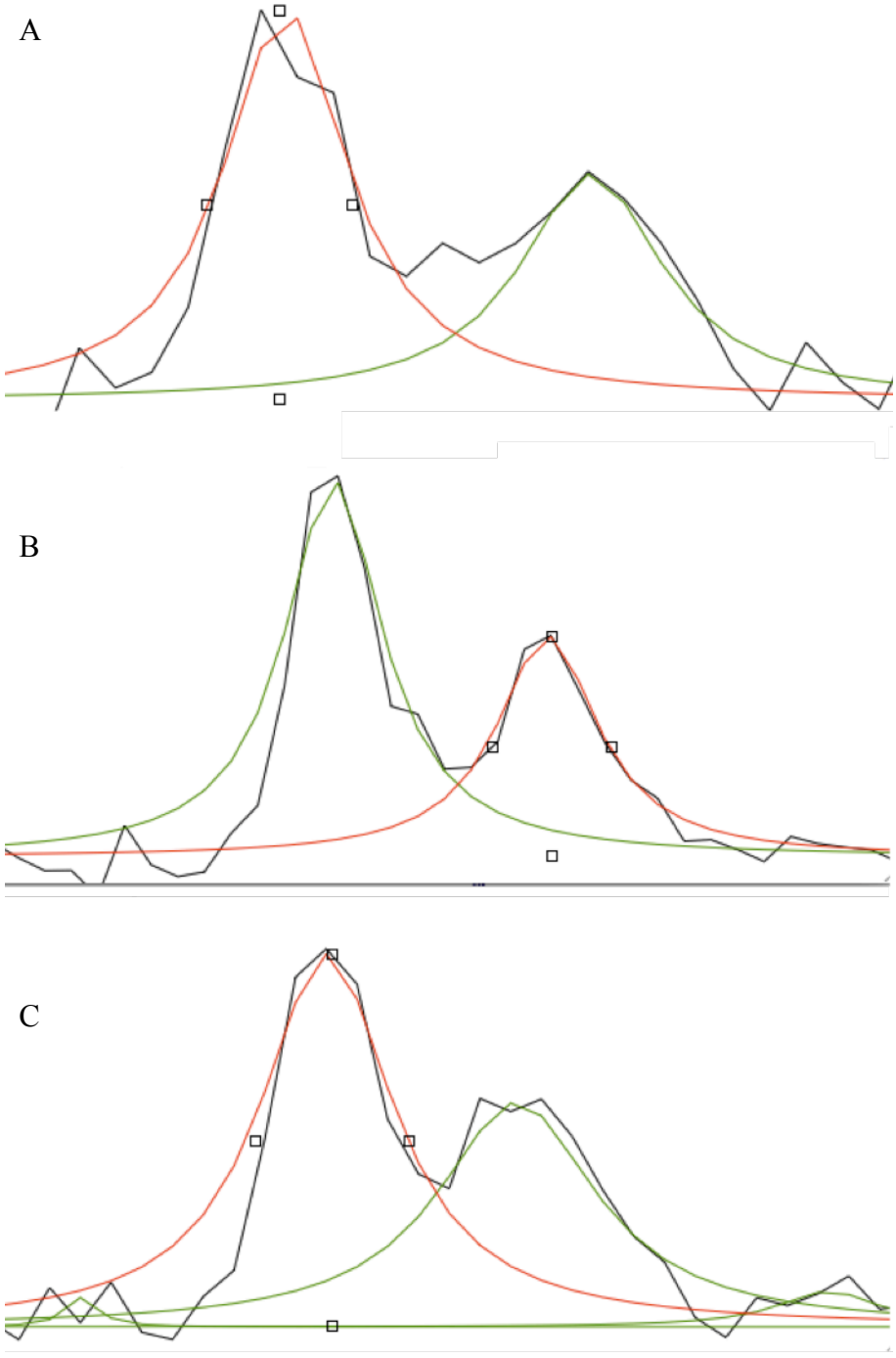

Figure 3.19 Deconvolution of the 1D ^{1}H-^{15}N HSQC spectrum of selectively ^{15}N-alanine labeled MCP at (A) 40°C (B) 45°C and (C) 50°C using a Lorentzian fit in iNMR. The sample contained 0.8 mM ^{15}N-alanine labeled MCP, 100 mM SDS, 10% D_2O at a pH of 5.5.

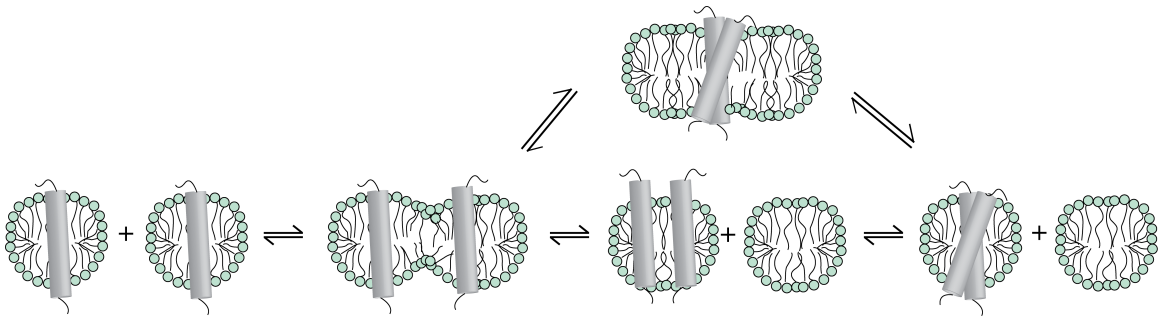

Figure 3.20 Hypothetical kinetic mechanisms proposed for of polypeptide transfer and dimerization in SDS micelles. If micelle fusion/fission events are slow relative to monomer-dimer exchange rates, then only the first state on the left (two PDCs, each containing a single polypeptide) will appear as a monomeric species in the NMR spectrum.

To understand the origin of slow exchange kinetics in a weak interaction, and how it might provide insight into the non-ideal behavior observed for the MCP interaction in SDS micelles, it is useful to consider the kinetic mechanism involved in the process of dimer association and dissociation. While the series of kinetically resolvable steps that are involved in dimerization of micelle-solubilized peptides is not known, it is likely that this would involve collisions between micelles. A fraction of these collisions would give rise to fusion of two micelles, followed by a dissociation of detergent into two micelles, either in a single step as shown in the scheme in Figure 3.20, or slowly through gradual loss of detergent to the solution. In cases where micelle fusion leads to the creation of a PDC containing two polypeptides, dimer formation between the two TM segments could proceed.

While the NMR spectra reveal a macroscopic exchange rate on the order of ~10 s^{-1}, this reflects the rate of the slowest step in the kinetic mechanism, which could be the step that transfers two MCP polypeptides into a single PDC, rather than the rate of polypeptide association. It is possible that, once encapsulated in a single micelle, the rate

of exchange between monomeric and dimeric MCP is actually fast on the NMR time-scale. In this case, the species that we have assigned as dimeric in the NMR spectrum would actually be a mixture of monomer and dimer, with its chemical shift appearing at the population-weighted average between the two states. This would give rise to an overestimation of the population of dimeric MCP, and therefore overestimate the affinity of dimerization. Moreover, this would suggest that the only way that a monomeric state could be observed is through the physical separation of a PDC containing 2 polypeptides into a 2 PDCs with a single polypeptide. Although the mechanism for this type of PDC fission is not known, one possibility is that it requires collision with an empty micelle. In this case, the apparent affinity of dimerization would be affected by the concentration of 'empty' polypeptide-free micelles, which would become limiting as protein:detergent ratios increase to saturating levels.

3.3.9 SDS micelles as a participant in the TM dimer dissociation reaction

If the empty micelle (Mic_{empty}) is an active participant in apparent dimer dissociation, then the calculation of the standard free energy of dimerization must be based on the following reaction:

$$2M \rightleftharpoons D + Mic_{empty} \quad (3)$$

As has been suggested in similar models (Mineev et al., 2014), the amount of detergent that is associated with monomeric and dimeric PDCs may differ, and may also be larger or smaller than the number of detergent molecules contained in an empty micelle. In this case, the amount of detergent that is released upon dimerization will not be the same as

the amount of detergent required to make up the empty micelle, which can be reflected by a modified equilibrium:

$$2M \rightleftharpoons D + Mic_{empty} + fMic_{empty} \qquad (4)$$

where f is the fraction of empty micelle aggregation number (e):

$$f = (2m - d - e)/e \qquad (5)$$

and where m and d are the number of detergent molecules that are associated with the PDC of the monomeric and dimeric forms. Note that in the case where additional detergent is required to form a complete empty micelle upon dimerization, the value of f would be negative.

Using this new equilibrium system, the constant for association becomes:

$$K_a = \frac{[D][Mic_{empty}]^{1+f}}{[M]^2} \qquad (6)$$

This can be converted into the standard free energy of interaction to give:

$$\begin{aligned} \Delta G^{\circ}_{micelle} &= -RT \ln K_a \\ &= -RT \ln K_{app} - (1+f)RT \ln[Mic_{empty}] \end{aligned} \qquad (7)$$

According to this equation, $\ln K_{app}$ should vary linearly with $\ln[Mic_{empty}]$ over a range of detergent concentrations, with the slope corresponding to $-(1+f)RT$ and intercept $\Delta G^{\circ}_{micelle}$ that corresponds to the standard free energy of the TM helix interaction at an empty micelle concentration of 1 M. Based on this analysis it is also possible to extract

the aggregation numbers *e, m* and *d* since the concentration of empty micelles is dependent on these values.

To see if MCP in SDS followed this model of self-interaction, we applied this analysis to MCP and MCP_{TM}. As shown in Figure 3.21, there was a linear relationship between the apparent free energy and ln [Mic_{empty}] in both cases. A similar relationship has also been observed for self-association of the TM segment from the fibroblast growth factor receptor 3 (FGFR3) in a 9:1 mixture of DPC:SDS, and the vascular endothelium growth factor receptor 2 (VEGFR2) in DPC (Mineev et al, 2014), suggesting that this is a common feature of TM-segment interactions in micelles.

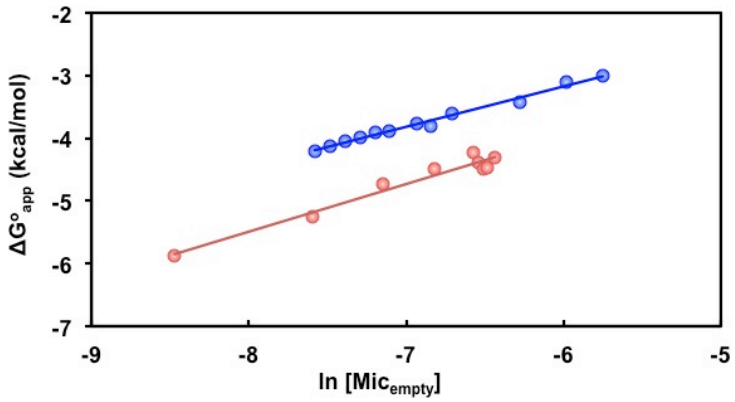

Figure 3.21 Apparent free energy of dimerization (ΔG^0_{app}) of MCP (blue) and MCP_{TM} (red) against ln[Mic_{empty}].

According to this analysis, the number of detergent molecules that make up the PDC when the protein is in its monomeric form versus its dimeric form is approximately the same for full-length MCP (Table 3.2), in line with our diffusion coefficient results. In contrast, the TM peptide showed differences in SDS aggregation numbers for the 2 PDCs, with a larger number of detergent molecules for PDCs containing just one peptide.

Sample	N_e	N_m	N_d	f	ΔG^o (kcal/mol)
MCP	78.4 ± 0.6	79.7 ± 0.5	77.2 ± 0.1	0.05 ± 0.01	0.70 ± 0.05
MCP_{TM}	78 ± 1	88 ± 1	80 ± 1	0.23 ± 0.03	0.6 ± 0.1
$MCP_{TM\text{ - }M28L/V31L}$	75 ± 1	89 ± 1	78 ± 1	0.33 ± 0.03	0.3 ± 0.1

Table 3.2 Summary of the calculated values of number of detergent molecules in empty micelles (N_e), micelles containing monomeric MCP (N_m), micelles containing dimeric MCP (N_d), fraction of empty micelles (f) and the standard free energy of interaction (ΔG^o) for MCP, MCP_{TM} and $MCP_{TM\text{-}M28L/V31L}$.

This change in aggregation number that occurs in the MCP_{TM} system indicates that there is an increased influence of the detergent on the equilibrium when the amphipathic helix is absent. To see if this was a general property of the MCP_{TM} peptide, the same analysis was carried out on data previously acquired on $MCP_{TM\text{-}M28/V31L}$, which is the same peptide as the MCP_{TM} peptide used in this thesis, except for the presence of two mutations that increase the affinity of TM helix interactions. In this system there is a similar increase in the SDS aggregation number in the monomer PDC relative to the dimeric PDC and empty micelle, indicating that differential detergent-binding is a property of the MCP_{TM} peptide. These differences in detergent binding properties indicate that the N-terminal amphipathic helix might play a role in determining the detergent-binding properties in PDC complexes.

This analysis also allowed us to estimate the values of the apparent free energy of association for MCP and MCP_{TM}, standardized to a concentration of free micelles of 1 M. As shown in Table 3.2 all the energies obtained were small and positive. Although it may seem counterintuitive for a specific interaction to have a positive free energy of association, this is a manifestation of the different standard state that was used with

respect to detergent, which is a standard concentration of 1M empty SDS micelle. Nonetheless, comparison of these free energies allow the affinity of TM helix interactions to be compared after removing contributions to the equilibrium that arise from differential detergent binding between monomer and dimer states. According to this analysis, the free energy of TM helix association is the same for both MCP and MCP_{TM} indicating that the difference in affinity measured using the continuum model can be attributed exclusively to differences in detergent-binding properties between the two systems. In contrast, the standard free energy of association for $MCP_{TM\ -\ M28L/V31L}$ was lower than the WT peptide, reflecting the fact that there is an increase the affinity of interaction, and that this increase comes about due to changes in the interactions between TM segments, and not detergent-binding properties.

The fact that our data could be fit to a model that assumes that the empty micelle is a reactant in dimer dissociation suggests that the rate determining step is micelle fusion/fission. However, the actual rates for SDS micelle fusion and fission in our system are not known, although they are predicted to be slow based on the electrostatic repulsion that is anticipated to occur between SDS headgroups of colliding micelles. The only measurement that approaches this comes from an experiment measuring the rate of fluorescence decay in a concentrated pyrene-labeled triglyceride diluted into SDS (Rharbi & Winnik, 2001). In these experiments the rate of SDS micelle fission at 23^{o}C was found to be very slow (~0.01 s^{-1} in 100 mM NaCl), and micelle fusion to be even slower. While this is much slower than the exchange rate measured for MCP in SDS micelles, it is possible that micelle fusion/fission rates in our system were accelerated by the incorporation of the MCP polypeptide into the micelle.

3.3.10 Investigation of the monomer-dimer equilibrium in DPC micelles

Following our characterization of MCP dimerization in SDS micelles, we wanted to investigate the properties of the monomer-dimer equilibrium of MCP in a different micelle system to determine the effect of the detergent on the self-interaction. A mild detergent with a polar headgroup was tested first, specifically the alkylmaltosides decylmaltoside (DM) and dodecylmaltoside (DDM).

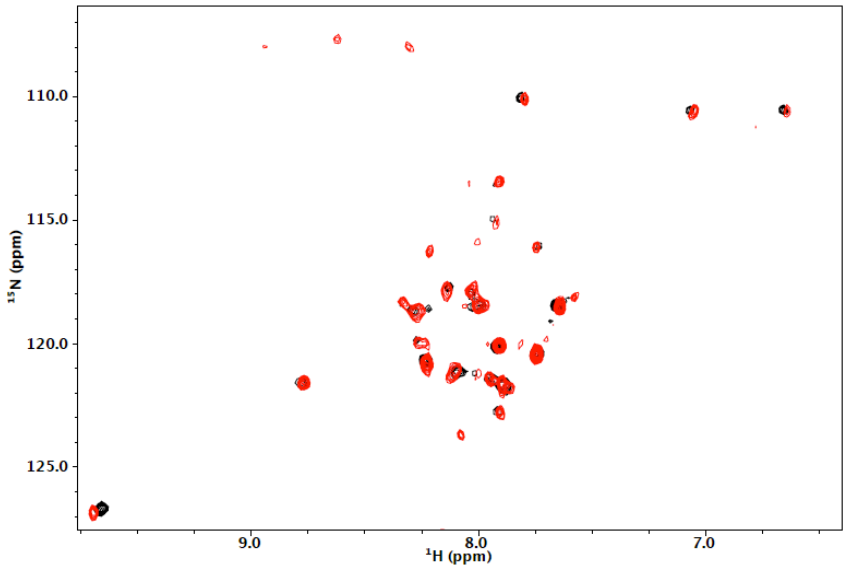

Figure 3.22 ^{1}H-^{15}N HSQC spectra obtained for 0.1 mM MCP in 50 mM DM (black) and 10% D_2O and 0.1 mM MCP in 500 mM DDM (red) and 10% D_2O at a pH of 5.5 at 40°C.

As shown in Figure 3.22, similar HSQC spectra were obtained for both detergents. In these samples the MCP concentrations was kept low to favor the monomeric form. In these spectra there were a small number of peaks compared to the number expected, with approximately 35 % of the expected backbone resonances being

observed. The small number of peaks, combined with the poor spectral quality, and low solubility in these detergents, suggested that these detergents would not be appropriate for a study of TM-helix interactions, and therefore we chose not to continue with these detergents.

We also investigated dimerization in the zwitterionic detergent dodecylphosphocholine (DPC), since it is usually described as more suitable than SDS due to its zwitterionic phosphocholine head group, which mimics lipids in the membrane better than the anionic sulfate head group of SDS. The secondary structure of MCP has also been studied in DPC, and spin-labeled micelles used to determine the position of MCP in these micelles. To study the dimer in this detergent, the same titration experiments that were performed in SDS were carried out and monitored by solution-state NMR in DPC.

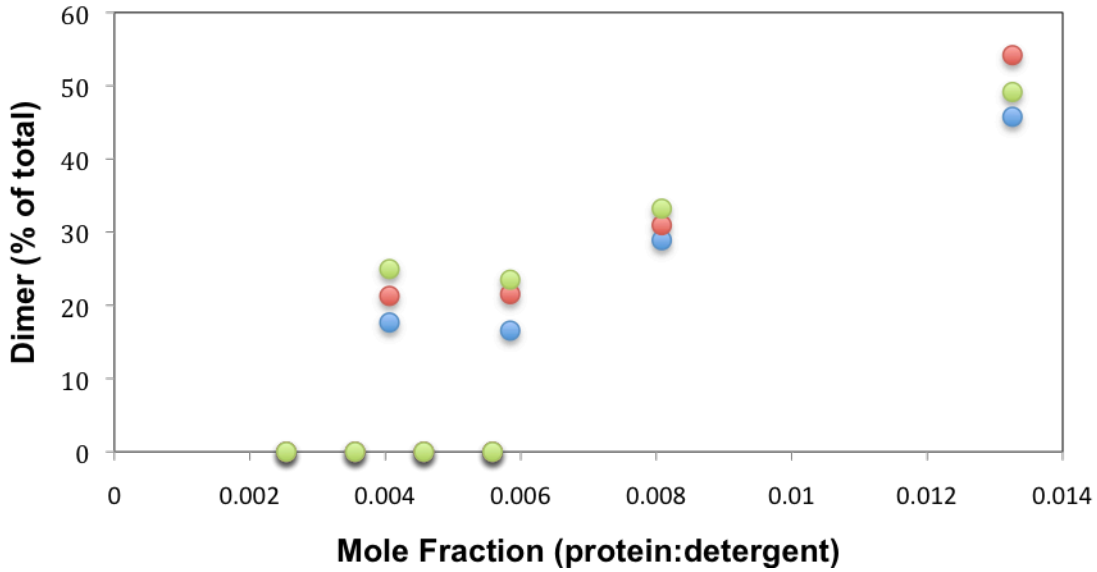

Figure 3.23 Fraction of the protein population that is dimeric at various molar ratios of MCP to DPC based on monomer and dimer peaks of Gly38 (blue), Val33 (red) and Ala35 (green).

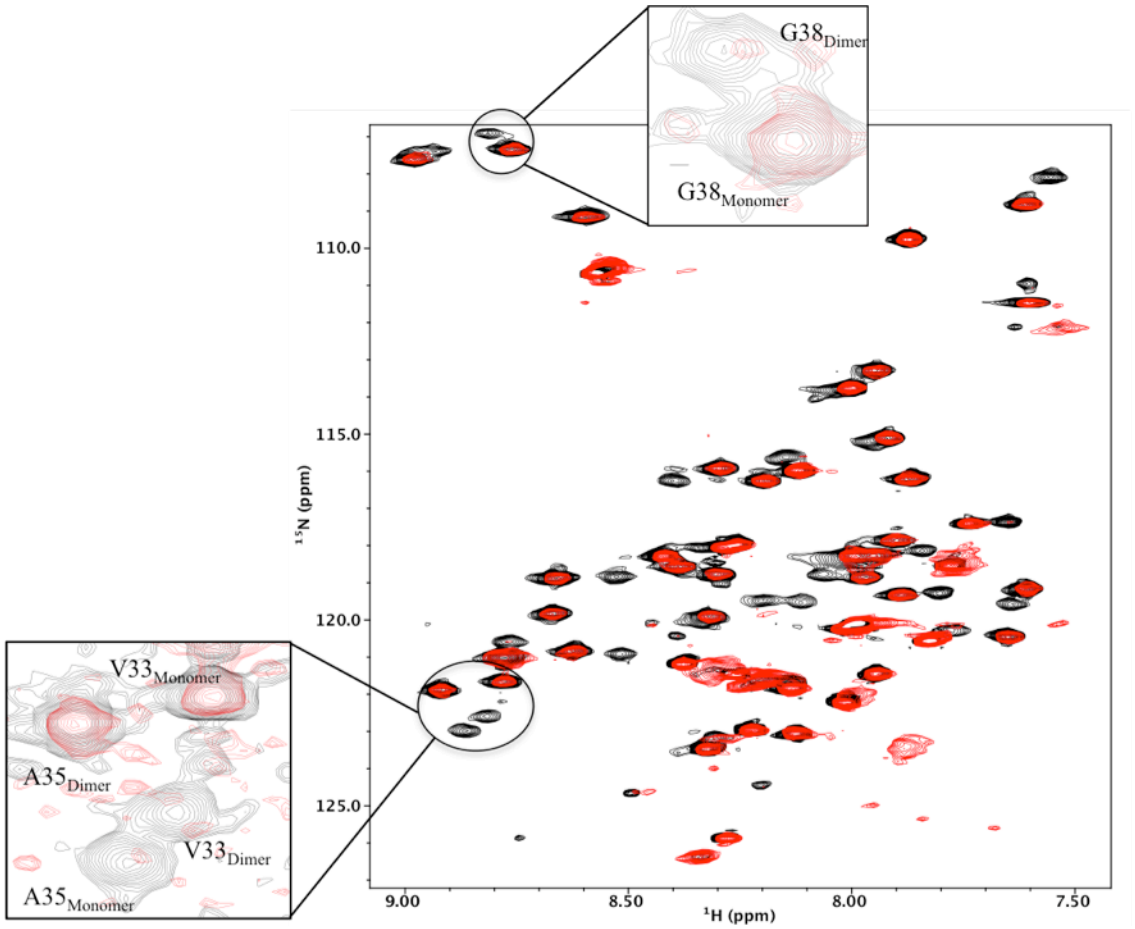

Figure 3.24 ^{1}H-^{15}N HSQC of MCP with 0.6 mM (black) and 0.5 mM (red) MCP in 100 mM DPC and 10% D_2O at a pH of 5.5. Monomer and dimer peaks corresponding to V33, A35 and G38 are circled and zoomed in at lower levels.

As shown in Figure 3.24, the ^{1}H-^{15}N HSQC spectrum of 0.6 mM MCP in 100 mM DPC showed good chemical shift dispersion similar to what has previously been reported (Papavoine et al., 1995). The spectra reported in the literature was for the monomeric form of MCP in DPC micelles and our starting conditions were selected to favour the dimeric species. Although chemical shift assignment experiments were not carried out, comparison to the published monomeric spectra for MCP in DPC and our SDS data suggested peaks for the dimeric state could be tentatively assigned for Val33, Ala35 and

Gly38. However, when we monitored the relative population of these two species at a range of protein to detergent molar ratios, there was no visible trend that would be consistent with a monomer-dimer equilibrium (Fig. 3.23). In some cases, no dimer species could be detected, even though other experiments performed at similar protein:detergent ratios did give rise to some dimer population (Fig. 3.24). The poor reproducibility of these experiments suggested some irreversibility to the oligomerization that we were observing suggesting that it would not be possible to extract thermodynamic data for MCP dimerization in DPC micelles. It is possible that differences in the arrangement of DPC detergent molecules around MCP compared to interactions between SDS and MCP (Papavoine et al., 1995) may be responsible for the poor ability of DPC to maintain an equilibrium state between monomeric and dimeric MCP.

3.4 Summary

In this study a number of fundamental insights into the nature of TM helix interactions studied in micelles were made. The main accomplishments of this work include:

- Development of an alternate purification technique to extract MCP from M13 bacteriophage.
- Determination of backbone chemical shift assignments for the dimeric form of MCP, allowing secondary shift analysis to determine that the secondary structure of monomeric and dimeric were the same. This was further confirmed by circular dichroism studies.
- Calculation of the standard free energy of interaction of MCP and MCP_{TM} in SDS micelles at high detergent:protein molar ratios using the continuum solvent

model. We found that removal of the N-terminal amphipathic helix increased the strength of interaction by 0.8 kcal/mol.

- Determination that SDS is a non-ideal solvent at lower detergent to protein molar ratios in the study of self-association by MCP and MCP_{TM}. Translational diffusion coefficient measurements indicated that large changes in the size of the PDC is not the origin of this non-ideal behaviour.

- Demonstration that empty SDS micelles must be explicitly included in the free energy of TM helix association for MCP. This analysis showed that the differences in affinity that were determined in the continuum model between different MCP constructs can be accounted for by differences in the detergent binding properties of monomeric and dimeric MCP and MCP_{TM}.

- DPC and alkylmaltosides were found to be poor membrane mimetic systems to study the thermodynamics of MCP dimerization.

Chapter 4: Purification and preliminary characterization of the p7 protein of the hepatitis C virus

4.1 Introduction

4.1.1 Hepatitis C

Hepatitis C is an infection that affects 170 million people worldwide, and approximately 250,000 people in Canada ("Canadian Liver Foundation - Hepatitis C," "World Health Organizaition - Hepatitis C,"). The infection is caused by the hepatitis C virus (HCV) that is typically transmitted via contact with blood of an infected individual, and targets mainly the liver, and in some cases kidney and skin cells (Galossi et al., 2007). Infection results when the virus enters the bloodstream, which can occur via drug use, improperly sterilized medical equipment and tainted blood samples from transfusions (Pawlotsky, 2004; Simmonds, 2004). The infection is considered to be chronic when the virus persists in the bloodstream for longer than six weeks, which can cause moderate to extreme damage to the liver, and ultimately lead to cirrhosis, liver failure and liver cancer (Galossi et al., 2007; Stauber & Stadlbauer, 2006).

4.1.1.1 HCV and current treatments

Some treatments are available for hepatitis C, but these are more commonly utilized in developed countries due to the high cost of these medications. A treatment is considered to be curative if HCV RNA is absent from the blood serum for at least six months after the completion of the treatment (Hoofnagle & Seeff, 2006). The most common treatment combines pegylated-interferon-α and ribavarin (Feld & Hoofnagle, 2005; Pawlotsky, 2004; Shepard et al., 2005). Interferons (IFN) function by activating genes that are interferon-stimulated, which leads to the production of proteins that act

against viral infection, such as nucleases that target double-stranded RNA and inhibitors of viral protein translation (Guo et al., 2004; Hoofnagle & Seeff, 2006). Ribavarin is a guanosine analogue that has previously been shown to have activity against RNA and DNA viruses (Hayashi & Takehara, 2006). Its mode of action is not clear, but it might be due to the incorporation of ribavarin into the HCV RNA, which inhibits its replication (Feld & Hoofnagle, 2005). These treatments are not effective in all patients, with approximately 50% of patients on these treatments being cured, as defined by the absence of viral particles in blood serum (Feld & Hoofnagle, 2005). There are also difficult side effects associated with these drugs, such as IFN-induced depression and ribavirin-induced anemia, which can make it challenging for a patient to adhere to protocols for complete treatment (Horsmans, 2005; Russmann et al.,, 2006).

More recently, new treatments have been implemented which have been classified as direct acting antiviral agents (DAA) (Poordad & Dieterich, 2012; Schinazi et al., 2014). These proteins target specific steps in the viral life cycle and consist of four classes: nonstructural proteins 3/4A (NS3/4A) protease inhibitors (PIs), NS5B nucleoside polymerase inhibitors (NPIs), NS5B non-nucleoside polymerase inhibitors (NNPIs), and NS5A inhibitors (Schinazi et al., 2014; Welsch et al., 2012).

4.1.2 HCV p7 protein

The HCV genome consists of a 9.6 kb single-stranded positive sense RNA, which encodes a polyprotein precursor protein that is 3000 amino acids in length. This polyprotein undergoes co- and post-translational modification by host and viral proteases which produces 10 proteins (Cocquerel et al., 2006; Penin, 2003). At the N-terminus

there are three proteins that are released by cellular peptidases in the endoplasmic reticulum (ER): the core protein, and the glycoproteins E1 and 2, all of which make up the structure of the virus (Francesco, 1999). At the C-terminal end there are six non-structural (NS) proteins (NS2, NS3, NS4A, NS4B, NS5A, NS5B). NS2 and NS3 are proteases required for processing of the remainder of the polyprotein (Francesco, 1999). NS4A is thought to be a cofactor of NS3 and the NS3-NS4A complex has been shown to be necessary to carry out the protease activity of NS3 (Hundt et al., 2013). NS5B has also been well studied, and based on structural and biochemical studies it was shown to be responsible for the replication of HCV RNA (Behrens et al., 1996; Bressanelli et al., 1999; Herian et al., 1997; Lesburg et al., 1999). While not much is known about the functions of NS4B, some studies have pointed towards its involvement in altering the membrane properties of the ER, which may be necessary for HCV replication (Egger et al., 2002; Harboring et al., 2003; Hundt et al., 2013).

Located between the structural and non-structural proteins is the p7 protein, which was the target of study in this chapter. p7 is an integral membrane protein that is 63 amino acids in length and localizes to the ER membrane (Carrère-Kremer et al., 2002; Haqshenas et al., 2007). The topology of p7, predicted by hydropathy and confirmed by antibody studies with Myc tagged p7, contains two transmembrane α-helices, with N- and C-termini on the cytosolic side of the ER membrane (Carrère-Kremer et al., 2002; Lin et al., 1994). It is highly conserved across the seven different genotypes (and subgroups) of HCV, which has highlighted the importance of understanding p7 structure and function in the development of new treatments against HCV (Kuiken & Simmonds, 2009).

p7 was initially not considered to be an independent protein since it was first found to be expressed as a fusion with the structural glycoprotein E2. It was found that the E2 glycoprotein could be expressed in the absence of the p7 protein by C-terminal truncation studies (Lin et al., 1994) identifying p7 as an independent protein. Following this discovery, there has been renewed interest in the structural and functional properties of p7. It has since been shown to be important in viral infectivity and accumulation. While deletion of p7 from the HCV genome does not inhibit replication of the viral genome (Lohmann et al., 1999), detectable levels of infectious HCV was reduced *in vitro*. This was in line with experiments showing that chimpanzees transfected with HCV missing the gene that encodes p7 and did not acquire hepatitis C (Jones et al., 2007; Sakai et al., 2003; Steinmann et al., 2007).

4.1.3 p7 as a viroporin and potential drug target

Significant evidence has pointed towards p7 as a viroporin that forms an oligomeric state that conducts ions across the ER membrane. This was first suggested by membrane conductance measurements with purified p7 in black lipid membranes that showed cation specific channel activity (Clarke et al., 2006; Griffin et al., 2003; Pavlovic et al., 2003; Premkumar et al., 2004). Oligomerization of p7 has also been demonstrated by cross-linking studies and transmission electron microscopy (TEM) (Clarke et al., 2006), supporting the idea that p7 is a viroporin. In addition there are similarities between p7 and other proteins that have been classified as viroporins. Proteins from this class tend to be relatively small in size (most lie within the range of ~60-120 amino acids) with at least one transmembrane α-helix. p7 also possesses a short sequence of basic amino acids linking the TM domains (Fig. 4.1), and a region rich in aromatic residues within the first

transmembrane segment that have both been shown to be important for destabilization of the membrane similar to other viroporins (Gonzalez & Carrasco, 2003). For example, when W30 and Y42 were mutated into phenylalanine, a decreased viral infectivity was observed (Steinmann et al., 2007). Also, alterations to basic residues K33 and R35 in the cytoplasmic loop were shown *in vivo* to prevent viral infectivity (using double mutant K33I/R35S) and alanine mutations resulted in decreased virus production *in vitro* (Jones et al., 2007; Sakai et al., 2003) (Figure 4.1).

Figure 4.1 Amino acid sequence of HCV p7 protein. Putative TM domains are shown in red. Aromatic residues shown to be important for p7 function are shown in green. Basic residues that have been shown to impact viral infectivity and virus production are shown in blue.

While viroporin function is not well understood, many of the well-studied systems, such as Vpu from HIV and M2 from the influenza A virus, have been shown to form oligomers that make hydrophilic pores in the membrane (Agirre et al., 2002; Fischer & Sansom, 2002; L. Pinto et al., 1997). Oligomerization states seem to be protein-dependent, with Vpu forming a pentamer, while M2 favors a tetrameric state (Grice et al., 1997; Holsinger & Lamb, 1991; Pinto & Lamb, 2006; Sugrue & Hay, 1991). Similarly p7 has been shown to form a heptameric cation selective channel (Clarke et al., 2006; Premkumar et al., 2004).

4.1.4 GST-p7-His construct

In this chapter we wanted to study the structure of the monomeric form of p7 and characterize its oligomerization properties. For this purpose, we used a fusion protein of p7 with a N-terminal glutathione-S-transferase (GST) tag and C-terminal hexahistidine (His) tag previously used in our lab (Haddad, 2007). This p7 construct has been shown to be suitable for studying the functional and structural properties of p7. Purified p7-His cleaved from its GST tag has previously been studied in black lipid membranes (BLM, also referred to as painted bilayers, formed over a pinhole in hydrophobic material such as Teflon) and it was found to form ion-conducting channels (Griffin et al., 2003). This study also showed a loss of channel activity when amantadine was introduced indicating that it is likely inhibited by amantadine, as has been previously demonstrated for p7 in HCV (Griffin et al., 2003). These results indicate that the p7-His construct is functionally relevant, and appropriate for in-depth structural studies on p7.

4.1.5 Chapter Aims

In this chapter we utilized *E. coli* protein expression, protein purification (Ni-affinity chromatography, RP-HPLC) and circular dichroism (CD) techniques to perform preliminary structural studies on the HCV p7 protein. Our aim was to ultimately carry out solution-state NMR studies to determine the structure and dynamics of p7.

4.2 Materials & Methods

4.2.1 Protein expression

Expression vectors which encoded the C-terminal hexahistadine tagged construct of GST-p7 (GST-p7-His) were transformed into *E. coli* BL21 cells and grown at 37°C in M9 minimal media (0.05 M Na_2HPO_4, 0.02 M KH_2PO_4, 0.009 M NaCl, 0.3% (w/v) glucose, 1% (v/v), 0.1% (w/v) ammonium chloride Gibco MEM vitamins, 0.1 mM $CaCl_2$, 0.1 mM $MgSO_4$) with 100 μg/mL ampicillin (Haddad, 2007). Once the optical density of the culture at 600 nm (OD_{600}) reached ~0.5-0.6 (usually in about 4 hours), expression was induced with 0.5 mM isopropyl B-D-thiogalactopyranoside (IPTG). The culture was then left in the incubated shaker overnight at 37°C to allow for expression of p7, and the bacterial pellet isolated after centrifugation the following day.

4.2.2 Purification of GST-p7-His construct

4.2.2.1 Cell lysis and solubilization of GST-p7-His construct

The bacterial pellet from a 1 L culture was resuspended in 25 mL of cell lysis buffer (20 mM Tris•Cl, 500 mM NaCl, pH 8.0), transferred to a 50 mL falcon tube and sonicated on ice twice for 1 minute each time. The sample was then centrifuged for 30 minutes at 22,000g at 4°C. The pellet was suspended in 10 mL of binding buffer (8 M urea, 0.1M $NaHPO_4$ buffer) and left on rotating shaker at 4°C overnight.

4.2.2.2 Nickel affinity chromatography

For nickel affinity purification, nickel nitriloacetic acid (Ni-NTA) resin (Novagen) was regenerated prior to every purification. This was carried out by washing

10 mL of resin with 20 mL of regeneration buffer (6 M guanidinium hydrochloride (Gu-HCl), 0.2 M acetic acid), followed by 5 washes with ddH_2O (10 mL each), 10 mL of 25% ethanol, 10 mL of 50% ethanol, 10 mL of 75% ethanol, 50 mL of 100% ethanol, 10 mL of 75% ethanol, 10 mL of 50% ethanol, 10 mL of 25% ethanol, 10 mL of ddH_2O, and 50 mL of 100 mM EDTA (5x10 mL increments). The resin was then washed extensively with ddH_2O prior to application of 20 mL of 200 mM $NiSO_4$. The column was then washed with ddH_2O and 20 mL of regeneration buffer. This was usually carried out the day before purification.

After cell lysis the sample was spun for 30 minutes at 30,000 g at 4°C, and the supernatant added to the column containing 10 mL of Ni-NTA resin that had been equilibrated with 10 mL of binding buffer. The flow-through was collected and reapplied to the column one time. Five wash steps were carried out, each with 20 mL of wash buffer (8M urea, 0.1M $NaHPO_4$, 0.03M imidazole, 0.02% SDS pH 8.0). 15 mL of elute buffer (8M urea, 0.1M $NaHPO_4$, 0.4M imdiazole, 0.02% SDS, pH 8.0) was then applied to elute GST-p7-His. The elution fractions were then transferred to dialysis tubing (30 kDa molecular weight cut off) (Spectrum), which was placed in dialysis buffer containing 4 L of 0.02% SDS at pH 8.8. The dialysis was carried out overnight at 4°C.

4.2.2.3 Thrombin cleavage

The dialyzed sample was concentrated using 10 kDa MWCO ultrafiltration units (Millipore). The filter was equilibrated with 0.02% SDS prior to addition of purified GST-p7-His sample. Concentration was done at 4000 g, 4°C to a final volume of approximately 1-1.5 mL. The concentration was determined using BCA™ Protein Assay

Kit (Pierce) according the manufacturer's instructions and was typically on the order of 7-8 mg/mL. Thrombin cleavage was performed with 6 Units (unit is the amount of thrombin required to cleave leader sequence from Cleavage Control Protein, as specified by the manufacturer) of thrombin (Novagen) for every milligram of p7 fusion protein in thrombin cleavage buffer (20 mM Tris-HCl, pH 8.4, 150 mM NaCl, 2.5 mM $CaCl_2$), overnight at 4°C.

4.2.2.4 RP-HPLC separation of p7-His from GST tag

HPLC was carried out on a Waters 600E Multisolvent Delivery System. An acetonitrile (MeCN)/ddH_2O gradient was implemented, using 0.01% (v/v) TFA in ddH_2O (Solvent A) and 80% acetonitrile with 0.01% (v/v) TFA (Solvent B). HPLC, or Optima, grade solvents (Fisher, Bioshop) were used and solvents filtered prior to use. A Zorbax 300SB-C3 semi-prep column (Agilent) was used for the purification after first being rinsed thoroughly with 100% MeCN (0.1% TFA).

After injection of p7, the column was run at 4 mL/min in 40% solvent B. A linear gradient was employed to raise the concentration of solvent B to 100% over 40 minutes, at a flow rate of 4 mL/min. At approximately 40 minutes, the peak corresponding to the cleaved p7-His was collected. It was stored at -80°C overnight, and lyophilized the following morning.

4.2.3 Circular dichroism (CD) spectroscopy

Far-UV circular dichroism (CD) spectra were recorded on a Jasco-810 instrument on samples of p7 at decreasing concentrations, prepared by dilution of a stock sample with concentrations of SDS maintained at 100 mM. Eight scans were performed from 250

to 200 nm with a step resolution of 0.2 nm, a speed of 20 nm/min, a bandwidth of 1.0 nm, and response time of 2 s using a 0.1 mm path-length quartz cuvette. Runs were carried out at 22°C and 40°C.

4.3 Results

4.3.1 Expression and purification of p7

The expression and purification of the C-terminal hexahistadine tagged construct of GST-p7 (GST-His-p7) construct was performed using protocols previously developed in the Goto lab (Haddad, 2007).

GST – GGGGGLVPRGSALENLVILNAASLAGTHGLVSFLVFFCFAW YLKGRWVPGAVYAFYGMWPLLLLLLALPQRAYAHHHHHH

Figure 4.2 Amino acid sequence of the GST-p7-His construct. The GST tag is followed by a five-glycine linker (blue), and the sequence for thrombin cleavage (green). The C-terminal hexahistidine tag is indicated in red.

As shown in Figure 4.3, the GST-p7-His construct was expressed in BL21(DE3) cells at high levels, as shown by the appearance of a significant band at ~34 kDa in the Coomassie stained gel of the whole cell lysate. Since the protein was localized to inclusion bodies, nickel affinity chromatography was carried out under denaturing conditions.

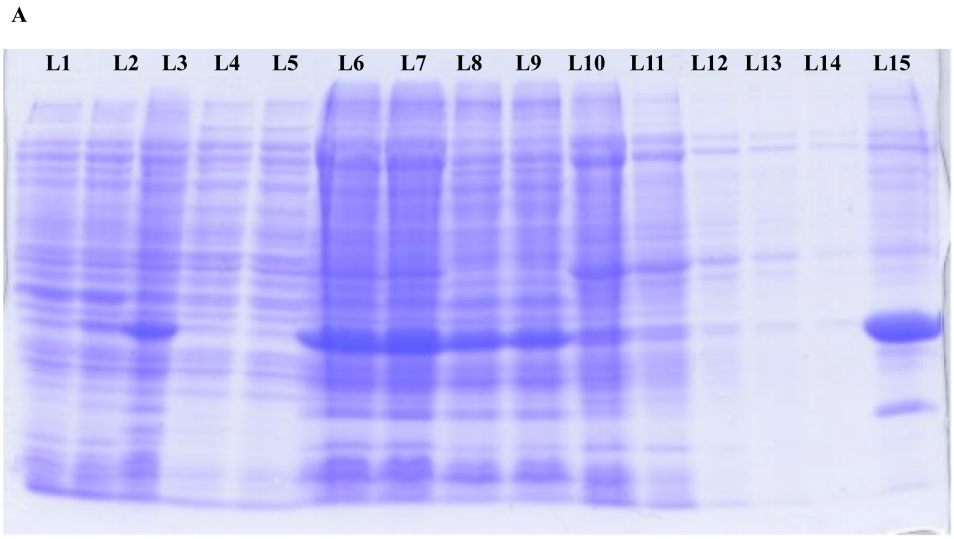

Figure 4.3 Coomassie-stained 15% SDS-PAGE gel of fractions taken at various stages of expression and purification of GST-p7-His construct: L1 – Whole cell lysate sample taken prior to induction with IPTG, and L2 – 1 hour after induction with IPTG, L3 – overnight expression. L4, L5 – supernatant of cell lysate, L6, L7 – insoluble fraction after overnight solubilization, L8- flow-through from application of solubilized sample to the nickel affinity resin, L9 – flow through from reapplication of solubilized fraction, L10– Wash #1, L11 – Wash #2, L12 - Wash #3, L13 – Wash #4, L14 - Wash #5, L15 – Elution

The isolated fusion protein was dialyzed overnight to remove imidazole and urea. It was necessary to include 0.02% SDS in the dialysis buffer to keep the fusion protein in solution. To remove the GST tag, thrombin was used. Almost complete cleavage was demonstrated by SDS-PAGE analysis. There was a decrease in intensity of the band at ~30 kDa, and the appearance of bands at 20 kDa and 8 kDa showing the GST tag and cleaved p7 protein, respectively, with almost complete cleavage being obtained after reaction overnight (Figure 4.4).

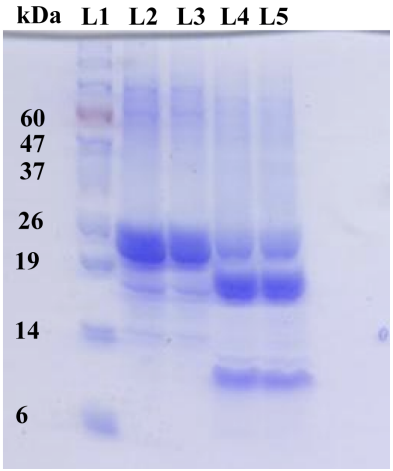

Figure 4.4 Coomasie stained 15% SDS-PAGE gel of fractions taken during thrombin cleavage of GST tag. L1 – protein ladder, L2, L3 – pre-thrombin cleavage, L4, L5 – overnight thrombin cleavage

Following thrombin cleavage, RP-HPLC was carried out to separate the GST tag from the remaining His-p7 protein. The profile of this purification and gel corresponding to elutes collected are shown in Figure 4.5. As was previously demonstrated, the p7-His peak eluted at an acetonitrile concentration of approximately 50%.

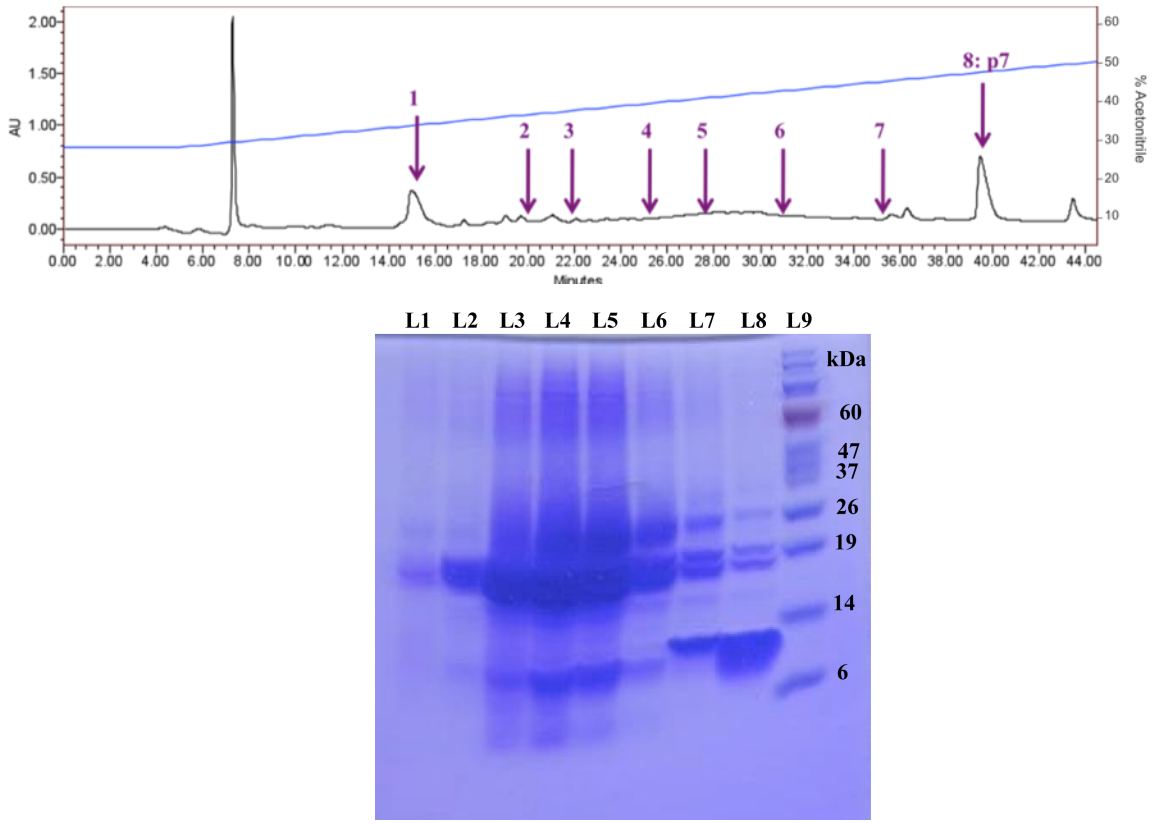

Figure 4.5 RP-HPLC purification of cleaved p7-His. (A) RP-HPLC profile of purified p7. Blue line shows gradient of acetonitrile/ddH_2O (0.01% TFA). The protein eluted at ~40 minutes. (B) Coomassie stained 15% SDS-PAGE gel of fractions taken from the HPLC purification, each lane corresponds to the fraction collected: L1 –1, L2- 2, L3- 3, L4- 4, L5-5, L6-6, L7-7, L8 – 8 (p7), L9 – protein ladder

The elution of the p7 peak at the expected concentration of acetonitrile, as well as the purity of the protein as evaluated by SDS-PAGE (~90%) (Figure 4.5) indicate that the purification was successful. However, yields were very low, with ~0.7 mg of fusion protein per litre produced before cleavage, which resulted in ~100 μg of p7-His being isolated after HPLC. Not only did the low yield make the production of an NMR sample time consuming, the cost would have been prohibitively high for structural studies. Therefore, it was necessary to optimize yields in the p7 purification.

On the SDS-PAGE gel it can be seen that a large amount of the expressed protein in the insoluble fraction does not get solubilized. In an attempt to increase the amount of protein solubilized we carried out the solubilization step for an extended time on the shaker in lysis buffer. This, however, had no effect on the amount of protein obtained. The SDS-PAGE analysis also shows that during Ni-affinity chromatography significant amounts of protein did not bind to the nickel-NTA resin, since significant amounts of fusion protein was found in flow-through fractions, even if re-applied to the column. We sought to optimize the binding of the fusion protein to the resin by increasing the amount of time that the guanidinium-solubilized cell extract was exposed to the resin. We also increased the amount of resin, increasing the number of sites available for the fusion protein to bind. In all cases the modifications did not seem to have any impact, with similar amounts of p7-His being obtained throughout.

We also investigated the possibility that p7-His was lost during the cleavage of the GST tag. For this purpose, fractions from the cleavage reaction were analyzed by SDS-PAGE at various points in time (Fig 4.6). The cleavage of the GST tag from the GST-p7-HIS construct could be observed by the appearance of a band corresponding to free GST after 1 hour incubation with thrombin. The cleavage reaction appears to reach completion overnight, with the loss of the band corresponding to the full-length fusion protein and the appearance of the p7 band at the expected molecular weight. This indicated that cleavage was being carried out as expected, with no detectable loss of p7-His from excess cleavage.

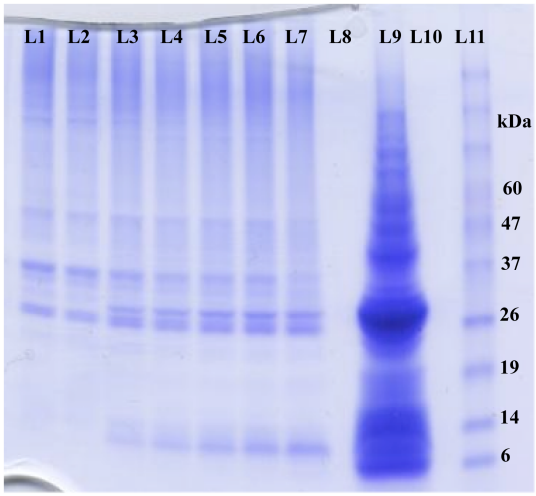

Figure 4.6 Coomasie stained 15% SDS-PAGE gel of thrombin cleavage of GST-p7-His. L1- GST-p7-His after concentration, L2- after concentration with buffer for thrombin cleavage, L3 – 1 hour after addition of thrombin, L4 – 2 hours after addition of thrombin, L5 – 3 hours after addition of thrombin, L6 – 4 hours after addition of thrombin, L7 – overnight after addition of thrombin, L8 – blank, L9 – GST expressed O/N, L10 – blank, L11 – ladder

Examination of the HPLC profile indicated that a significant source of protein loss during the purification was the elution of a broad peak over 18-36 minutes that contained both full-length fusion protein in addition to p7-His (Fig. 4.5). To test whether this loss could be prevented, different aspects of the HPLC procedure were evaluated. For example, we tested the effect of using different sample solubilization conditions prior to injection, different amounts of protein injected in each run, as well as the effect of avoiding the concentration step prior to loading. The results of these experiments are summarized in Table 4.1, with the quantity of p7-His obtained for different RP-HPLC conditions.

Trial #	Fusion Protein Injected (mg)	Fraction of sample that is p7-His (mg)	Concentrated[1] (Y/N)	Injection Conditions	p7 Eluted (mg)
1	4.3	1.0	Y	0.02% SDS	0.06
2	2.2	0.55	Y	0.02% SDS	0.06
3	0.9	0.23	Y	0.02% SDS	0.02
4	1.3	0.33	Y	0.02% SDS	0.03
5	1.2	0.3	N	0.02% SDS	0.01
6	2	0.5	N	0.02% SDS	0.02
7	1.4	0.35	N	0.02% SDS	0.01
8	1.4	0.35	N	0.02% SDS and 40% acetonitrile	0.02
9	0.7	0.18	N	0.02% SDS and 40% acetonitrile	0.02
10	1.3	0.33	N	0.02% SDS and 40% acetonitrile	0.02
11	1.3	0.33	N	0.02% SDS and 10% acetonitrile	0.009

Table 4.1. Summary of different RP-HPLC protocols implemented to optimize the recovery of desired p7 protein during separation of p7-His from GST tag. [1]- whether protein sample was concentrated to a smaller volume to prior to thrombin cleavage (Y-yes, N-No)

While there was some increase in the yield of purified p7 by decreasing the amount of protein that was injected on the column, it was not a large enough improvement to allow an NMR sample to be routinely produced. Typical yields based on the expected amount of p7 based on calculations of total protein content and proportion of the fusion protein sequence that comes from p7-His ranged between 5-9%, regardless of how the sample was run.

During the course of these experiments, samples from several purifications were accumulated (~ 15 samples), giving rise to a sufficient quantity for solution NMR analysis. At that point, structural data was becoming available on p7 in micelles, with no information yet available on its structure in the more bilayer-like conditions of a small isotropic bicelle. Unfortunately, when the bicelle solution was added to the lyophilized sample, the sample was not solubilized, making it impossible to pursue these studies. Given the time constraints of the project, and cost of sample production, we chose to not continue this line of investigation, instead turning to a biophysical method more amenable to the study of smaller quantities of protein.

4.3.2 Circular dichroism studies of p7

Since it has been suggested that p7 oligomerizes to form an ion channel to carry out its biological function, and the structural and functional properties of p7 are often studied in detergent systems, we carried out circular dichroism (CD) to look at the effect of shifting the equilibrium of p7 from the monomeric to oligomeric form. While it has been shown by SDS-PAGE analysis, transmission electron microscopy (TEM) and planar lipid bilayer conductance studies that p7 forms an oligomeric complex that functions as an ion channel (Clarke et al., 2006) it had not been demonstrated whether the assembly and function of this oligomer induced conformational changes in p7. Conformational changes of p7 by oligomerization can be detected by CD. By decreasing the ratio of detergent to protein, we can change the effective concentration of p7 since p7 is confined to the detergent phase, and therefore promote its oligomerization. The effect of oligomerization on the secondary structure of p7 was studied in SDS micelles, since SDS was able to maintain the highly hydrophobic p7 in solution.

Conditions that would be expected to favour the monomeric form of p7, with a protein: detergent ratio of 1:25000, gave rise to a CD spectrum characteristic of alpha helical content, with minima at 208 and 222 nm (Figure 4.7, red spectrum). Increasing the ratio of protein to detergent to favour the formation of oligomer showed an increase in the MRE at these wavelengths, indicating a slight increase in alpha helical character at 40°C (green spectrum). A similar trend was also seen at 22°C (data not shown).

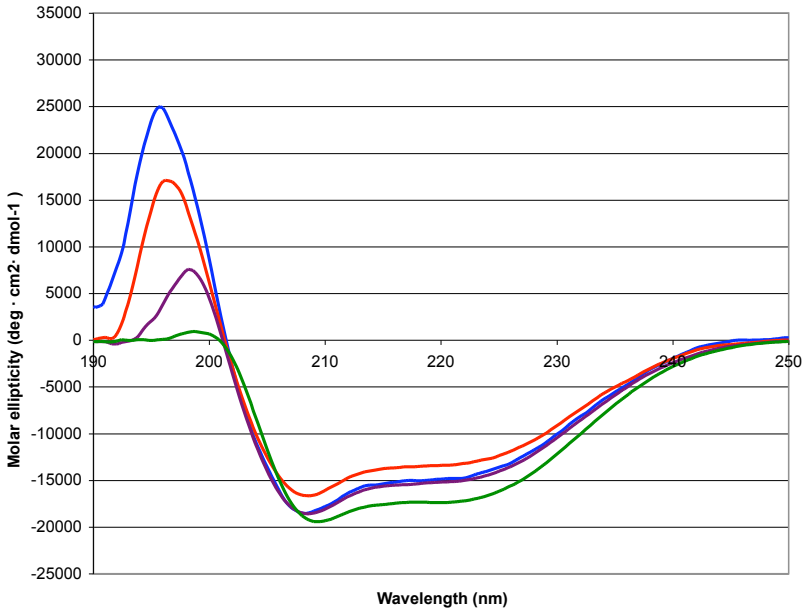

Figure 4.7 Circular dichroism studies on p7, at concentrations of 29 μM (green), 17 μM (purple), 9 μM (blue) and 4 μM (red) in 100 mM SDS micelles, pH 8.0 at 40°C.

4.4 Discussion

4.4.1 Optimization of p7-His purification

In this chapter the goal was to obtain NMR samples to carry out experiments aimed at studying the structure of the monomeric form and oligomerization properties of the membrane protein p7 in detergent and bicelle systems. While we were able to obtain p7-His at 90% purity, the yield of protein produced from each 1L expression growth was very low, requiring large quantities of media with significant costs in time to obtain a single NMR sample.

In our analysis of the purification, a significant loss of p7-His appeared to occur at the RP-HPLC step, following thrombin cleavage of the GST fusion protein. A significant proportion of p7-His appeared to elute with the GST tag, potentially due to non-specific interactions. It is a possibility that cleaved p7-His product may be interacting with uncleaved fusion protein to form dimers or higher order oligomers, leading to poor separation of the desired p7 product from the GST tag. Increasing the concentration of SDS in the starting conditions may help to break apart any dimers or higher order oligomers, or could help indicate whether this is the source of protein loss.

Since the time of this study, there have been various solution NMR studies reporting on the dynamic and structural properties of p7 (Cook et al., 2013; Foster et al., 2014; OuYang et al., 2013). These studies have utilized different protein constructs and variations in purification methods. The first of these structures was of the monomeric form of p7 in DHPC micelles solved by the Opella group in 2013 (Cook et al., 2013) (Figure 4.8A). The His_6-TrpΔLE fusion protein was selected for the expression of p7,

which directed the protein into inclusion bodies. Ni-NTA chromatography was carried out in denaturing conditions to purify the fusion protein, which was then cleaved by cyanogen bromide, and His_6-TrpΔLE and p7 were separated by dialysis (Cook et al., 2012). The solution NMR structure contained a short N-terminal helix preceding the first TM helix, followed by a short loop and a kinked TM helix, in line with previous hydropathy-based predictions on the topology of p7. Chemical shift changes upon addition of amantadine, a known inhibitor of p7 channel activity, were also monitored, with large changes observed in the terminal regions, suggesting a specific interaction with the monomeric form.

The His_6-TrpΔLE-p7 construct was also used to study the structure of the hexameric state of p7 in DPC micelles (OuYang et al., 2013a) (Figure 4.8C). The p7 construct had five mutations introduced to increase the stability of the protein. The purification technique was similar to that carried out for the structure of the monomeric form of p7, but instead of dialysis to separate the cleaved fusion protein, RP-HPLC was used (OuYang et al., 2013). Samples made in this way were used to determine that Leu20 within the channel is the amantadine binding site, which differs from the binding site identified in the monomeric state. They proposed, based on NOE studies, that the p7 hexamer forms a pore that has an open and closed state, and that the binding of amantadine maintains the closed state of the pore. Virtual screening identified candidate molecules that could interact at this site and act as inhibitors by stabilizing the closed form of the channel through an allosteric mechanism. This allosteric mechanism is described as similar to what has been suggested for M2, another viroporin channel (Cady et al., 2010).

The Griffin group solved a different structure of p7 in 2014, using a construct with p7 fused to a FLAG tag (FLAG-p7) in MeOH (Foster et al., 2014) (Figure 4.8B). To express FLAG-p7 a GST fusion tag was used. The GST-FLAG-p7 fusion protein was effectively expressed and the GST tag was cleaved off by PreScission protease and the cleaved products were separated by RP-HPLC (Clarke et al., 2006). This structure showed two TM helices connected by a short loop, and a short helix at the C-terminus. This structure was used to model a hexameric oligomer that forms a flower-like structure. This oligomeric form of p7 was similar to that proposed in an separate study by electron microscopy (EM) using a chemically synthesized p7 protein (Luik et al., 2009).

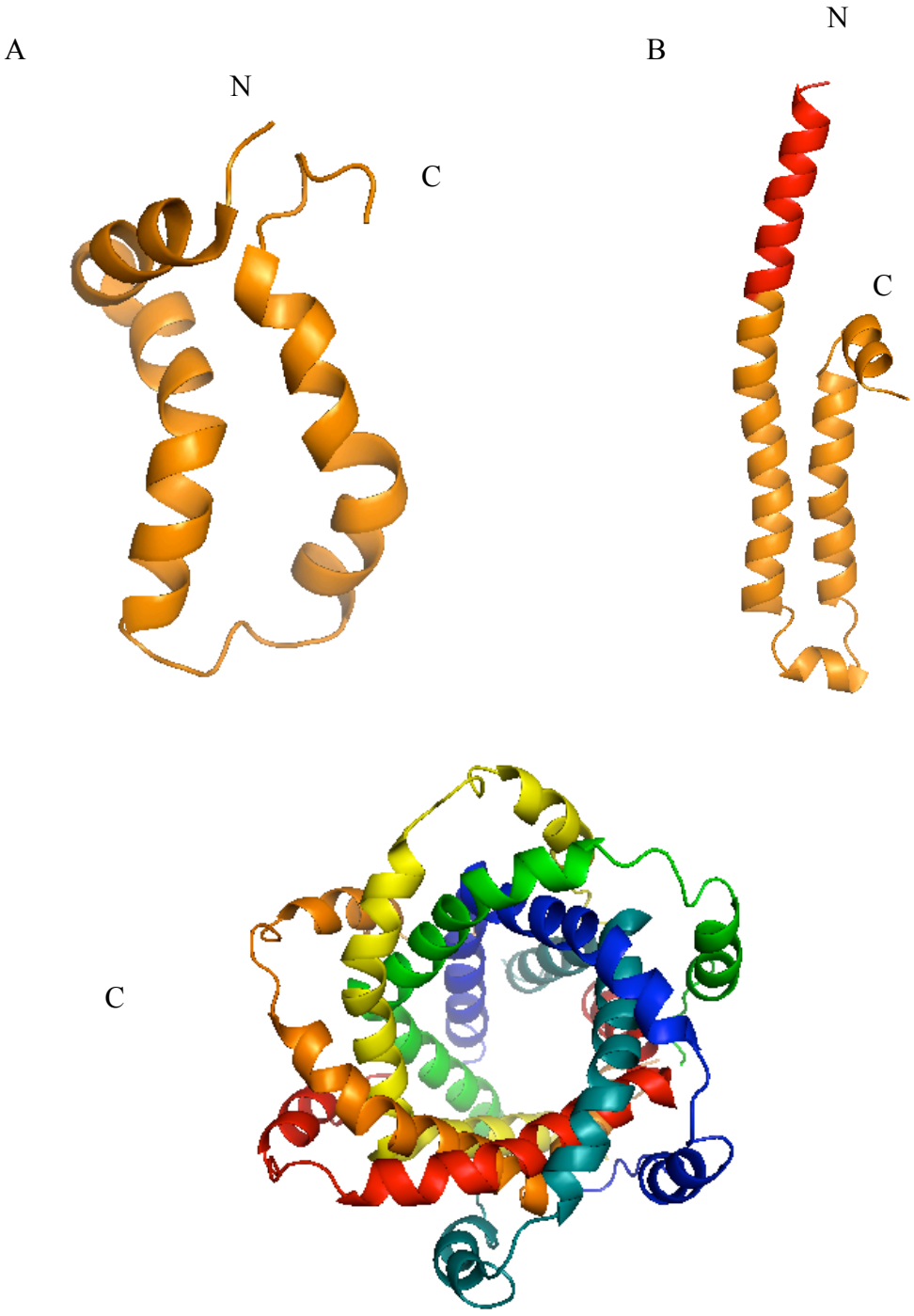

Figure 3.1 Three-dimensional structure of p7. (A) Solution-state NMR structure of monomeric p7 in DHPC micelles (PDB ID: 2MTS). (B) Solution-state NMR structure of FLAG-p7 (FLAG tag in red) in MeOH (PDB ID: 3ZD0). (C) Solution-state NMR structure of the p7 hexamer in DPC micelles (PDB ID: 2M6X).

4.4.2 Evidence of p7 oligomerization in SDS micelles

It was found that as the ratio of protein:detergent is increased there is an increase in α-helical secondary structure. The α-helical structure was in line with what has previously been published for p7 in SDS micelles (Montserret et al., 2010). The increase in α-helical character in conditions that favor oligomerization was similar to what has been seen for TM helix 5 of the human adensoine A_{2A} receptor ($A_{2A}R$), belonging to the GPCR family (Lazarova et al., 2005). As protein concentration was increased relative to the concentration of DMPC vesicles, there was an increase in α-helical character detected by CD. The change in secondary structure content indicates that there are changes to the conformation of the protein, which could be the result of helix formation promoted by oligomerization. The oligomeric structure of p7 solved in DPC micelles, showed that the secondary structure of the monomeric form was consistent with the structure solved in DHPC micelles and organic solvent, however since these are different solvents it is difficult to compare the impact of oligomerization on secondary structure content (Cook & Opella, 2011; Montserret et al., 2010; OuYang et al., 2013). However, the secondary structure of p7 in DHPC determined by solution NMR show that while there are two TM helices, the dynamics suggest a break in each helix that induces kinks not seen in the oligomeric state (Cook & Opella, 2011) . This would result in less α-helical character when p7 is in its monomeric form, which is what we see in our studies.

4.4.3 Future directions

Continued efforts in the optimization of the purification of GST-p7-His would be required to study this construct by solution-state NMR. Other groups that have successfully used RP-HPLC to purify p7 following cleavage from the His-Trp leader sequence (OuYang et al., 2013). In those studies a C18 column was utilized, which is more hydrophobic than the C3 column used in our purification. While it is possible that the use of a different stationary phase might have been required to separate p7-His from GST, the GST carrier itself may have had a tendency to aggregate with cleaved p7-His. If there was a reduced tendency for nonspecific interactions between p7-His and the Trp leader peptide carrier protein, this might explain the success of the purification from that fusion protein.

While structural data has been obtained about the monomeric and oligomeric forms of p7, these have been carried out in different solvent systems. In order to effectively understand the formation of the oligomeric form of p7 and how it functions to propogate HCV within the body, it would be interesting to study the transition between the monomeric and oligomeric form of p7 since this could provide insight into how to prevent its formation. SDS has shown promise as a detergent where the formation of the oligomer could be favored by increasing the protein:detergent ratio. It would be beneficial to look at this trend in other detergent systems (including DPC and DHPC). If the purification technique can be optimized to produce a higher yield of protein then NMR would be a powerful technique to monitor the equilibrium between monomeric and oligomeric states of p7 in various membrane-mimetic media.

Chapter 5: Summary and Future Directions

In this thesis the dimerization properties of the M13 bacteriophage MCP was studied using SDS micelles as the membrane-mimetic system. Under conditions of excess micelles, SDS exhibited near-ideal solvent behavior, similar to observations that have been made for TM segments studied in phospholipid bicelles (Bocharov et al., 2012). These measurements allowed quantification of the increase in apparent affinity in MCP TM helix interactions that are brought about by the removal of its N-terminal amphipathic helix. However, a detailed analysis of this interaction using a model that treats the empty micelle as an active species in the reaction allowed us to determine that the affinity difference can be attributed to differences between the interactions of SDS molecules with MCP versus MCP_{TM}.

While our findings help explain the differences in migration behavior seen for MCP versus the TM peptide in SDS-PAGE gels (Rath et al., 2009), they also highlight the influence of the membrane mimetic system used to measure free energies of interaction. Although attempts were made to perform similar types of measurements in other detergents that are commonly employed in TM peptide interaction studies (including a series of experiments not described in the thesis using phospholipid bicelles), these efforts were largely unsuccessful, with practical considerations preventing the measurement of apparent affinities by NMR. Perhaps more importantly, the central role of the solvent system in the affinity measurement indicates that the most relevant measurements of TM helix association will be those done under conditions most similar to those found *in vivo*. Moreover, cellular membranes are not only made up of phospholipids, but contain a number of other macromolecules that contribute to the

overall structure, fluidity and physical properties of the membrane (Ellis, 2001; Ellis, 2001; Minton, 2000). Therefore, in order to determine the impact of the amphipathic helix on TM interactions for MCP, it would be ideal to study the system in *E. coli* lipid membranes.

Methods for the measurement of membrane protein dimerization affinities in biological membranes have been developed for mammalian systems using FRET and giant plasma membrane vesicles (GPMV). These vesicles are derived from mammalian cells by induced blebbing of the plasma membrane (Baumgart et al., 2007; Sengupta et al., 2008). In one such study, GpA with fluorescent donor or acceptor tags were expressed together in Chinese hamster ovary (CHO) cells, which was then used to produce GPMVs that could be used to measure FRET between GpA protein in these membranes (Chen et al., 2010). Dimer and monomer populations could be calculated based on FRET intensities, allowing free interaction energies to be derived. The calculated value was -3.9±0.2 kcal/mol, this was lower than what has been previously measured in micelle systems indicating that the presence of other macromolecules in the lipid vesicle weakens the strength of the interaction (Fisher et al., 1999). A similar study was carried out recently for receptor tyrosine kinase (RTK) FGFR3 (Sarabipour et al., 2015), which allowed them to study dimerization of the TM domain of FGFR3, as well as the impact of ligand binding on this dimerization.

The application of this technique to well studied dimer systems such as GpA and RTK shows great promise for the study of MCP. Since MCP is incorporated in bacterial cell membranes it would be beneficial to carry out similar studies in vesicles made of bacterial cell membranes. The ability to carry out blebbing of *E. coli* membranes has been

demonstrated (Katsui et al., 1982) and these vesicles were found to have lipopolysaccharide to phospholipid ratio similar to that of native *E. coli* membranes. It would allow measurement of more physiologically relevant affinities, and could provide further support for the influence of the N-terminal helix in modulating the strength of TM helix self-association.

Despite structural studies being carried out on the oligomeric form of p7, there is little information about how the channel assembles and whether there is more than one oligomeric state. The structures that were solved for the p7 monomer and oligomer were carried out in DPC and DHPC micelles. DPC micelles are often chosen for membrane protein structure determination since its zwitterionic headgroup is less likely to participate in interactions that would disrupt the protein structure (Kallick et al., 1995). It has been used to study large multispanning membrane proteins by solution-state NMR, including diacylglycerol kinase (DAGK), the phospholamban (PLM) pentamer and mitochondrial uncoupling protein 2 (UCP2) (Berardi et al., 2011; Traaseth et al., 2007; W. Van Horn et al., 2009) Similarly, DHPC micelles posses a zwitterionic headgroup, with a phosphoglycerol backbone in the headgroup that is the same as occurs in natural lipid bilayers (Hauser, 2000). While these two detergent systems have been described as better mimics of lipid bilayer, it is also possible to study oligomeric protein structures in SDS micelles. For example, SDS has been shown to be suitable for the structural studies of the tetrameric channel KcsA potassium channel (Chill et al., 2006). It was found to be the optimal detergent for solubilization of KcsA, based on CD studies evaluating secondary structure content. However, the impact of changing the protein:detergent ratios to promote oligomer formation has not been explored, and could allow association

energetics and mechanism to be defined. If purification of p7 can be optimized it would be beneficial to carry out similar studies as we did in MCP in SDS micelles, to study the oligomeric forms of p7 by solution-state NMR and determine the self association energies for p7 by NMR or analytical ultracentrifugation (AUC). The GPMV system could also be applied directly to the study of the channel formation of p7. This would provide new information about the interactions of p7 monomers, as well as corroborate *in vitro* studies on channel formation. Understanding channel formation and interaction between p7 monomers will ultimately help elucidate the function of the oligomeric form of p7, and how its formation can be inhibited.

Through the studies conducted in this thesis we were able to probe the suitability of SDS micelles for the study of transmembrane domain interaction. Also, we were able to quantify the impact of the amphipathic helix on M13 MCP TM segment dimerization affinities. SDS was also shown to have potential for the study of the oligomerization of HCV p7. Overall, these results suggest new approaches to be considered for the goal of enhancing understanding about the interactions of TM domains and the influence of extramembranous domains of MCP, p7 and potentially other protein systems.

References

Abragam, A. (1961). *Principles of Nuclear Magnetism*. Oxford University Press, Oxford.

Agirre, A., Barco, A., Carrasco, L., & Nieva, J. L. (2002). Viroporin-mediated membrane permeabilization. Pore formation by nonstructural poliovirus 2B protein. *The Journal of biological chemistry*, *277*(43), 40434–41.

Almeida, F. C., & 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. *Journal of molecular biology*, *270*(3), 481–495.

Altieri, A., & Byrd, R. (1995). Randomization approach to water suppresion in multidimensional NMR using pulsed field gradients. *Journal of Magnetic Resonance B*, *107*(3), 260–266.

Anbazhagan, V., Munz, C., Tome, L., & Schneider, D. (2010). Fluidizing the membrane by a local anesthetic: phenylethanol affects membrane protein oligomerization. *Journal of molecular biology*, *404*(5), 773–7.

Andersen, O. S., & Koeppe, R. E. (2007). Bilayer thickness and membrane protein function: an energetic perspective. *Annual review of biophysics and biomolecular structure*, *36*, 107–30.

Aue, W., Bartholdi, E., & Ernst, R. (1976). Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *The Journal of Chemical Physics*, *64*(5), 2229–2246.

Baneyx, F. (1999). Recombinant protein expression in Escherichia coli. *Current Opinion in Biotechnology*, *10*(5), 411–421.

Barrett, P. J., Song, Y., Van Horn, W. D., Hustedt, E. J., Schafer, J. M., Hadziselimovic, A., & Sanders, C. R. (2012). The amyloid precursor protein has a flexible transmembrane domain and binds cholesterol. *Science (New York, N.Y.)*, *336*(6085), 1168–71.

Baumgart, T., Hammond, A. T., Sengupta, P., Hess, S. T., Holowka, D. A., Baird, B. A., & Webb, W. W. (2007). Large-scale fluid / fluid phase separation of proteins and lipids in giant plasma membrane vesicles. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(9), 3165–3170.

Behrens, S., Tomei, L., & De Francesco, R. (1996). Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *The EMBO Journal*, *15*(1), 12–22.

Berardi, M. J., Shih, W. M., Harrison, S. C., & Chou, J. J. (2011). Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature*, *476*(7358), 109–13.

Binding, L., & Kiefer, H. (1996). Expression of an Olfactory Receptor in Escherichia coli : Purification ,. *Biochemistry*, *2960*(96), 16077–16084.

Biological Magnetic Resonance Data Bank. Retrieved from http://www.bmrb.wisc.edu/ref_info/statful.htm

Bloom, M., Evans, E., & Mouritsenj, O. (1991). Physical properties of the fluid lipid-bilayer component of cell membranes : a perspective. *Quarterly reviews of biophysics*, *3*, 293–397.

Bodenhausen, G. & Ruben, D. J. (1980). Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett*, *69*, 185-189

Bocharov, E. V., Mineev, K. S., Goncharuk, M. V, & Arseniev, A. S. (2012). Structural and thermodynamic insight into the process of "weak" dimerization of the ErbB4 transmembrane domain by solution NMR. *BBA - Biomembranes*, *1818*(9), 1–13.

Bocharov, E. V, Pustovalova, Y. E., Pavlov, K. V, Volynsky, P. E., Goncharuk, M. V, Ermolyuk, Y. S., & Arseniev, A. S. (2007). Unique dimeric structure of BNip3 transmembrane domain suggests membrane permeabilization as a cell death trigger. *The Journal of biological chemistry*, *282*(22), 16256–16266.

Bocharov, Eduard V, Lesovoy, D. M., Goncharuk, S. A., Goncharuk, M. V, Hristova, K., & Arseniev, A. S. (2013). Structure of FGFR3 transmembrane domain dimer: implications for signaling and human pathologies. *Cell*, *21*(11), 2087–93.

Bodenhausen, G., & Ruben, D. (1980). Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters*, *69,* 185-189.

Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R. L., Mathieu, M., & Francesco, R. De. (1999). Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(23), 13034–13039.

Cady, S. D., Schmidt-Rohr, K., Wang, J., Soto, C., Degrado, W., & Hong, M. (2010). Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature*, *463*(7281), 689–692.

Canadian Liver Foundation - Hepatitis C. Retrieved July 12, 2015, from http://www.liver.ca/liver-disease/types/viral_hepatitis/Hepatitis_C.aspx

Carr, D. (2002). *The handbook of Analysis and Purificatino of Peptides and Proteins by Reversed-Phase HPLC*. California, USA: GraceVydac.

Carrère-Kremer, S., Montpellier-Pala, C., Cocquerel, L., Wychowski, C., Penin, F., & Dubuisson, J. (2002). Subcellular localization and topology of the p7 polypeptide of hepatitis C virus. *Journal of Virology*, *76*(8), 3720–3730.

Cavanagh, J., Fairbrother, W. J., Palmer, W., Rance, R. & Skelton, N. (2007). *Protein NMR spectroscopy: principles and practice*. San Diego: Academic Press.

Chen, B. G., Ray, R., Dubik, D., Shi, L., Cizeau, J., Bleackley, R. C., & Greenberg, A. H. (1997). Mitochondrial protein that activates apoptosis. *The Journal of Experimental Medicine*, *186*(12), 1975–1983.

Chen, G., Cizeau, J., Velde, C. Vande, Park, J. H., Bozek, G., Bolton, J., & Greenberg, A. (1999). Nix and Nip3 form a subfamily of pro-apoptotic mitochondrial proteins. *The Journal of Biological Chemistry*, *274*(1), 7–11.

Chen, L., Novicky, L., Merzlyakov, M., Hristov, T., & Hristova, K. (2010). Measuring the energetics of membrane protein dimerization in mammalian membranes. *Journal of the American Chemical Society*, *132*(21), 3628–3635.

Chill, J., Louis, J., Miller, C., & Bax, A. (2006). NMR study of the tetrameric KcsA potassium channel in detergent micelles. *Protein Science*, *15*, 684–698.

Chou, J. J., Baber, J. L., & Bax, A. (2004). Characterization of phoshpolipid mixed micelles by translational diffusion. *Journal of biomolecular NMR*, *29*, 299–308.

Chou, J., Kaufman, J., Stahl, S., Wingfield, P., & Bax, A. (2011). Micelle-induced curvature in a water-insoluble HIV-1 Env peptide revealed by NMR Dipolar coupling measurement in stretched polyacrylamide gel. *Journal of the American Chemical Society Communications*, *124*, 2450–2451.

Claridge, T. (1999). *High-resolution NMR techniques in organic chemistry*. New York: Pergamon.

Clarke, D., Griffin, S., Beales, L., Gelais, C. S., Burgess, S., Harris, M., & Rowlands, D. (2006). Evidence for the formation of a heptameric ion channel complex by the hepatitis C virus p7 protein in vitro. *The Journal of biological chemistry*, *281*(48), 37057-37068

Cocquerel, L., Voisset, C., & Dubuisson, J. (2006). Hepatitis C virus entry: potential receptors and their biological functions. *The Journal of general virology*, *87*(Pt 5), 1075–1084.

Cohn, E. J., & Conant, J. B. (1926). The molecular weights of proteins in phenol. *Proceedings of the National Academy of Sciences of the United States of America*, *12*, 433–438.

Columbus, L., Lipfert, J. A. N., Klock, H., Millett, I. A. N., Doniach, S., & Lesley, S. A. (2006). Expression, purification, and characterization of Thermotoga maritima membrane proteins for structure determination. *Protein Science*, *15*, 961–975.

Cook, D. N., Pisetsky, D. S., & Schwartz, D. A. (2004). Toll-like receptors in the pathogenesis of human disease. *Nature Immunology*, *5*(10), 975–979.

Cook, G. A., Dawson, L. A., Tian, Y., & Opella, S. J. (2013). Three-dimensional structure and interaction studies of hepatitis C virus p7 in 1,2-dihexanoyl- sn -glycero-3-phosphocholine by solution nuclear magnetic resonance. *Biochemistry*, *52*, 5295–5303.

Cook, G., Stefer, S., & Opella, S. J. (2012). Expression and purification of the membrane protein p7 from hepatitis C virus. *Biopolymers*, *96*(1), 32–40.

Cook, & Opella, S. (2011). Secondary structure, dynamics, and architecture of the p7 membrane protein from hepatitis C virus by NMR spectroscopy. *Biochimica et biophysica acta*, *1808*(6), 1448–1453.

Crick, D. J., Wang, J. X., Graham, B., Swarbrick, J. D., Mott, H. R., & Nietlispach, D. (2015). Integral membrane protein structure determination using pseudocontact shifts. *Journal of biomolecular NMR*, *61*(3-4), 197–207.

Cunningham, F., & Deber, C. M. (2007). Optimizing synthesis and expression of transmembrane peptides and proteins. *Methods (San Diego, Calif)*, *41*(4), 370–380.

Cymer, F., & Schneider, D. (2012). Oligomerization of polytopic α-helical membrane proteins: causes and consequences. *Biological chemistry*, *393*(11), 1215–30.

Cymer, F., Veerappan, A., & Schneider, D. (2012). Transmembrane helix – helix interactions are modulated by the sequence context and by lipid bilayer properties. *BBA - Biomembranes*, *1818*(4), 963–973.

Dawson, J. P., Melnyk, R. A., Deber, C. M., & Engelman, D. M. (2003). Sequence context strongly modulates association of polar residues in transmembrane helices. *Journal of Molecular Biology*, *331*(1), 255–262.

Deber, C. M., Khan, A. R., Li, Z., Joensson, C., Glibowicka, M., & Wang, J. (1993). Val Ala mutations selectively alter helix-helix packing in the transmembrane segment of phage M13 coat protein. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(24), 11648–11652.

Delaglio, FS, Grzesiek, G., Vuister, G., Zhu, G., Pfeifer, J., & Bax, A. (1995). NMRPipe: a multidimensinal spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR*, *6*, 277–293.

Dingley, A., Mackay, J., Chapman, B., Morris, M., Kuchel, P., Hambly, B., & King, G. (1995). Measuring protein self-association using pulsed-field-gradient NMR spectroscopy : Application to myosin light chain 2. *Journal of Biomolecular NMR*, *6*(3), 321–328.

Domingues, C., Boelens, R., & Bonvin, A. (2003). HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*, 6150–6156.

Drew, D., Fröderberg, L., Baars, L., & de Gier, J. W. L. (2003). Assembly and overexpression of membrane proteins in Escherichia coli. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1610*(1), 3–10.

Drin, G., & Antonny, B. (2010). Amphipathic helices and membrane curvature. *FEBS letters*, *584*(9), 1840–1847.

Duong, M. T., Jaszewski, T. M., Fleming, K. G., & MacKenzie, K. R. (2007). Changes in apparent free energy of helix-helix dimerization in a biological membrane due to point mutations. *Journal of molecular biology*, *371*(2), 422–434.

Duquesne, K., & Sturgis, J. N. (2010). Heterologous Expression of Membrane Proteins. *Methods in Molecular Biology*, *601*, 205–217.

Dutta, S., Morrison, E. A., & Henzler-Wildman, K. A. (2014). EmrE dimerization depends on membrane environment. *Biochimica et biophysica acta*, *1838*(7), 1817–1822.

Egger, D., Wölk, B., Gosert, R., Bianchi, L., Blum, H. E., Moradpour, D., & Bienz, K. (2002). Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex, *76*(12), 5974–5984.

Ellis, R. J. (2001). Macromolecular crowding : an important but neglected aspect of the intracellular environment. *Current opinion in structural biology*, *11*, 114–119.

Ellis, R. J. (2001). Macromolecular crowding : obvious but underappreciated. *Trends in Biochemical Sciences*, *26*(10), 597–604.

Fantl, W. J., Johnson, D. E., & Williams, L. T. (1993). Signalling by receptor tyrosine kinases. *Annual Review of Biochemistry*, *62*, 453–481.

Fassina, G., Merli, S., Germani, S., Ciliberto, G., & Cassani, G. (1994). Human yield expression and purification of human endothelin-1. *Protein Expression and Purification*, *5*, 559–568.

Feigenson, G. W. (2009). Phase boundaries and biological membranes. *Annual Review of Biophysics and Biomolecular Structure*, *36*, 63–77.

Feld, J. J., & Hoofnagle, J. H. (2005). Mechanism of action of interferon and ribavirin in treatment of hepatitis C. *Nature*, *436*(7053), 967–972.

Feng, J., Russel, M., & Model, P. (1997). A permeabilized cell system that assembles filamentous bacteriophage, *94*(April), 4068–4073.

Fernández, C., & Wider, G. (2003). TROSY in NMR studies of the structure and function of large biological macromolecules. *Current opinion in structural biology*, *13*(5), 570–580.

Fischer, W. B., & Sansom, M. S. P. (2002). Viral ion channels : structure and function. *Biochimica et biophysica acta*, *1561*(1), 27–45.

Fisher, L. E., Engelman, D. M., & Sturgis, J. N. (1999). Detergents Modulate Dimerization, but not helicity, of the Glycophorin A transmembrane domain. *Journal of molecular biology*, *293*, 639–651.

Fleming, K. (2002). Standardizing the free energy change of transmembrane helix–helix interactions. *Journal of Molecular Biology*, *323*(3), 563–571.

Fleming, K. G., Ackerman, A. L., & Engelman, D. M. (1997). The effect of point mutations on the free energy of transmembrane a -Helix dimerization. *Journal of Molecular Biology*, *272*, 266–275.

Fleming, KG, & Engelman, D. (2001). Specificity in transmembrane helix–helix interactions can define a hierarchy of stability for sequence variants. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(25), 14340–14344.

Foster, T. L., Thompson, G. S., Kalverda, A. P., Kankanala, J., Bentham, M., Wetherill, L. F., & Griffin, S. (2014). Structure-guided design affirms inhibitors of hepatitis C virus p7 as a viable class of antivirals targeting virion release. *Hepatology (Baltimore, Md.)*, *59*(2), 408–422.

Francesco, D. (1999). Molecular virology of the hepatitis C. *Journal of Hepatology*, *31*, 47–53.

Fromme, P., & Grotjohann, I. (2008). Structure of photosystems I and II. *Results and Problems in Cell Differentiation*, *45*, 33–72.

Galossi, A., Guarisco, R., Bellis, L., & Puoti, C. (2007). Extrahepatic manifestations of chronic HCV infection. *Journal of Gastrointestinal Liver Disease*, *16*(1), 65–73.

Garavito, R. M., & Ferguson-Miller, S. (2001). Detergents as tools in membrane biochemistry. *The Journal of Biological Chemistry*, *276*(35), 32403–32406.

Gardner, K. H., & Kay, L. E. (1998). The use of 2H, 13C, 15N multidemensional NMR to study the structure and dynamics of proteins. *Annual Review of Biophysics and Biomolecular Structure*, *27*, 357–406.

Gautier, A., Mott, H. R., Bostock, M. J., Kirkpatrick, J. P., & Nietlispach, D. (2010). Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nature Publishing Group*, *17*(6), 768–774.

Glucksmant, M. J., & Bhattacharjee, S. (1992). Three-dimensional structure of a cloning vector: X-ray diffraction studies of a filamentous bacteriophage M13 at 7 A resolution. *Journal of Molecular Biology*, *226*, 455–470.

Goh, F. G., & Midwood, K. S. (2012). Intrinsic danger: activation of toll-like receptors in rheumatoid arthritis. *Rheumatology (Oxford, England)*, *51*(1), 7–23.

Gonzalez, M., & Carrasco, L. (2003). Viroporins. *FEBS Letters*, *552*, 28–34.

Gounarides, J. S., Chen, A., & Shapiro, M. J. (1999). Nuclear magnetic resonance chromatography : applications of pulse field gradient diffusion NMR to mixture analysis and ligand – receptor interactions. *J Chromatography B: Biomedical Sciences and Applications*, *725*(1), 79–90.

Grasberger, B., Minton, A. P., DeLisi, C., & Metzger, H. (1986). Interaction between proteins localized in membranes. *Proceedings of the National Academy of Sciences of the United States of America*, *83*, 6258–6262.

Greenfield, N. J. (1996). Methods to Estimate the Conformation of Proteins and Polypeptides from Circular Dichroism Data. *Analytical biochemistry*, *10*(235), 1–10.

Grice, A., Kerr, I., & Sansom, M. S. (1997). Ion channels formed by HIV-1 Vpu: a modelling and simulation study. *FEBS Letters*, *405*(3), 299–304.

Griffin, S. D. C., Beales, L. P., Clarke, D. S., Worsfold, O., Evans, S. D., Jaeger, J., & Rowlands, D. J. (2003). The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Letters*, *535*(1-3), 34–38.

Grzesiek, S, & Bax, A. (1992). Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *Journal of the American Chemical Society*, *114*(16), 6291–6293.

Grzesiek, Stephan, & Bax, A. (1992). An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *Journal of Magnetic Resonance (1969)*, *207*(99), 201–207.

Guan, Y., Zhang, H., & Wang, A. H. (1995). Electrostatic potential distribution of the gene V protein from Ff phage facilitates cooperative DNA binding : A model of the GVP-ssDNA complex. *Protein Science*, *4*, 187–197.

Guo, J.T., Sohn, J. a, Zhu, Q., & Seeger, C. (2004). Mechanism of the interferon alpha response against hepatitis C virus replicons. *Virology*, *325*(1), 71–81.

Haddad, L. (2007). *Towards Structure Determination of the Hepatits C Virus*. University of Ottawa.

Haigh, N. G., & Webster, R. E. (1998). The major coat protein of filamentous bacteriophage f1 specifically pairs in the bacterial cytoplasmic membrane. *Journal of molecular biology*, *279*(1), 19–29.

Haqshenas, G., Mackenzie, J. M., Dong, X., & Gowans, E. J. (2007). Hepatitis C virus p7 protein is localized in the endoplasmic reticulum when it is encoded by a replication-competent genome. *The Journal of general virology*, *88*(Pt 1), 134–42.

Harboring, H.-C., Replicons, S., Gosert, R., Egger, D., Lohmann, V., Bartenschlager, R., & Moradpour, D. (2003). Identification of the Hepatitis C Virus RNA Replication Complex in. *Journal of Virology*, *77*(9), 5487–5492.

Hauser, H. (2000). Short-chain phospholipids as detergents. *Biochimica et Biophysica Acta*, *1508*, 164–181.

Hayashi, N., & Takehara, T. (2006). Antiviral therapy for chronic hepatitis C: past, present, and future. *Journal of Gastroenterology*, *41*(1), 17–27.

Helenius, A., McCaslin, D. R., Fries, E., & Tanford, C. (1979). Properties of detergents. *Methods in enzymology*, *56*, 734–749.

Helms, V. (2002). Forces behind transmembrane protein folding and supramolecular complex assembly. *European Molecular Biology Organization Reports*, *3*(12), 1133–1138.

Hemminga, M. A., Sanders, J. C., & Spruijt, R. B. (1992). Spectroscopy of lipid-protein interactions: structural aspects of two different forms of the coat protein of

bacteriophage M13 incorporated in model membranes. *Progress in Lipid Research*, *31*(3), 301–333.

Henry, G. D., & Sykes, B. D. (1992). Assignment of amide 1H and 15N NMR resonances in detergent-solubilized M13 coat protein: a model for the coat protein dimer. *Biochemistry*, *31*(23), 5284–5297.

Henry, G. D., Weiner, J. H., & Sykes, B. D. (1986). Backbone dynamics of a model membrane protein: 13C NMR spectroscopy of alanine methyl groups in detergent-solubilized M13 coat protein. *Biochemistry*, *25*(3), 590–598.

Herian, U., Lohmann, V., Ko, F., & Bartenschlager, R. (1997). Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *Journal of Virology*, *71*(11), 8416–8428.

Holsinger, L. J., & Lamb, R. A. (1991). Influenza virus M, integral membrane protein is a homotetramer stabilized by formation of disulfide bonds. *Virology*, *183*(1), 32–43.

Hong, H., & Bowie, J. U. (2011). Dramatic destabilization of transmembrane helix interactions by features of natural membrane environments. *Journal of the American Chemical Society*, *133*, 11389–11398.

Hoofnagle, J., & Seeff, L. (2006). Peginterferon and ribavirin for chronic hepatitis C. *New England Journal of Medicine*, *355*(23), 2444–2451.

Horsmans, Y. (2005). Chronic hepatitis C, depression and interferon. *Journal of Jepatology*, *42*(6), 788–789.

Hou, Z., Kelly, E., & Robia, S. (2008). Phosphomimetic mutations increase phospholamban oligomerization and alter the structure of its regulatory complex. *Journal of Biological Chemistry*, *283*(43), 28996–29003.

Hundt, J., Li, Z., & Liu, Q. (2013). Post-translational modifications of hepatitis C viral proteins and their biological significance. *World Journal of Gastroenterology*, *19*(47), 8929–39.

Jin, M. S., Kim, S. E., Heo, J. Y., Lee, M. E., Kim, H. M., Paik, S.G., & Lee, J.-O. (2007). Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell*, *130*(6), 1071–82.

Johnson, B. (2004). Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol*, *278*, 313–352.

Johnson, B. A., & Blevins, R. A. (1994). NMRView : A computer program for the visualization and analysis of NMR data. *Journal of biomolecular NMR*, *4*(5), 603–614.

Johnson, R. M., Rath, A., & Deber, C. M. (2006). The position of the Gly-xxx-Gly motif in transmembrane segments modulates dimer affinity. *Biochemistry and cell biology*, *84*(6), 1006–1012.

Johnson, R., Rath, A., Melnyk, R., & Deber, C. (2006). Lipid solvation effects contribute to the affinity of Gly-xxx-Gly motif-mediated helix-helix interactions. *Biochemistry*, *45*(28), 8507–8515.

Jones, C. T., Murray, C. L., Eastman, D. K., Tassello, J., & Rice, C. M. (2007). Hepatitis C virus p7 and NS2 proteins are essential for production of infectious virus. *Journal of Virology*, *81*(16), 8374–83.

Kallick, D., Tessmer, M., Watts, C., & Li, C. (1995). The use of dodecylphosphocholine micelles in solution NMR. *Journal of Magnetic Resonance B*, *109*, 60–65.

Kanelis, V., Forman-Kay, J. D., & Kay, L. E. (2001). Multidimensional NMR Methods for Protein Structure Determination. *IUMBM Life, 4*(291), 291–302.

Kang, J. Y., Nan, X., Jin, M. S., Youn, S.J., Ryu, Y. H., Mah, S., & Lee, J.O. (2009). Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity*, *31*(6), 873–884.

Kaptein, R., & Wagner, G. (2015). NMR studies of membrane proteins. *Journal of Biomolecular NMR*, *61*(3-4), 181–184.

Karim, C. B., Stamm, J. D., Karim, J., Jones, L. R., & Thomas, D. D. (1998). Cysteine Reactivity and Oligomeric Structures of Phospholamban and Its Mutants. *Biochemistry*, *2960*(98), 12074–12081.

Katsui, N., Tsuchido, T., Hiramatsu, R., Fujikawa, S., & Shibasaki, I. (1982). Heat-induced blebbing and vesiculation of the outer Membrane of Escherichia coli. *Journal of Bacteriology*, *151*(3), 1523–1531.

Kawai, T., & Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature immunology*, *11*(5), 373–384.

Kehoe, J. W., & Kay, B. K. (2005). Filamentous phage display in the new millennium. *Chemical reviews*, *105*(11), 4056–4072.

Kim, H. J., Howell, S. C., Van Horn, W. D., Jeon, Y. H., & Sanders, C. R. (2009). Recent advances in the application of solution NMR spectroscopy to multi-span integral

membrane proteins. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *55*(4), 335–360.

Klepsch, M. M., Persson, J. O., & de Gier, J. W. L. (2011). Consequences of the overexpression of a eukaryotic membrane protein, the human KDEL receptor, in Escherichia coli. *Journal of Molecular Biology*, *407*(4), 532–42.

Knippers, R., & Hoffmann-Berling, H. (1966). A coat protein from bacteriophage fd. II. Interaction of the protein with DNA in vitro. *Journal of molecular biology*, *21*(2), 293–304.

Kobus, F. J., & Fleming, K. G. (2005). The GxxxG-containing transmembrane domain of the CCK4 oncogene does not encode preferential self-interactions. *Biochemistry*, *44*(5), 1464–1470.

Koehorst, R. B. M., Spruijt, R. B., Vergeldt, F. J., & Hemminga, M. A. (2004). Lipid bilayer topology of the transmembrane alpha-helix of M13 Major coat protein and bilayer polarity profile by site-directed fluorescence spectroscopy. *Biophysical journal*, *87*(3), 1445–1455.

Konings, R. N. H., Folmer, R. H. A., Folkers, J. M., Nilges, M., & Hilbers, C. W. (1995). Three-dimensional structure of the single-stranded DNA-binding protein encoded by gene V of the filamentous bacteriophage M13 and a model of its complex with single-stranded DNA. *FEMS Microbiology Reviews*, *17*, 57–72.

Konings, R., Ward, R., Francke, B., & Hofschneider, P. (1970). Gene order of RNA bacteriophage M12. *Nature*, *226*, 604–607.

Krishnan, J., Lee, G., & Choi, S. (2009). Drugs targeting Toll-like receptors. *Archives of Pharmacal Research*, *32*(11), 1485–502.

Krueger-Koplin, R. D., Sorgen, P. L., Krueger-Koplin, S. T., Rivera-Torres, I. O., Cahill, S. M., Hicks, D. B., & Girvin, M. E. (2004). An evaluation of detergents for NMR structural studies of membrane proteins. *Journal of biomolecular NMR*, *28*(1), 43–57.

Kuiken, C., & Simmonds, P. (2009). Hepatitis C: Methods and Protocols. (H. Tang, Ed.)*Hepatitis C: Methods and Protocols, Second Edition*, *510*, 33–53.

Kukar, T. L., Ladd, T. B., Robertson, P., Pintchovski, S. a, Moore, B., Bann, M. A, & Golde, T. E. (2011). Lysine 624 of the amyloid precursor protein (APP) is a critical determinant of amyloid β peptide length: support for a sequential model of γ-secretase intramembrane proteolysis and regulation by the amyloid β precursor protein (APP) juxtamembrane region. *The Journal of Biological Chemistry*, *286*(46), 39804–39812.

Laage, R., & Langosch, D. (2001). Strategies for Prokaryotic Expression of Eukaryotic Membrane Proteins. *Traffic*, *2*(2), 99–104.

Landry, Y., & Gies, J.P. (2008). Drugs and their molecular targets: an updated overview. *Fundamental & Clinical Pharmacology*, *22*(1), 1–18.

Lawrie, C. M., Sulistijo, E. S., & Mackenzie, K. R. (2010). Intermonomer Hydrogen Bonds Enhance GxxxG-Driven Dimerization of the BNIP3 Transmembrane Domain: Roles for Sequence Context in Helix–Helix Association in Membranes. *Journal of Molecular Biology*, *396*(4), 924–936.

Lazarova, T., Roberts, M. F., & Robinson, C. R. (2005). Oligomerization of the fifth transmembrane domain from the adenosine A 2A receptor. *Protein Science*, *14*, 2177–2186.

Lee, R., Doughty, S., Ashman, K., & Walker, J. (1996). Purification of hydrophobic integral membrane proteins from Mycoplasma hyopneumoniae by reversed-phase high-performance liquid chromatography. *Journal of Chromatography A*, *737*, 273–279.

Lemmon, M. A., & Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. *Cell*, *141*(7), 1117–1134.

Lemmon, M., Flanagan, J., & Hunt, J. (1992). Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices. *The Journal of Biological Chemistry*, *267*(11), 7683–7689.

Lemmon, M., Flanagan, J., & Treutlein, H. (1992). Sequence specificity in the dimerization of transmembrane. alpha.-helixes. *Biochemistry*, *31*, 12719–12725.

Lemmon, M., Treutlein, H., Adams, P. D., Brunger, A. T., & Engelman, D. M. (1994). A dimerization motif for transmembrane alpha-helices. *Nature Structural Biology*, *1*(3), 157–163.

Leonard, J. N., Ghirlando, R., Askins, J., Bell, J. K., Margulies, D. H., Davies, D. R., & Segal, D. M. (2008). The TLR3 signaling complex forms by cooperative receptor dimerization. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(1), 258–263.

Leong, J. M., Spring, C., Symp, H., Mackenzie, K. R., Prestegard, J. H., & Engelman, D. M. (1997). A transmembrane helix dimer : structure and implications. *Science*, *276*, 131–134.

Lesburg, C. A., Cable, M. B., Ferrari, E., Hong, Z., Mannarino, A. F., & Weber, P. C. (1999). Crystal structure of the polymerase from hepatitis C virus reveals a fully encircled active site. *Nature Structural Biology*, *6*(10), 937–943.

Levchenko, A. (2003). Dynamical and integrative cell signaling: challenges for the new biology. *Biotechnology and Bioengineering*, *84*(7), 773–82.

Li, E., & Hristova, K. (2006). Role of receptor tyrosine kinase transmembrane domains in cell signaling and human pathologies. *Biochemistry*, *45*(20), 6241–6251.

Li, W., Wimley, W. C., & Hristova, K. (2013). Transmembrane helix dimerization: beyond the search for sequence motifs. *Biochimica et biophysica acta*, *1818*(2), 183–193.

Lin, C., Lindenbach, B. D., Pragat, B. M., Mccourt, D. W., & Rice, C. M. (1994). Processing in the Hepatitis C Virus E2-NS2 Region : Identification of p7 and Two Distinct E2-Specific Products with Different C Termini. *Journal of Virology*, *68*(8), 5063–5073.

Liu, L., Botos, I., Wang, Y., Leonard, J. N., Shiloach, J., David, M., & Davies, D. R. (2009). Structural basis of Toll-Like Receptor 3 signaling with double-stranded DNA. *Science*, *320*(5874), 379–381.

Lohmann, V., Ko, F., Koch, J., Herian, U., & Theilmann, L. (1999). Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science*, *285*(5424), 110–114.

Luik, P., Chew, C., Aittoniemi, J., Chang, J., Wentworth, P., Dwek, R.A., Biggin, P.C., Venien-Bryan, C., & Zitzmann, N. (2009). The 3-dimensional structure of a hepatitis C virus p7 ion channel by electron microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(31), 1–5.

Ma, C. H. E., Marassi, F. M., Jones, D. H., Straus, S. K., Bour, S., Strebel, K., & Opella, S. J. (2002). Expression , purification , and activities of full-length and truncated versions of the integral membrane protein Vpu from HIV-1. *Protein Science*, *11*(3), 546–557.

MacKenzie, K., Prestegard, J., & Engelman, D. (1997). A transmembrane helix dimer: structure and implications. *Science*, *5309*(April), 131–134.

MacKenzie, K. R., Prestegard, J. H., & Engelman, D. M. (1997). A transmembrane helix dimer: structure and implications. *Science (New York, NY)*, *276*(5309), 131–133.

MacLennan, D. H., & Krania, E. G. (2003). Phospholamban: a crucial regulator of cardiac contracility. *Nature Reviews Molecular Cell Biology*, *4*, 566-577

Mancia, F., & Love, J. (2010). High-throughput expression and purification of membrane proteins. *Journal of structural biology*, *172*(1), 85–93.

Mao, L., Tang, Y., Vaiphei, S. T., Shimazu, T., Kim, S.-G., Mani, R., & Inouye, M. (2009). Production of membrane proteins for NMR studies using the condensed single protein (cSPP) production system. *Journal of structural and functional genomics*, *10*(4), 281–289.

Marassi, F. M., Ma, C., Gratkowski, H., Straus, S. K., Strebel, K., Montal, M., & Opella, S. J. (1999). Correlation of the structural and functional domains in the membrane protein Vpu from HIV-1. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(25), 14336–14341.

Marassi, F., Ramamoorthy, A., & Opella, S. J. (1997). Complete resolution of the solid-state NMR spectrum of a uniformly 15 N-labeled membrane protein in phospholipid bilayers. *Proceedings of the National Academy of Sciences of the United States of America*, *94*, 8551–8556.

Marciano, D. K., Russel, M., & Simon, S. M. (1999). An aqueous channel for filamentous phage export. *Science, 284*(May), 1516–1520.

Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. & Bax, A. (1989). Three - dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *Journal of the American Chemical Society, 111*(1515-1517)

Marsh, D. (2008). Energetics of hydrophobic matching in lipid-protein interactions. *Biophysical Journal*, *94*(10), 3996–4013.

Marvin, D. A. (1998). Filamentous phage structure , infection and assembly. *Current Opinion in Structural Biology*, *8*, 150–158.

Marvin, D., & Wachtel, E. (1976). Structure and assembly of filamentous bacterial viruses. *Philosphical Transactions of the Royal Society B: Biological Sciences*, *276*, 81–98.

Marvint, D. A., Halej, R. D., Cb, C., Nave, C., Wa, W., Citterich, M. H., & Vergata, T. (1994). Molecular models and structural comparisons of native an mutant class I filamentous bacteriophages Ff (fd, f1, M13), If1 and IKe. *Journal of Molecular Biology*, *235*, 260–286.

McDonnell, P. A., Shon, K., Kim, Y., & Opella, S. J. (1993). fd coat protein structure in membrane environments. *Journal of Molecular Biology*, *233*, 447–463.

Medzhitov, R. (2001). Toll-like receptors and innate immunity. *Nature Reviews Immunology*, *1*, 135–145.

Melnyk, R, Partridge, A., & Deber, C. (2002). Transmembrane domain mediated self-assembly of major coat protein subunits from Ff bacteriophage. *Journal of Molecular Biology*, *315*(1), 63–72.

Melnyk, R. a, Kim, S., Curran, a R., Engelman, D. M., Bowie, J. U., & Deber, C. M. (2004). The affinity of GXXXG motifs in transmembrane helix-helix interactions is modulated by long-range communication. *The Journal of Biological Chemistry*, *279*(16), 16591–16597.

Melnyk, RA, Partridge, A., Yip, J., Wu, Y., Goto, N. K., & Deber, C. M. (2003). Polar residue tagging of transmembrane peptides. *Biopolymers*, *71*(6), 675–685.

Membrane proteins of known structure determined by NMR. Retrieved September 30, 2015, from http://www.drorlist.com/nmr/MPNMR.html#alpha

Meyenburg, K. Von, Jorgensen, B. B., Michelsen, O., Sorensen, L., & Mccarthy, J. E. G. (1985). Proton conduction by subunit a of the membrane bound ATP synthase of Escherichia coli revealed after induced overproduction I. *The EMBO Journal*, *4*(9), 2357–2363.

Mineev, K S, Bocharov, E. V, Volynsky, P. E., Goncharuk, M. V, Tkach, E. N., & Ermolyuk, Y. S. (2011). Dimeric Structure of the Transmembrane Domain of Glycophorin A in Lipidic and Detergent Environments. *Acta Naturae*, *3*(9), 90–98.

Mineev, Konstantin S, Goncharuk, S. a, & Arseniev, A. S. (2014). Toll-like receptor 3 transmembrane domain is able to perform various homotypic interactions: An NMR structural study. *FEBS letters*, *588*(21), 3802–7.

Mineev, Konstantin S, Lesovoy, D. M., Usmanova, D. R., Goncharuk, S. a, Shulepko, M. a, Lyukmanova, E. N., & Arseniev, A. S. (2014). NMR-based approach to measure the free energy of transmembrane helix-helix interactions. *Biochimica et Biophysica Acta*, *1838*, 164–72.

Minton, A. P. (2000). Implications of macromolecular crowding for protein assembly. *Current opinion in structural biology*, *10*, 34–39.

Montero Vega, M. T., & de Andrés Martín, A. (2009). The significance of toll-like receptors in human diseases. *Allergologia et Immunopathologia*, *37*(5), 252–263.

Montserret, R., Saint, N., Vanbelle, C., Salvay, A. G., Simorre, J.-P., Ebel, C., & Penin, F. (2010). NMR structure and ion channel activity of the p7 protein from hepatitis C virus. *The Journal of biological chemistry*, *285*(41), 31446–31461.

Mouritsen, O. G., & Bloom, M. (1984). Mattress model of lipid-protein interactions in membranes. *Biophysical journal*, *46*(2), 141–53.

Nagler, C., Nagler, G., & Kuhn, A. (2007). Cysteine residues in the transmembrane regions of M13 procoat protein suggest that oligomeric coat proteins assemble onto phage progeny. *Journal of Bacteriology*, *189*(7), 2897–905.

Nakamura, M., Tsumoto, K., Kumagai, I., & Ishimura, K. (2003). A morphologic study of filamentous phage infection of Escherichia coli using biotinylated phages. *FEBS Letters*, *536*(1-3), 167–172.

Nazarov, P. V, Koehorst, R. B. M., Vos, W. L., Apanasovich, V. V, & Hemminga, M. A. (2006). FRET study of membrane proteins: simulation-based fitting for analysis of membrane protein embedment and association. *Biophysical Journal*, *91*(2), 454–466.

Nazarov, P. V, Koehorst, R. B. M., Vos, W. L., Apanasovich, V. V, & Hemminga, M. A. (2007). FRET study of membrane proteins: determination of the tilt and orientation of the N-terminal domain of M13 major coat protein. *Biophysical journal*, *92*(4), 1296–305.

Neill, L. A. J. O., Bryant, C. E., & Doyle, S. L. (2009). Therapeutic targeting of toll-like receptors for infectious and inflammatory diseases and cancer. *Pharmacological Reviews*, *61*(2), 177–197.

Ohi, N., Tokunaga, A., Tsunoda, H., Motoyama, N., & Nakajima, T. (1999). A novel adenovirus E1B19K-binding protein B5 inhibits apoptosis induced by Nip3 by forming a heterodimer through the C-terminal hydrophobic region. *Cell Death and Differentiaion*, *6*, 314–325.

Opalka, N., Beckmann, R., Boisset, N., Simon, M. N., Russel, M., & Darst, S. S. (2003). Structure of the Filamentous Phage pIV Multimer by Cryo-electron Microscopy. *Journal of Molecular Biology*, *325*(3), 461–470.

OuYang, B., Xie, S., Berardi, M. J., Zhao, X., Dev, J., Yu, W., & Chou, J. J. (2013a). Unusual architecture of the p7 channel from hepatitis C virus. *Nature*, *498*(7455), 521–525.

Oxenoid, K., & Chou, J. J. (2005). The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(31), 10870–10875.

Papavoine, C H, Konings, R. N., Hilbers, C. W., & van de Ven, F. J. (1994). Location of M13 coat protein in sodium dodecyl sulfate micelles as determined by NMR. *Biochemistry*, *33*(44), 12990–12997.

Papavoine, C., Christiaans, B., Forlmer, R. H., Konings, R. N., & Hilbers, C. W. (1998). Solution structure of the M13 major coat protein in detergent micelles: a basis for a model of phage assembly involving specific residues. *Journal of molecular biology*, *282*, 401–419.

Papavoine, Christina HM, Aelen, J. M., Konings, R. N., Hilbers, C. W., & Van de Ven, F. J. (1995). NMR studies of the major coat protein of bacteriophage M13 Structural information of gVIIIp in dodecylphosphocholine micelles. *European Journal of Biochemistry*, *232*, 490–500.

Park, B. S., Song, D. H., Kim, H. M., Choi, B.-S., Lee, H., & Lee, J.-O. (2009). The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature*, *458*(7242), 1191–1195.

Park, S H, Son, W. S., Mukhopadhyay, R., Valafar, H., & Opella, S. (2009). Phage-induced alignment of membrane proteins enables the measurement and structural analysis of residual dipolar couplings with dipolar waves and lambda-maps. *Journal of the American Chemical Society*, *131*(40), 14140–14141.

Park, Sang Ho, Mrse, A. a., Nevzorov, A. a., Mesleh, M. F., Oblatt-Montal, M., Montal, M., & Opella, S. J. (2003). Three-dimensional structure of the channel-forming trans-membrane domain of Virus Protein "u" (Vpu) from HIV-1. *Journal of Molecular Biology*, *333*(2), 409–424.

Passos-Bueno, M., Wilcox, W. R., Jabs, E. W., Sertié, A. L., Alonso, L. G., & Kitoh, H. (1999). Clinical spectrum of fibroblast growth factor receptor mutations Clinical Spectrum of Fibroblast Growth Factor Receptor Mutations. *Human Mutation*, *14*, 115–125.

Pavlovic, D., Neville, D., Argaud, O., Blumberg, B., Dwek, R. A., Fischer, W. B., & Zitzmann, N. (2003). The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(10), 6104–6108.

Pawlotsky, J.M. (2004). Pathophysiology of hepatitis C virus infection and related liver disease. *Trends in Microbiology*, *12*(2), 96–102.

Penin, F. (2003). Structural biology of hepatitis C virus Franc. *Clin Liver Dis*, *7*(1), 1–21.

Pinto, L., Dieckmann, G., Gandhi, C., Papworth, C., Braman, J., Shaughnessy, M., & DeGrado, W. (1997). A functionally defined model for the M2 proton channel of influenza A virus suggests a mechanism for its ion selectivity. *Proceedings of the National Academy of Sciences of the United States of America*, *94*, 11301–11306.

Pinto, L. H., & Lamb, R. a. (2006). The M2 proton channels of influenza A and B viruses. *The Journal of biological chemistry*, *281*(14), 8997–9000.

Poget, S. F., & Girvin, M. E. (2007). Solution NMR of membrane proteins in bilayer mimics: small is beautiful, but sometimes bigger is better. *Biochimica et Biophysica Acta*, *1768*(12), 3098–3106.

Poordad, F., & Dieterich, D. (2012). Treating hepatitis C: current standard of care and emerging direct-acting antiviral agents. *Journal of Viral Hepatitis*, *19*(7), 449–464.

Premkumar, A., Wilson, L., Ewart, G., & Gage, P. (2004). Cation-selective ion channels formed by p7 of hepatitis C virus are blocked by hexamethylene amiloride. *FEBS Letters*, *557*(1-3), 99–103.

Qureshi, T., & Goto, N. K. (2012). Contemporary methods in structure determination of membrane proteins by solution NMR. *Topics in Current Chemistry*, *326*, 123-186.

Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G., & Deber, C. M. (2009). Detergent binding explains anomalous SDS-PAGE. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(6), 1760–1765.

Ray, R., Chen, G., Velde, C. Vande, Cizeau, J., Park, J. H., Reed, J. C., & Greenberg, A. H. (2000). BNIP3 heterodimerizes with Bcl-2 / Bcl-X L and induces cell death independent of a Bcl-2 Homology 3 (BH3) domain at both mitochondrial and nonmitochondrial sites. *The Journal of Biological Chemistry*, *275*(2), 1439–1448.

Reddy, L. G., Jones, L. R., & Thomas, D. D. (1999). Depolymerization of phospholamban in the presence of calcium pump : a fluorescence energy transfer study. *Biochemistry*, *38*, 3954–3962.

Rharbi, Y., & Winnik, M. A. (2001). Solute exchange between surfactant micelles by micelle fragmentation and fusion. *Advances in Colloid and Interface Science*, *89-90*, 25–46.

Robertson, S. C., Tynan, J. A., & Donoghue, D. J. (2000). RTK mutations and human syndromes when good receptors turn bad. *Trends in Genetics*, *16*(6), 265–271.

Roth, T. A., Weiss, G. A., Eigenbrot, C., & Sidhu, S. S. (2002). A minimized M13 coat protein defines the requirements for assembly into the bacteriophage particle. *Journal of Molecular Biology*, *322*(2), 357–367.

Rowe, S. M., Miller, S., & Sorscher, E. J. (2005). Cystic Fibrosis. *The New England Journal of Medicine*, *352*, 1992–2001.

Rule, G. S. & Hitchens, T. K. (2006). Fundamentals of Protein NMR Spectroscopy (Focus on Structural Biology), Springer, Netherlnds, 411-414

Russ, W. P., & Engelman, D. M. (2000). The GxxxG motif: a framework for transmembrane helix-helix association. *Journal of Molecular Biology*, *296*(3), 911–919.

Russmann, S., Grattaglian, P., Palmieri, V., & Palasciano, G. (2006). Ribavirin-induced anemia: mechanisms, risk factors and related targets for future research. *Current Medicinal Chemistry*, *13*, 3351–3357.

Sakai, A., Claire, M. S., Faulk, K., Govindarajan, S., Emerson, S. U., Purcell, R. H., & Bukh, J. (2003). The p7 polypeptide of hepatitis C virus is critical for infectivity and contains functionally important genotype-specific sequences. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(20), 11646–11651.

Sanders, C R, & Sönnichsen, F. (2006). Solution NMR of membrane proteins: practice and challenges. *Magnetic Resonance in Chemistry : MRC*, *44*(S1), S24–S40.

Sanders, Charles R, & Myers, J. K. (2004). Disease-related misassembly of membrane proteins. *Annual Review of Biophysics and Biomolecular Structure*, *33*, 25–51.

Sanguinetti, M. C., & Tristani-Firouzi, M. (2006). hERG potassium channels and cardiac arrhythmia. *Nature*, *440*(7083), 463–9.

Sarabipour, S., Del Piccolo, N., & Hristova, K. (2015). Characterization of Membrane Protein Interactions in Plasma Membrane Derived Vesicles with Quantitative Imaging Förster Resonance Energy Transfer. *Accounts of chemical research*, *48*(8), 2262–2269.

Sato, T., Kawakami, T., Akaji, K., Konishi, H., Mochizuki, K., Fujiwara, T., & Aimoto, S. (2002). Synthesis of a membrane protein with two transmembrane regions. *Journal of Peptide Science*, *8*(4), 172–180.

Sattler, M. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *34*(2), 93–158.

Schinazi, R., Halfon, P., Marcellin, P., & Asselah, T. (2014). HCV direct-acting antiviral agents: the best interferon-free combinations. *Liver international : Official Journal of the International Association for the Study of the Liver*, *34 Suppl 1*, 69–78.

Schlessinger, J. (1992). Growth factor signaling by receptor tyrosine kinases. *Neuron*, *9*, 383–391.

Schlessinger, Joseph. (2000). Cell signaling by receptor tyrosine kinases A large group of genes in all eukaryotes encode for. *Cell*, *103*, 211–225.

Schneider, D., Engelman, D. M., & Haven, N. (2004). Motifs of two small residues can assist but are not sufficient to mediate transmembrane helix interactions. *Journal of molecular biology*, *343*, 799–804.

Schneider, W. M., Inouye, M., Montelione, G. T., & Roth, M. J. (2009). Independently inducible system of gene expression for condensed single protein production (cSPP) suitable for high efficiency isotope enrichment. *Journal of Structural and functional Genomics*, *10*(3), 219–225.

Schneider, W. M., Tang, Y., Vaiphei, S. T., Mao, L., Maglaqui, M., Inouye, M., & Montelione, G. T. (2010). Efficient condensed-phase production of perdeuterated soluble and membrane proteins. *Journal of Structural and Functional Genomics*, *11*(2), 143–54.

Seddon, A. M., Curnow, P., & Booth, P. J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta*, *1666*(1-2), 105–17.

Segrest, J. P., De Loof, H., Dohlman, J. G., Brouillette, C. G., & Anantharamaiah, G. (1990). Amphipathic Helix Motif : Classes and Properties. *Proteins: Structure, Function, and Bioinformatics*, *117*, 103–117.

Senes, A, Gerstein, M., & Engelman, D. M. (2000). Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *Journal of Molecular Biology*, *296*(3), 921–936.

Senes, Alessandro, Ubarretxena-belandia, I., & Engelman, D. M. (2001). The Ca- HO hydrogen bond : A determinant of stability and specificity in transmembrane helix interactions. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(16), 9056–9061.

Sengupta, P., Hammond, A., Holowka, D., & Baird, B. (2008). Structural determinants for partitioning of lipids and proteins between coexisting fluid phases in giant plasma membrane vesicles. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1778*(1), 20–32.

Shepard, C. W., Finelli, L., & Alter, M. J. (2005). Global epidemiology of hepatitis C virus infection. *Lancet Infectious Disease*, *5*(9), 558–567.

Shimizu, S., & Tsujimoto, Y. (2000). Proapoptotic BH3-only Bcl-2 family members induce cytochrome c release, but not mitochondrial membrane potential loss, and do not directly modulate voltage-dependent anion channel activity. *Biochemistry*, *97*(2), 577–582.

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

Simmerman, H. K. B., Kobayashi, Y. M., Autry, J. M., & Jones, L. R. (1996). A leucine zipper stabilizes the pentameric membrane domain of phospholamban and forms a

coiled-coil pore structure. *The Journal of Biological Chemistry*, *271*(10), 5941–5946.

Simmonds, P. (2004). Genetic diversity and evolution of hepatitis C virus--15 years on. *The Journal of General Virology*, *85*(Pt 11), 3173–3188.

Smith, S. O., Song, D., Shekar, S., Groesbeek, M., Ziliox, M., & Aimoto, S. (2001). Structure of the transmembrane dimer interface of glycophorin A in membrane. *Biochemistry*, *40*(22), 6553–6558.

Snyder, D. A., Chen, Y., Denissova, N. G., Acton, T., Aramini, J. M., Ciano, M., & Montelione, G. T. (2005). Comparisons of NMR spectral quality and success in crystallization demonstrate that NMR and X-ray crystallography are complementary methods for small protein structure determination, *127*, 16505–16511.

Sørensen, H. P., & Mortensen, K. K. (2005). Advanced genetic strategies for recombinant protein expression in Escherichia coli. *Journal of Biotechnology*, *115*(2), 113–28.

Spruijt, R. B., Wolfs, C. J., & Hemminga, M. A. (1989). Aggregation-related conformational change of the membrane-associated coat protein of bacteriophage M13. *Biochemistry*, *28*(23), 9158–9165.

Spruijt, R., & Meijer, A. (2000). Localization and rearrangement modulation of the N-terminal arm of the membrane-bound major coat protein of bacteriophage M13. *Biochimica et Biophysica Acta 1509*, 311–323.

Sreerama, N, & Woody, R. (1994). Protein secondary structure from circular dichroism spectroscopy. Combining variable selection principle and cluster analysis with neural network, ridge regression and self-consistent methods. *Journal of molecular biology*, *242*(4), 497–507.

Sreerama, Narasimha, & Woody, R. W. (2004). Computation and Analysis of Protein Circular Dichroism Spectra. *Methods in Enzymology*, *383*(1), 318–351.

Stauber, R. E., & Stadlbauer, V. (2006). Novel approaches for therapy of chronic hepatitis C. *Journal of Clinical Virology : The Official Publication of the Pan American Society for Clinical Virology*, *36*(2), 87–94.

Steinmann, E., Penin, F., Kallis, S., Patel, A. H., Bartenschlager, R., & Pietschmann, T. (2007). Hepatitis C virus p7 protein is crucial for assembly and release of infectious virions. *PLoS Pathogens*, *3*(7), 0962-0971.

Stejskal, E., & Tanner, J. (1965). Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*, *42*(1), 288–292.

Stopar, D., Spruijt, R. B., Wolfs, C. J., & Hemminga, M. A. (2003). Protein-lipid interactions of bacteriophage M13 major coat protein. *Biochimica et Biophysica Acta*, *1611*(1-2), 5–15.

Sugrue, R., & Hay, A. (1991). Structural characteristics of the M2 protein of influenza A viruses: evidence that it forms a tetrameric channel. *Virology*, *180*, 617–624.

Suh, Y., & Checler, F. (2002). Amyloid precursor protein, presenilins, and a -synuclein: molecular pathogenesis and pharmacological applications in alzheimer's disease. *Pharmacology Reviews*, *54*(3), 469–525.

Sulistijo, E S, & Mackenzie, K. R. (2009). Structural basis for dimerization of the BNIP3 transmembrane domain. *Biochemistry*, *48*(23), 5106–5120.

Sulistijo, E. S., Jaszewski, T. M., & MacKenzie, K. R. (2003). Sequence-specific dimerization of the transmembrane domain of the "BH3-only" protein BNIP3 in membranes and detergent. *Journal of Biological Chemistry*, *278*(51), 51950–51956.

Sulistijo, Endah S, Jaszewski, T. M., & MacKenzie, K. R. (2003). Sequence-specific dimerization of the transmembrane domain of the "BH3-only" protein BNIP3 in membranes and detergent. *The Journal of Biological Chemistry*, *278*(51), 51950–51956.

Sulistijo, Endah S, & MacKenzie, K. R. (2006). Sequence dependence of BNIP3 transmembrane domain dimerization implicates side-chain hydrogen bonding and a tandem GxxxG motif in specific helix-helix interactions. *Journal of Molecular Biology*, *364*(5), 974–990.

Suzuki, M., Zhang, J., Liu, M., Woychik, N. a., & Inouye, M. (2005). Single protein production in living cells facilitated by an mRNA interferase. *Molecular Cell*, *18*(2), 253–261.

Takami, M., Nagashima, Y., Sano, Y., Ishihara, S., Morishima-Kawashima, M., Funamoto, S., & Ihara, Y. (2009). Gamma-secretase: successive tripeptide and tetrapeptide release from the transmembrane domain of beta-carboxyl terminal fragment. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(41), 13042–13052.

Traaseth, N. J., Verardi, R., Torgersen, K. D., Karim, C. B., Thomas, D. D., & Veglia, G. (2007). Spectroscopic validation of the pentameric structure of phospholamban. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(37), 15676–14681.

Vaiphei, S. T., Tang, Y., Montelione, G. T., & Inouye, M. (2011). The use of the condensed single protein production system for isotope-labeled outer membrane proteins, OmpA and OmpX in E. coli. *Molecular Biotechnology*, *47*(3), 205–210.

Van de Ven, F. J., van Os, J. W., Aelen, J. M., Wymenga, S. S., Remerowski, M. L., Konings, R. N., & Hilbers, C. W. (1993). Assignment of 1H, 15N, and backbone 13C resonances in detergent-solubilized M13 coat protein via multinuclear multidimensional NMR: a model for the coat protein monomer. *Biochemistry*, *32*(32), 8322–8328.

Van Horn, W. D., Kim, H.-J., Ellis, C. D., Hadziselimovic, A., Sulistijo, E. S., Karra, M. D., & Sanders, D. (2009). Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. *Science*, *324*, 1726–1729.

Van Horn, W., Kim, H., Ellis, C., Hadziselimovic, A., Sulistijo, E., Karra, M., & Sanders, C. (2009). Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. *Science*, *324*(5935), 1726–1729.

Verardi, R., Shi, L., Traaseth, N. J., Walsh, N., & Veglia, G. (2011). Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(22), 9101–9106.

Vos, W. L., Nazarov, P. V., Koehorst, R. B., Spruijt, R. B., & Hemminga, M. A. (2009). From "I" to "L" and back again: the odyssey of membrane-bound M13 protein. *Trends in biochemical sciences*, *34*(5), 249–255.

Wagner, S., Baars, L., Ytterberg, a J., Klussmeier, A., Wagner, C. S., Nord, O., Nygren, P-A., van Wijk, K. J., & de Gier, J.-W. (2007). Consequences of membrane protein overexpression in Escherichia coli. *Molecular & Cellular Proteomics : MCP*, *6*(9), 1527–1550.

Wagner, S., Bader, M. L., Drew, D., & de Gier, J.-W. (2006). Rationalizing membrane protein overexpression. *Trends in Biotechnology*, *24*(8), 364–371.

Wang, C., & Deber, C. M. (2000). Peptide mimics of the M13 coat protein transmembrane segment. Retention of helix-helix interaction motifs. *The Journal of Biological Chemistry*, *275*(21), 16155–16159.

Wang, D.-N., Safferling, M., Lemieux, M. J., Griffith, H., Chen, Y., & Li, X.-D. (2003). Practical aspects of overexpressing bacterial secondary membrane transporters for structural studies. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1610*(1), 23–36.

Wang, Y. a, Yu, X., Overman, S., Tsuboi, M., Thomas, G. J., & Egelman, E. H. (2006). The structure of a filamentous bacteriophage. *Journal of Molecular Biology*, *361*(2), 209–215.

Welling, G. W., Zee, R. V. D. Z., & Welling-Wester, S. (1987). Column liquid chromatography of integral membrane proteins. *Journal of Chromatography*, *418*, 223–243.

Welsch, C., Jesudian, A., Zeuzem, S., & Jacobson, I. (2012). New direct-acting antiviral agents for the treatment of hepatitis C virus infection and perspectives. *Gut*, *61 Suppl 1*(Suppl 1), i36–46.

Wickner, W. (1975). Asymmetric orientation of a phage coat protein in cytoplasmic membrane of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(12), 4749–4753.

Williams, K. A., Glibowicka, M., Li, Z., Li, H., Khan, A. R., Chen, Y. M., & Deber, C. M. (1995). Packing of coat protein amphipathic and transmembrane helices in filamentous bacteriophage M13: role of small residues in protein oligomerization. *Journal of molecular biology*, *252*(1), 6–14.

Wishart, D. S., Sykes, B. D., & Richard, F. M. (1991). Relationship between Nuclear Magnetic Resonance Chemical Shift and Protein Secondary Structure. *Journal of Molecular Biology*, *222*(2), 311–333.

Wishart, D., & Sykes, B. D. (1994). The 13C Chemical-Shift Index : A simple method for the identification of protein secondary structure using. *Journal of Biomolecular NMR*, *4*(2), 171–180.

World Health Organizaition - Hepatitis C. (n.d.). Retrieved July 12, 2015, from http://www.who.int/mediacentre/factsheets/fs164/en/

Wu, Y., Shih, S. C. C., & Goto, N. K. (2007). Probing the structure of the Ff bacteriophage major coat protein transmembrane helix dimer by solution NMR. *Biochimica et Biophysica Acta - Biomembranes*, *1768*(12), 3206–3215.

Yasuda, M., Theodorakis, P., Subramanian, T., & Chinnadurai, G. (1998). Adenovirus E1B-19K / BCL-2 Interacting Protein BNIP3 Contains a BH3 Domain and a Mitochondrial Targeting Sequence. *The Journal of Biological Chemistry*, *273*(20), 12415–12421.

Yee, A. A., Savchenko, A., Ignachenko, A., Lukin, J., Xu, X., Skarina, T., Edvokimova, E., Liu, C. S., Semesi, A., Guido, V., & Arrowsmith, C. H. (2005). NMR and x-ray crystallography , complementary tools in structural proteomics of small proteins. *Journal of the American Chemical Society, 127*(4), 16512–16517.

Yuen, C. T. K., Davidson, A. R., & Deber, C. M. (2000). Role of aromatic residues at the lipid - water interface in micelle-bound bacteriophage M13 major coat protein. *Biochemistry*, *12*, 16155–16162.

Zamoon, J., Mascioni, A., Thomas, D. D., & Veglia, G. (2008). NMR solution structure and topological orientation of monomeric phospholamban in dodecylphosphocholine micelles. *Biophysical Journal*, *85*(4), 2589–2598.

Zeri, C., Mesleh, M. F., Nevzorov, A. A., & Opella, S. J. (2003). Structure of the coat protein in fd filamentous bacteriophage particles determined by solid-state NMR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(11), 6458–6493.

Zhou, Y., Cierpicki, T., Jimenez, R. H., Lukasik, S. M., Ellena, J. F., Cafiso, D. S., Kadokura, H., Beckwith, J., & Bushweller, J. H. (2008). NMR solution structure of the integral membrane enzyme DsbB: functional insights into DsbB-catalyzed disulfide bond formation. *Molecular cell*, *31*(6), 896–908.

List of Publications

1. Qureshi T. & N.K. Goto (2015) Non-ideality of SDS can promote transmembrane helix-helix interactions **(Manuscript in preparation)**

   Statement of Contribution: All experiments, except for obtaining CBCA(CO)NH, HNCACB and ^{15}N-NOESY-HSQC, which were graciously carried out by Dr. Suave and Dr Aubin, were carried out by me. Analysis of all data was carried out by me. I am writing this paper with Dr. Goto.

2. Houman, G., Ka-Cheong Kwok, J.K.-C., Sherratt A.R., Foo, A.C.Y., Qureshi T., & N.K. Goto (2014) Micelle-catalyzed domain swapping in the GlpG rhomboid protease cytoplasmic domain. *Biochemistry*, 53:5907-5915

   Statement of Contribution: My contribution to this co-author paper was to carry out NMR experiment for ^{13}C- NOESY-HSQC data.

3. Qureshi, T. & N.K. Goto (2012) "Contemporary methods in membrane protein structure determination by solution NMR" *G. Zhu ed Topics in Current Chemistry Springer*, 326:123-185

   Statement of Contribution: I wrote this book chapter with Dr. Goto.

Curriculum Vitae

Tabussom Qureshi

EDUCATION

Doctorate (PhD – Chemistry) Sept 2007-2016
University of Ottawa (Ottawa, ON, Canada)

Honours BSc (Biopharmaceuticals, Medicinal Chemistry) 2003-2007
University of Ottawa (Ottawa, ON, Canada)

PUBLICATIONS:

1. Qureshi T. & N.K. Goto. "Non-ideality of SDS can promote transmembrane helix-helix interactions" **(Manuscript in preparation)**
2. Houman, G., Ka-Cheong Kwok, J.K.-C., Sherratt A.R., Foo, A.C.Y., Qureshi T., & N.K. Goto. "Micelle-catalyzed domain swapping in the GlpG rhomboid protease cytoplasmic domain" in *Biochemistry*, 53:5907-5915, published in 2014
3. Qureshi, T. & N.K. Goto, "Contemporary methods in membrane protein structure determination by solution NMR" in *G. Zhu ed Topics in Current Chemistry Springer*, 326:123-185, published in 2012

ORAL PRESENTATIONS:

1. "Modulation of TM Helix Oligomerization by Amphipathic-Helix Membrane Interactions". Biomolar Seminar, November 2012
2. "Modulation of TM Helix Oligomerization by Amphipathic-Helix Membrane Interactions". MOOT Conference, University of Laval, October 2012
3. "Diffusion coefficient measurements of protein-detergent complexes.", Undergraduate Thesis Seminar, University of Ottawa, May 2007
4. "Monitoring cellular functions via *in vivo* fluorescent imaging.", Graduate Seminar, University of Ottawa, February 2009

POSTER PRESENTATIONS:

1. **Tabussom Qureshi** and Natalie K. Goto. Modulation of TM Helix Oligomerization by Amphipathic Helix-Membrane Interactions. Methods in Protein Structure Analysis (MPSA), Ottawa, Ontario (Jun 26-28, 2012)
2. **Tabussom Qureshi** and Natalie K. Goto. Modulation of TM Helix Oligomerization by Amphipathic Helix-Membrane Interactions. Groupe de Recherche Axé sur la Structure des Protéines (GRASP) Symposium, Montreal, Quebec (Nov 21, 2011)