Structural and functional characterization of the human respiratory syncytial virus small hydrophobic protein

Gan, Siok Wan

2010

Gan, S. W. (2010). Structural and functional characterization of the human respiratory syncytial virus small hydrophobic protein. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/42676

https://doi.org/10.32657/10356/42676

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE HUMAN RESPIRATORY SYNCYTIAL VIRUS SMALL HYDROPHOBIC PROTEIN

GAN SIOK WAN

School of Biological Sciences

A thesis submitted to the Nanyang Technological University
In partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2010

Acknowledgements

I would like to express my sincere appreciation to my supervisor, Dr. Jaume Torres for being an exceptional mentor. This thesis would not be complete without his generous support, encouragement, and scientific guidance. Despite many difficulties that I have faced during this project, he always encourages me to try many different things and techniques to go through the end of the project.

I wish to extent my sincere thanks to our collaborators that have helped me with experiments along the way: Dr. Alex Gong and Ms. Ng LiFang for planar lipid bilayers recordings; Dr. Soong Tuck Wah and Ms. Yu DeJie for patch clamp experiment; Dr. Julian Lescar and Mr. Luo DaHai for x-ray crystallography; Dr. Liu Chuan Fa and Mr. Yang Ren Liang for technical assistant in native chemical ligation process; Dr. Konstantin Pervushin and Mr. Edward Tan for NMR structure determination.

I would like to thank all the members in our lab: Dr. Ardcharaporn Varanatanavet, Dr. Lin Xin, Parthasarathy Krupakar, and Mukesh Mahajan, project students and all our previous colleagues for their friendship and helpful discussion.

I would also like to thank Ian Brett for sharing the plasmids and protocols for membrane proteins expression and purification. Thanks also go to Eliza for helping with the peptide synthesis, and Kuang Lung for helpful discussion on NMR peaks assignment.

Last but not least, this work would not have been possible without the help and support of my friends and family. I thank my mother for her encouragement and great love. I thank my sisters and brothers, Sin Hui, Shiau Ping, Chee Tian, and Chee Min, who are always available for support and encouragement. And my nieces Zhi Yi and Yong Han, and my nephew Yong Le, who provide endless entertainment and remind me of the things that are truly important. My uncle and aunt, Mr. Tan Seng Lip and Madam Lau Kwee Cheng for their care, support and encouragement.

Table of Contents

Abstract

The small hydrophobic (SH) protein is a transmembrane surface glycoprotein encoded by the respiratory syncytial virus (RSV). It is 64 amino acids long with one putative transmembrane domain. Although SH protein is important for viral infectivity, its exact role during viral infection is not clear. In this study, we have examined the structure, oligomerization, and function of SH protein and the transmembrane domain (SH-TM) using biochemical, biophysical and computational approaches.

We have shown that SH is able to assemble into pentamers and the process is driven by the transmembrane domain. We have carried out a thermodynamic study using sedimentation equilibrium analytical ultracentrifugation to determine the propensity of SH to self-associate in C14SB, C8E5, and DPC micelles. SH pentamerization is most favorable in C8E5, with standard free energy of association of -24.89 kcal/mol, and a conserved histidine at the transmembrane domain is important for the pentamerization specificity.

Investigation with polarized attenuated total reflection-Fourier transform infrared (PATR-FTIR) shows that the SH-TM is predominantly α-helical and the C-terminal part of SH is mostly β structure in planar lipid bilayers. The rotational orientation of SH-TM was determined by site specific infrared dichroism (SSID) at two consecutive isotopically labeled residues. This orientation is consistent with that of an evolutionary conserved pentameric model obtained from a global search protocol using thirteen homologous sequences of RSV.

Finally, conductance studies of SH and the transmembrane domain indicate a pH sensitive ion channel activity, presumably gated by His51. Thus, our results provide experimental evidence that SH is a pentameric ion channels.

List of Figures

List of Tables

Abbreviations

Amp	Ampicillin
ATR-FTIR	Attenuated total reflectance Fourier transform infrared spectroscopy
AUC-SE	Analytical ultracentrifugation
BLM	Black lipid membrane
Boc	*Tert*-butoxycarbonyl
C8E5	Pentaoxyethylene octyl ether
C14SB	3-(N,N-dimethylmyristyl-ammonio)propanesulfonate
Chl	Chloramphenicol
CMC	Critical micelle concentration
DHPC	1,2-Dihexanoyl-*sn*-glycero-3-phosphocholine
DMF	N,N- Dimethylformamide
DMPC	1,2-Dimyristoyl-*sn*-glycero-3-phosphocholine
DNA	Deoxyribonucleic acid
DPC	Dodecyl phosphocholine
DSS	2, 2-dimethyl-2-silapentane-5-sulfonate
DTT	Dthiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetraacetic acid
EPR	Electron paramagnetic resonance
FRET	Fluorescent resonance energy transfer

FMOC	N-9-fluorenyl methyloxycarbonyl
GFP	Green fluorescent protein
GSMD	Global search molecular dynamics
HCV	Hepatitis C virus
HEK	Human embryonic kidney cells
HF	Hydrogen fluoride
HFIP	Hexa fluoro isopropanol
HIV-1	Human immunodeficiency virus-type 1
HMA	Hexamethyleneamiloride
hRSV	Human respiratory syncytial virus
HSQC	Heteronuclear single quantum coherence
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Infrared
LB	Luria-Bertani
LMPG	Lyso-myristoyl phosphatidylglycerol
LPPG	Lyso-palmitoyl phosphatidylglycerol
MALDI	Matrix-assisted laser desorption/ionization
MBHA	4-methyl-benzhydrylamine
MBP	Maltose binding protein
MW	Molecular weight
NCL	Native chemical ligation
Ni^{2+}-NTA	Nickel-nitrilotriacetic acid

NMM	N-Methyl morpholine
NMR	Nuclear magnetic resonance
β-OG	Octyl-β-D-glucopyranoside
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein data bank
PFO	Perfluorooctanoic acid
POPC	1-Palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine
POPE	1-Palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine
POPS	1-Palmitoyl-2-oleoyl-*sn*-glycero-3-[Phospho-L-serine]
ppm	Parts per million
PyBop	Benzotriazole-1-yl-oxy-tris-pyrrolidinophosphonium hexafluorophosphate
RMSD	root mean square deviation
RP-HPLC	Reverse phase-high pressure liquid chromatography
SARS	Severe acute respiratory syndrome
SAX	Small-angle X-ray scattering
SE	Sedimentation equilibrium
SDS	Sodium dodecyl sulfate
SH	Small hydrophobic
SHprot	Small hydrophobic protein with additional three amino acids at the N-terminal, SNA
SH-TM	Small hydrophobic protein transmembrane domain
SSID	Site specific infrared spectroscopy
SSPS	Solid Phase Peptide Synthesis

TCA	Trichloroacetic acid
TEV	Tobacco etch virus
TFA	Trifluoro acetic acid
TM	Transmembrane
TROSY	Transverse relaxation optimized spectroscopy
UV	Ultra violet

1) INTRODUCTION

The goal of this study is to characterize the small hydrophobic (SH) protein, a transmembrane surface glycoprotein encoded by the human respiratory syncytial virus (hRSV), both functionally and structurally. SH protein is a small, conserved integral membrane protein consisting of 64 amino acids and a single transmembrane α-helix. The role played by SH protein in hRSV life cycle and during infection is not well characterized. In this study we show that SH is a pentameric ion channel using various biochemical, biophysical and computational approaches. SH protein may be a potential drug target to treat hRSV infection.

1.1) Membrane Proteins

Membrane proteins play important functions in cellular process. They act as receptors and signal transducers on the surface of cell membrane to allow for communication between cells and their surroundings. In a way they are the regulator of the intracellular environment. The passage of ions and molecules across the membrane is controlled by ion channel formed by membrane proteins. Vital cellular processes such as cell assembly, fusion of cells or viruses require specific membrane proteins to take part. These activities of membrane proteins make them an interesting target for drug design strategies. For example, inhibition or controlling these activities could stop the infection process of bacteria or viruses and combat human diseases caused by defect in receptor activity.

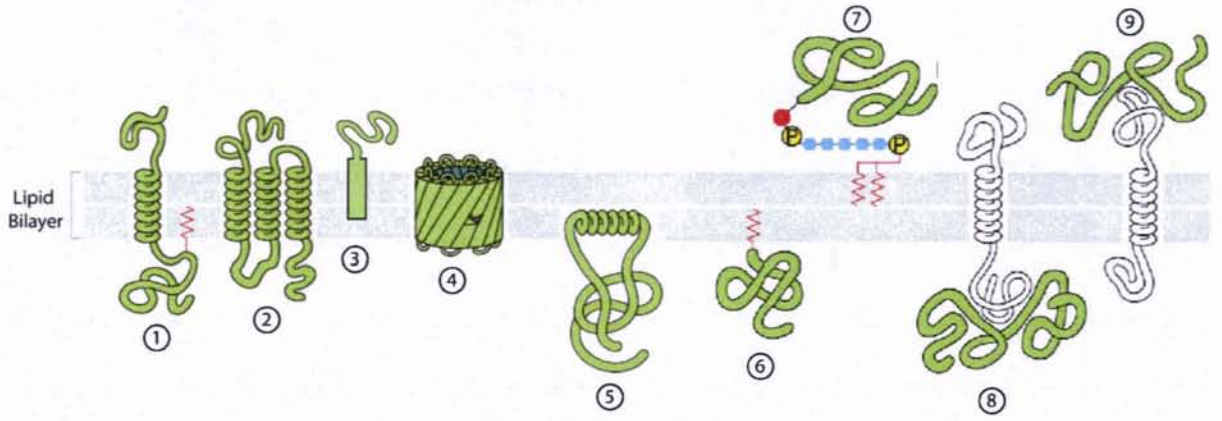


Figure 1-1. Schematic diagram of several types of membrane proteins in biological membrane. Intrinsic membrane proteins (1-5): 1) bitopic, 2) polytopic, 3&5) monotopic, 4) β-barrel. Extrinsic membrane protein (6-9): 6) anchored by a fatty acid, 7) anchored by GPI, 8&9) non-covalently attachment to integral membrane proteins. The picture is adapted from Molecular Biology of The Cell (Alberts et al. 2002).

Membrane proteins are classified into two classes, integral (or intrinsic) and peripheral (or extrinsic) based on their interactions with the cell membrane (see Fig. 1-1). Integral membrane proteins have one or more transmembrane domains inserted into the hydrophobic core of the lipid bilayer, and can be removed only through the use of detergent or denaturants. Integral membrane protein can be divided into three groups based on the way they inserted into the lipid bilayer: monotopic, bitopic, and polytopic. Monotopic proteins are inserted in the membrane but do not pass through the membrane. Bitopic and polytopic proteins span the lipid bilayer once or multiple times, respectively.

In contrast, peripheral membrane proteins do not penetrate the lipid bilayer, they interact with the cell membrane in two ways: non-covalently attachment to the polar head group of the lipid bilayer or the integral membrane proteins; or through covalent attachment to either a fatty acid or a glycolipid headgroup. Peripheral membrane proteins can be extracted from the lipid bilayer without using detergent or denaturant and can be treated in the same way as soluble proteins.

The lipid bilayer is a native environment for membrane proteins. It consists of amphipathic phospholipids with a hydrophilic headgroup and hydrophobic fatty acid tail. They associate with each other forming a bilayer with the hydrophobic part facing the center and the polar headgroup exposing to the aqueous environment. Typically, the thickness hydrophobic core of lipid bilayer is about 30Å and forms a low dielectric environment. These properties of lipid bilayer impose certain structural features to integral membrane proteins, in particular bitopic and polytopic groups that have to span the bilayer reaching both side of the membrane. Two types of secondary structure have been observed in integral membrane proteins, α-helices and β-strands. Both structures have a regular array of hydrogen bonding between the peptide backbones to satisfy hydrogen bonds in the low dielectric environment where van-der-waals and electrostatic interactions dominate. Most integral membrane proteins cross the bilayer as α-helix, and they have a broader range of functionalities and complexities. β-strands are found in outer membrane of gram-negative bacteria, mitochondria, and chloroplasts, forming a rigid pore called β-barrels.

The low dielectric environment of lipid bilayer also affects the amino acids distribution in membrane proteins. Statistic study reveals that most of the transmembrane α-helix comprises of apolar residues such as Leu, Ile, Ala, Val, and Phe, and Gly (Senes et al. 2000). Aromatic residues such as Trp and Try are found located preferentially at the lipid headgroup-fatty acid chain interface (von Heijne 1996). The regions flanking both ends of transmembrane domain are dominated by polar residues with the cytoplasmic side being enriched in Arg and Lys. The presence of a positively charged cytoplasmic tail could help in the insertion into membranes during membrane protein biogenesis. This results in the so-called the positive-inside rule, which is applied in membrane protein topology prediction (von Heijne 1996).

In certain integral membrane proteins, the transmembrane α-helix is amphipathic. This type of α-helix has a distinctive side of more polar region and an apolar region in the opposite

side. The specific arrangement of amino acids in the amphipathic α-helix can be revealed by drawing helical wheels. The existence of amphipathic α-helix offers a structural versatility to the transmembrane domain. The more hydrophilic part of amphipathic α-helix is involved in interhelical interactions or lining a pore of a channel to provide a polar environment for the transport of ions or molecules. Transmembrane helix-helix interactions have received much attention recently.

1.2) Self-Association of Membrane Proteins

A key feature of membrane proteins is that they have a high tendency to associate with defined stoichiometries and orientations, forming oligomers. Oligomerization and helix-helix interactions of the transmembrane domain affect proper folding and function of many membrane proteins. Membrane proteins can oligomerize to form pores or ion channels, as seen in the potassium channel (a tetrameric ion channel) (Doyle et al. 1998), influenza proton channel M2 (tetrameric) (Kovacs et al. 2000), and phospholamban (pentameric) (Arkin et al. 1994). Some are able to fuse membranes or disrupt membrane permeability, e.g., influenza haemagglutinin (a trimeric viral fusion protein) (Rosenthal et al. 1998), and in some cases oligomerization of membrane proteins controls signal transduction, as seen in integrins (dimeric glycoprotein) (Li et al. 2004).

Detection and determination of oligomeric size is important in characterizing membrane proteins. It provides information of a possible function performed by the protein. Several techniques have been developed to study the oligomeric state of membrane proteins. These include cross-linking assays followed by gel electrophoresis, fluorescent resonance energy transfer (FRET) (Adair and Engelman 1994; Fisher et al. 1999), small-angle X-ray scattering (SAXS) (Bu and Engelman 1999), transmission electron microscopy (TEM) (Clarke et al.

2006), analytical ultracentrifugation (AUC) (Fleming 2008), and TOXCAT assays (Russ and Engelman 1999). Each of these methods has its advantages and disadvantages, and accurate assessment of oligomeric size is still a challenging task. Cross-linking, FRET, TEM, and TOXCAT offer detection of oligomers in lipid bilayers, an ideal condition for membrane proteins. In contrast, gel electrophoresis, AUC, and SAXS are performed in the presence of detergent micelles. In this study, the focus is on gel electrophoresis as an initial investigation on oligomer detection and stability, and AUC to obtain thermodynamic information.

1.2.1) SDS-PAGE and PFO-PAGE. Gel electrophoresis is a quick and simple method to detect oligomer formation. SDS is a common detergent in PAGE analysis; it denatures the proteins and helps in the migration during electrophoresis. SDS-PAGE is routinely employed in the lab for estimating molecular weights. However, it is not uncommon for membrane proteins to migrate anomalously in the present of SDS (Rath et al. 2009). Moreover, SDS can induce non-specific aggregation. Despite these drawbacks, some membrane proteins still maintain their native oligomeric structure in SDS, e.g. glycophorin A and phospholamban (Lemmon et al. 1992; Simmerman et al. 1996). However, there are very few cases of SDS-resistant oligomers. Often, SDS destabilizes homo-oligomer interaction.

In contrast to SDS, PFO is a milder detergent that may protect weak interactions and maintain the native oligomeric species. PFO-PAGE has been developed to be a complementary method to study oligomeric structure of membrane proteins (Ramjeesingh et al. 1999). The number of oligomers detected by PFO-PAGE is increasing, and it is becoming a powerful tool to study oligomer formation (Cleverley et al. 2008). The application of gel electrophoresis can be extended to map inter-helix interactions. If the oligomer is stable in the presence of either one of these detergents, PAGE analysis can become a useful method for

screening of mutants in which each residue in the TM is mutated and the effect on oligomerization can be accessed by gel electrophoresis (Lemmon et al. 1992; Simmerman et al. 1996).

1.2.2) Analytical Ultracentrifugation Sedimentation Equilibrium. AUC is a classical method for characterizing water-soluble proteins, and this method has been adapted to study membrane protein solubilized in detergent micelles. In contrast to all the above mentioned tools to study oligomer structure, analytical ultracentrifugation sedimentation equilibrium experiments can provide unambiguous information on the assembly state of membrane proteins. Not only it can be used to determine the oligomeric size, but also thermodynamic data. If a system is reversibly associating in solution, the equilibrium constant can be obtained and distribution of oligomeric species can be predicted over a wide range of concentrations.

However, micelles have mass and they contribute to the observed molecular weight of membrane protein solubilized in micelles. This introduces complexity into the analysis of sedimentation equilibrium data and proper care must be taken to separate the mass contribution of the protein from those of the bound detergents. Reynolds & Tanford have developed the density-matching strategy to resolve this problem. In this approach, the effective contribution of the bound detergents is eliminated by adjusting the solvent density to match that of the detergent (Reynolds and Tanford 1976). AUC is used to determine the buoyant molecular weight (M_b) of a protein in solution and it is defined as:

$$M_b = M_p(1 - \nu'\rho)$$

Equation 1-1

where M_p is the molecular weight of the protein, v' is the effective partial specific volume (ml/g) and ρ is the solvent density. Since the membrane protein is solubilized in detergent micelles, Equation 1-1 is rewritten to take into account the micelles contribution:

$$M_b = M_p[(1 - \nu_{Pr}\rho) + \delta_{Det}(1 - \nu_{Det}\rho)]$$

Equation 1-2

Where ν_{Pr} and ν_{Det} are the partial specific volumes of the protein and detergent respectively, and δ_{Det} is the amount of detergent bound to the protein. The density matching strategy manipulates the solvent density so that$\rho = 1/\nu_{Det}$, leading to $(1 - \nu_{Det}\rho) = 0$, and Equation 1-2 is just a analysis of the protein mass alone. There is no contribution from the detergent, and it does not matter how many detergent molecules are bound. The density matching strategy can be achieved by adding D_2O to the buffer. Since the density of D_2O is 1.1 g/ml, only detergents with partial specific volumes between 0.9 and 1 ml/g can be matched effectively in a H_2O-/D_2O mixture. This limits the usage of detergents to be investigated with AUC. Octyl-β-D-glucopyranoside (β-OG) and dodecylmaltoside (DDM) which are widely used for membrane protein solubilization and crystallization have a partial specific volume in the range 0.81 to 0.86 ml/g. These detergents are excluded in AUC experiments. Nevertheless, there are several detergents that can be density-matched with D_2O such as 3-(N,N-dimethylmyristyl-ammonio)propanesulfonate (C14SB), dodecyl phospholipid (DPC), and octaethylene glycol monododecyl ether (C12E8). In addition, the pentaoxyethylene octyl ether (C8E5) is neutrally buoyant (ν =0.993 ml/g), i.e., density matching can be easily achieved by adjusting the buffer density without the need of adding D_2O.

To accurately define the model, it is important to gather enough data over a range of initial concentrations and speeds to allow for global analysis. This is useful in reducing both model and parameter ambiguity. Usually, nine data collected from three different initial concentrations run at three different speeds are used for global data fitting by treating the association constant as a global parameter. A system that is not in equilibrium or irreversibly associating, will result in a poor fit, as shown by non-randomness of the residuals and poor fit statistics.

1.3) Structural Study of Membrane Protein

Determining structure of a protein is crucial for understanding their function at a molecular level. Also, design of potential drugs is dependent on structural information. Although membrane proteins constitute approximately one-third of all proteins expressed in the cell, only less than 1% of atomic-resolution structures deposited in the Protein Data Bank (PDB) are membrane proteins. To date, there are about 551 coordinate files of membrane proteins in the PDB with 202 unique membrane protein structures (Dr Stephen White's lab, *http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html*). The scarceness of membrane protein structures reflects the difficulties in obtaining a sample, and in applying classical techniques such as x-ray crystallography and nuclear magnetic resonance (NMR) to study membrane proteins in a membrane-mimicking environment.

1.3.1) x-ray crystallography is a very powerful tool for determining membrane protein structures. Most of the solved membrane protein structures are obtained with this procedure. However, obtaining a well-ordered, three-dimensional crystal, a prerequisite for

crystallography, is still a challenging task for membrane proteins. Membrane proteins are amphipathic molecules consisting of both hydrophilic and hydrophobic regions. Therefore, unlike water soluble proteins, membrane proteins need a membrane-mimicking environment to keep them fully solubilized and folded.

In principle, there are three possibilities for arranging membrane proteins in the form of 3D crystals (see Figure 1-2). In Type I, a 2D crystal is formed in the plane of a reconstituted membrane, and then stacking of these crystals in an ordered way to form the 3D lattice. This can be achieved by using the lipid cubic phase method to crystallize membrane proteins. This strategy is attractive because the membrane protein is residing in the native environment.

The type II is the most common and favorable way to obtain membrane protein crystals: to crystallize membrane proteins within detergent micelles. The hydrophobic part of the membrane protein is protected within the micelle and the more hydrophilic part of the membrane protein forms polar interactions that facilitate type II crystal growth. The size and chemistry of the detergent play an essential role in crystal packing. Additional steps, such as detergent screenings, need to be performed to find a most suitable combination for obtaining high quality crystals.

Type III crystals are a new crystal packing system that has been reported for membrane protein. It was observed in the major light-harvesting pigment-protein complex (LHC-II) from spinach chroplast crystallized in the presence of a nonbilayer forming lipid, digalactosyl diacylglycerol (DGDG) (Liu et al. 2004).

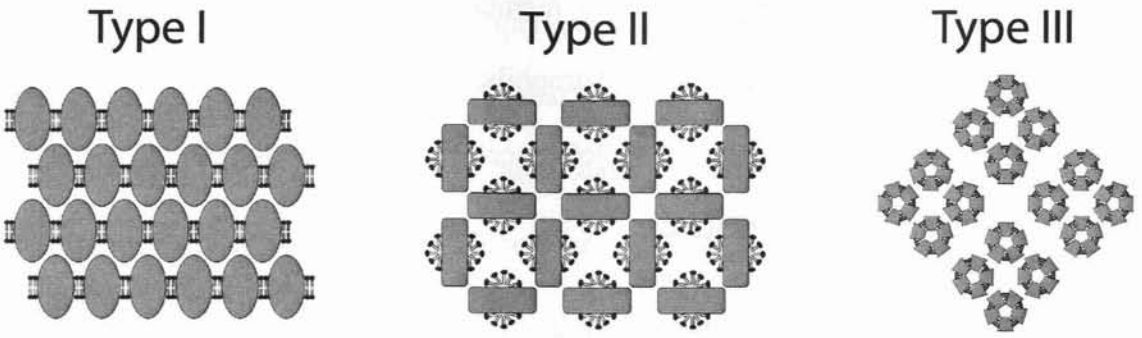


Figure 1-2. Schematic representation of the three main crystal types for membrane proteins. Type I crystals are formed by the stacking of 2D crystalline membranes ordered in the third dimension (membrane crystals), the lipids surround the proteins in a lipid bilayer-like fashion. Type II crystals are formed in the presence of detergents, the hydrophilic surface of the membrane proteins form polar interactions. Type III crystals resemble the crystals formed when icosahedral viruses crystallize. The membrane protein is ordered within a microliposome, and then these vesicles crystallize to form the macrocrystal lattice. The picture is adapted from Gardiner et al 2006 (Gardiner et al. 2006).

1.3.2) NMR spectroscopy is another well-established method to solve protein structures. To date, 33 out of 202 atomic-resolution membrane protein structures are solved by NMR (Dr Dror Warschawski's lab, *http://www.drorlist.com/nmr/MPNMR.html*). In solution NMR, membrane proteins can be studied in the presence of detergents or organic solvent. However, slow tumbling of large protein-detergent complexes lead to band broadening and loss of intensity due to fast relaxation. Innovative spectroscopy strategies such as the transverse relaxation optimized spectroscopy (TROSY) and high magnetic fields have accelerated the progress in solving membrane protein structures by solution NMR. In addition, segmented or selectively labeling technology facilitates structural study of membrane protein by solution NMR.

Solid-state NMR is another attractive tool to solve membrane protein structures, especially when the protein is forming oligomers. Solid state NMR characterizes proteins in lipid bilayers. In solid-state NMR, limitation of sample size and relaxation is overcome by using

aligned samples or magic-angle spinning samples. The number of membrane protein structures solved by this approach is increasing. In fact, for M2 protein, only one structure has been reported using solution NMR, the rest have been obtained by solid state NMR (Wang et al. 2001; Hu et al. 2007a; Schnell and Chou 2008).

1.3.3) Fourier transform infrared (FTIR) spectroscopy is another useful tool for the structural analysis of membrane proteins, in particular the application of attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy. As mentioned above, it is still difficult to determine the structure of membrane proteins by x-ray crystallography, or NMR spectroscopy. Although FTIR is a low-resolution technique, it has a particular advantage for membrane proteins over NMR or x-ray crystallography. ATR-FTIR can be used to study structural and conformational changes of membrane proteins that reside in their natural membrane environment. This technique has been applied to study lipid-protein interactions, protein-protein interactions, protein orientation, and protein dynamics in membranes.

Infrared is a form of light with long wavelength that is able to excite molecular vibrations to higher energy level and FTIR spectroscopy probes the frequencies of molecular vibrations. It is sensitive to vibrations that modulate a molecule electric dipole moment. Vibrational stretching and bending modes of the amide and carbonyl groups of peptide backbone show apparent absorption peaks in IR spectra, specifically the C=O stretching (amide I, 1600-1700 cm^{-1}), the N-H bending (amide II, 1510-1580 cm^{-1}), and the N-H stretching (amide A, ~3300 cm^{-1}). These amide vibration groups are conformation-sensitive; the frequency of IR spectrum is altered by their interactions. Amide I is widely used to determine protein secondary structure as it shows a more intense absorption band. Different secondary

structures give rise to different frequencies within the narrow range of the amide I absorption region (Table 1-1). By performing Fourier self deconvolution on the amide I region, individual secondary structures can be sorted out and the populations can be quantified. In addition, the amino acid side-chains give absorption peaks in IR spectrum. Most of them absorb in or near the amide I and amide II regions. Protein conformational changes involving side-chain interactions can be monitored by FTIR spectroscopy. For example, the sulphydryl group of cysteine has been used to map the interface of helix-helix interactions (Arkin et al. 1996).

	Band position in H_2O/cm^{-1}		**Band position in D_2O/cm^{-1}**	
Secondary structure	**Average**	**Extremes**	**Average**	**Extremes**
α-helix	1654	1648-1657	1652	1642-1660
β-sheet	1633	1623-1641	1630	1615-1638
β-sheet	1684	1674-1695	1679	1672-1694
Turns	1672	1662-1686	1671	1653-1691
Disordered	1654	1642-1657	1645	1639-1654

Table 1-1. Frequencies of Amide I band for protein secondary structure in H_2O and D_2O environments. The Table is adapted from Barth and Zscherp 2002 (Barth and Zscherp 2002).

The experimental design of ATR-FTIR is illustrated in Figure 1-3. In the ATR-FTIR experiment, protein-lipid samples are deposited on the surface of an internal reflection plate, normally germanium is used. Infrared radiation is focused into the plate where it undergoes a series of total internal reflections, creating an exponentially decaying evanescent radiation on the surface of the plate. The supported lipid bilayers and the reconstituted proteins absorb this evanescent wave and provide a wealth of information derived from both lipids and proteins in

the ATR-FTIR spectrum. When polarized light is used in conjunction with ATR-FTIR, one can evaluate the orientation of the proteins with respect to the supported membrane.

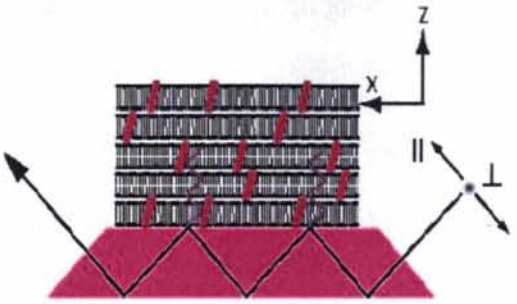


Figure 1-3. Schematic representation of ATR-FTIR experiment with transmembrane peptide samples inserted into the supported lipid bilayers.

ATR-FTIR is a powerful tool to study lipid-protein interactions. Lipids absorb in many different regions of the IR spectrum. Information of lipid orientational order parameter and physical state such as gel or liquid-crystalline phase can be determined by ATR-FTIR spectroscopy; the frequencies of the absorption bands of lipid methylene stretching vibration change upon gel-to-liquid crystal phase transition. The ester carbonyl stretching band is very sensitive to hydrogen bonding and can be used to monitor lipid hydration.

One of the most exciting applications of ATR-FTIR spectroscopy is to obtain orientation of lipid or protein in supported bilayers. Orientation information for these structural elements on a supported material can be determined by polarized ATR-FTIR spectroscopy from the dichroic ratio:

$$R = \frac{A_{\parallel}}{A_{\perp}}$$

Equation 1-3

where $A_{\parallel}$ and $A\perp$ are the integrated absorbance of a given band at parallel and perpendicular polarizations of the incident infrared light, respectively. This piece of information is used to obtain the order parameters, *S*, describing lipid, or protein order in supported bilayers:

$$S = \frac{2(E_x^2 - RE_y^2 + E_z^2)}{(3cos^2\alpha - 1)(E_x^2 - RE_y^2 - 2E_z^2)}$$

Equation 1-4

where E_x, E_y, *and* E_z are the components of the evanescent electric field at the surface of internal reflection element. For lipid bilayer deposited on the surface of germanium plate, these values are calculated as $E_x^2 = 1.969$, $E_y^2 = 2.249$, and $E_z^2 = 1.892$. α is the angle between the transition dipole moment analyzed and the long axis of the structure, which is 39° for α-helix amide I vibration and 90° for lipid methylene stretching vibration.

Another application of ATR-FTIR spectroscopy is the ability to monitor the kinetics of hydrogen/deuterium exchange. The amide hydrogen-deuterium exchange provides valuable information about the protein dynamics. As the hydrogen-deuterium exchange is focused on the amide proton only, one can measure the number of membrane embedded residues in the proteins by analyzing spectral changes in amide II or amide A bands. The amide hydrogen-deuterium exchange results in downshifts of the frequencies of amide I by ~10 cm^{-1}, amide II by ~100 cm^{-1}, and amide A by ~700 cm^{-1}. The kinetics of exchange can be calculated by measuring the area of amide II band relative to amide I band:

$$\text{Non-exchanged (\%)} = \frac{\left(\frac{A_{amide\ II}}{A_{amide\ I}}\right)_{D_2O}}{\left(\frac{A_{amide\ II}}{A_{amide\ I}}\right)_{Dry}} \times 100$$

Equation1-5

1.3.4) Site specific infrared dichroism (SSID). Innovative biophysical and computational methods have been developed in recent years to improve structure determination of membrane proteins. SSID is a new method to obtain the orientation restrains of TM α-helices in lipid bilayers (Arkin et al. 1997). Helix tilt, rotational orientation of TM α-helix can be determined from SSID by labeling two consecutive residues with $^{13}C=^{18}O$ carbonyl group, or labeling only one residue with either double deuterated CD_2 group of glycine or the triple deuterated methyl group of alanine (Torres et al. 2000b; Torres and Arkin 2002). The selected isotopically labeled residue gives rise to an isolated and unique infrared absorption band on the ATR-FTIR spectrum and the dichroic ratio of this band is related to helix orientation. SSID has been used to obtain orientational restraints of TM from various membrane proteins such as glycophorin A (Arkin et al. 1997), phospholamban (a putative Ca^+ channel) (Torres et al. 2000a), M2 protein of the influenza A virus (a H^+-selective ion channels) (Kukol et al. 1999), CD3-ζ (a component of the T-cell receptor) (Torres et al. 2002b), Vpu protein of the HIV-1 virus (a putative ion channel) (Kukol and Arkin 1999), MHC class II-associated invariant chain (Kukol et al. 2002), and E protein of the SARS coronavirus (a putative protein involved in host apoptosis and virus morphogenesis) (Torres et al. 2006). Among these TMs, NMR data is available for integrin (MacKenzie et al. 1997), M2 (Kovacs et al. 2000), Vpu (Park et al. 2003), and phospholamban (Oxenoid and Chou 2005). The TM structure obtained by this method is in close agreement to data obtained by NMR. In addition to the above, molecular dynamics simulations are important tools for understanding the physical basis of the structure and interaction of TMs. By combining SSID data with global search molecular dynamic (GSMD) simulations, we can obtain a model of the TM with oligomeric size, handedness, helix tilt, and rotational orientation.

1.3.5) GSMD simulation with incorporation of evolutionary conservation data. Alternatively, one can obtain the TM structural model independent of experimental data. This method uses evolutionary conservation data in GSMD simulations (Briggs et al. 2001). In this method, GSMD simulation is applied to a particular peptide sequence and its close homologues (>50% identity), which are likely to share the same backbone structure (Briggs et al. 2001). Theoretically, if enough variants are used, only one low- energy model that corresponds to the native structure will persist in all searches. Using this method, 7 transmembrane models have been analyzed, namely glycophorin A (Briggs et al. 2001), EmrE (a multidrug-transporter from bacteria) (Torres and Arkin 2000), CD3-ζ (Torres et al. 2002a), M2 (Torres et al. 2002c), phospholamban (mutagenesis data) (Torres et al. 2002c), integrin (homo- and hetero-oligomer search) (Lin et al. 2006a; b), and SARS-CoV E protein (Torres et al. 2005). The TM domain of glycophorin A obtained with this method differed from the published NMR structure by a Cα RMSD of < 1.0 Å.

1.4) Human Respiratory Syncytial Virus

Human respiratory syncytial virus (hRSV), enveloped viruses with linear negative-strand RNA, are the causative agent of severe lower respiratory tract disease in infants, elderly and immunocompromised populations worldwide. The virus was first identified in 1956 (Blount et al. 1956), but still there are currently no vaccine and effective antiviral drugs available. RSV-neutralizing antibodies are the only weapon available to fight this disease (Groothuis and Nishida 2002). HRSV can cause repeated reinfections throughout life and little is known about its molecular mechanism of pathogenesis. Effective treatments of RSV infection thus require the discovery of new drugs, the characterization of new drug target and inhibition mechanism.

HRSV is classified in the family *Paramyxoviridae*, subfamily *Pneumovirinae*, genus *Pneumovirus*. Members of this genus include bovine, caprine, and ovine RSVs and pneumonia virus of mice. Observation of the virus by electron microscope reveals two types of viral particles, spherical and filamentous (Bachi and Howe 1973).

The genome of hRSV encodes for 11 viral proteins, three of which are transmembrane surface glycoproteins: the fusion (F) protein, the attachment (G) protein, and the small hydrophobic (SH) protein. The inner envelope face is covered by a layer of matrix (M) protein. The virus nucleocapsid contains four proteins: the nucleoprotein (N), the phosphoprotein (P), the anti-termination factor (M2-1), and RNA-polymerase. HRSV encodes two non-structural proteins, NS1 and NS2, which are unique to members of the genus *Pneumovirus*. A schematic diagram showing the structural organization of hRSV is shown in Fig. 1-4.

HRSV enter the cells by fusion of the viral and cell membranes in the lipid-raft region (Werling et al. 1999). Attachment of virions to the cell membrane is mediated mainly by binding of G protein to cell-surface glycosaminoglycan (GAGs), a popular virus cell receptor (Krusat and Streckert 1997; Feldman et al. 1999). In addition, the F protein has been shown to bind to RhoA, suggesting an alternative way to attach to cell membrane (Pastey et al. 1999; Pastey et al. 2000). Following binding of the virus to the cell surface, membrane fusion is triggered by activation of F protein (Lamb 1993; Ruiz-Arguello et al. 2004). The fusion process is pH-independent and the mechanism is still not well understood. HRSV replicates in the cytoplasm without nuclear involvement. The virus assembles at lipid raft domains of the cell membrane and is released from the infected cell by budding (Brown et al. 2002; McCurdy and Graham 2003; Brown et al. 2004).

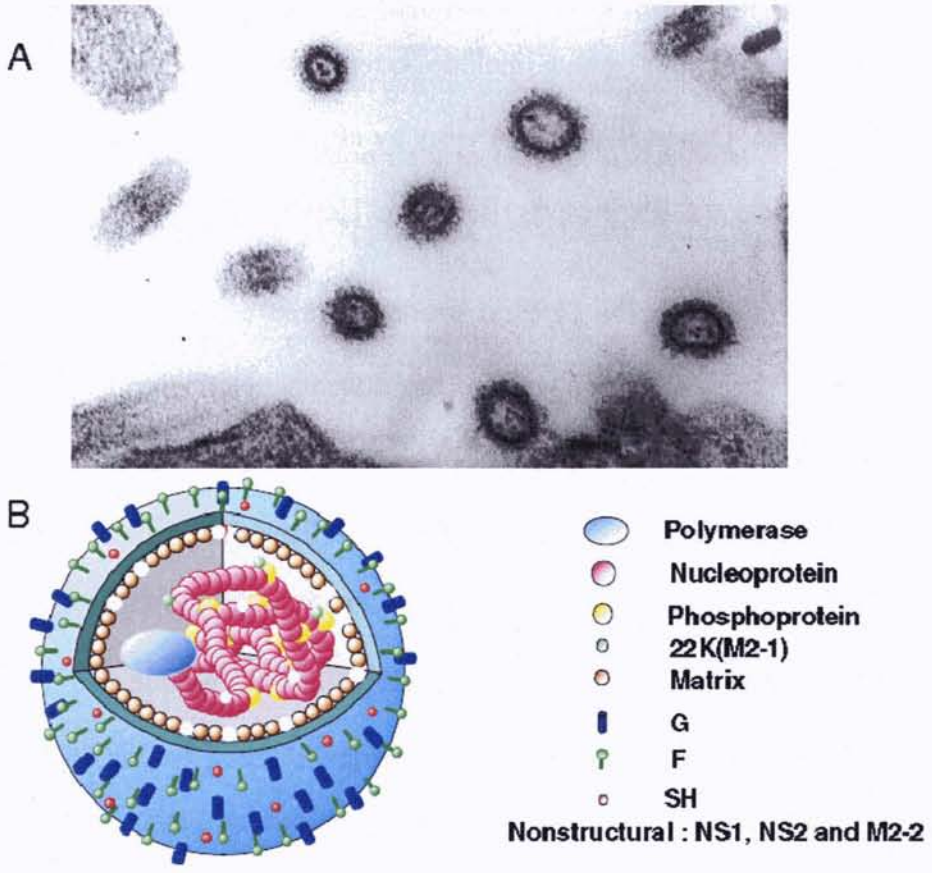


Figure 1-4. Electron microscopy (A) and schematic representation of the hRSV virion (B). The structural (colour-coded) and non-structural proteins are listed in part B of the figure. The picture is adapted from Melero 2007 (Melero 2007).

1.5) Small Hydrophobic Protein

The small hydrophobic (SH) protein is a transmembrane surface glycoprotein encoded by hRSV. During infection, the majority of the SH protein accumulates at the lipid-raft structures of the Golgi complex and also at the cell surface (Rixon et al. 2004). Only a very low amount of SH protein is associated with the viral envelope (Rixon et al. 2004). SH is categorized as an accessory protein of hRSV, as it is not essential for viral replication in vitro (Bukreyev et al. 1997). However, deletion of the SH gene leads to attenuation in mouse and chimpanzee models, suggesting it is important for infectivity (Bukreyev et al. 1997; Whitehead et al. 1999).

The SH protein is 64 to 65 amino acids long, depending on the viral strain; 64 amino acids are found in subgroup A, and 65 amino acids in subgroup B. Biochemical studies have shown that the SH protein has a single hydrophobic region that spans the membrane, with the more hydrophilic C-terminus oriented extracellularly (Collins and Mottet 1993). Analysis of SH protein amino acid sequence by TMHMM, a method of transmembrane helix prediction based on the Hidden Markov Model (Krogh et al. 2001) indicates the transmembrane α-helix of SH protein spans from residue 20 to 42 (Fig. 1-5).

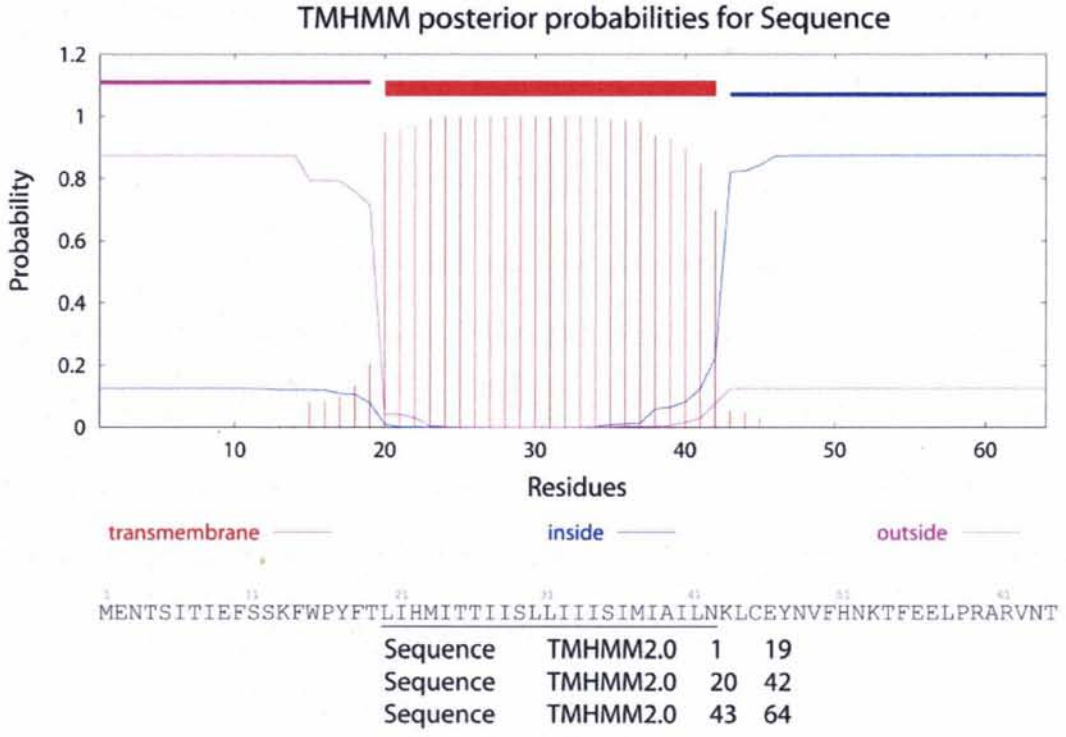


Figure 1-5. Prediction of transmembrane helices for SH protein by TMHMM program (Krogh et al. 2001) (http://www.cbs.dtu.dk/services/YMHMM/). Blue line represents residue inside the cell, magenta line indicates residues outside the cell and red line represent residues in the transmembrane helix. However, the topology prediction by TMHMM is wrong.

Several forms of SH protein are present during infection, which vary in their glycosylation status (Olmsted and Collins 1989); two non-glycosylated forms, full length 7.5 kDa (SH_0)

and a truncated 4.5 kDa species (SH_t), an N-linked glycosylated form (SH_g), and a polylactosaminoglycan-modified form (SH_p). All these, except the truncated SH_t, are incorporated at the surface of the infected cells, where non-glycosylated SH_0 appears to be the most abundant form (Collins and Mottet 1993). In addition to these modifications, the tyrosine residues of SH protein are phosphorylated during infection, and this modification affects its cellular distribution (Rixon et al. 2005).

Although the role played by the SH protein during viral infection is currently unknown, some studies suggest an ancillary role in virus-mediated cell fusion (Heminway et al. 1994; Techaarpornkul et al. 2001). More recently, cell studies have suggested that SH protein can inhibit apoptosis in several mammalian cell lines by evading the host immune response (Fuentes et al. 2007).

In addition to the above, it has been suggested that SH could form ion channels, because SH protein expressed in *Escherichia coli* can modify the permeability of the bacterial membrane to low-molecular-weight compounds (Perez et al. 1997). Also, in SDS, cross-linking of SH protein generated at least four oligomers that include dimers, trimers, tetramers and pentamers (Collins and Mottet 1993; Rixon et al. 2005). Thus, SH protein has been predicted to be a viroporin, a group of small, highly hydrophobic virus proteins that can oligomerize and form pores at the cell membrane (Gonzalez and Carrasco 2003).

Our preliminary functional studies in collaboration with Dr Soong's group at NUS suggest that SH has a channel activity. A detailed structural analysis of this protein will help to understand the assembly of SH channel, possible ion conduction mechanism, and its role in hRSV life cycle. In this study, we have determined the SH protein oligomeric state and structure using experimental and computational approaches.

2) MATERIALS AND METHODS

2.1) Sample Preparations

Several techniques were employed to produce the protein or peptide samples of SH protein (see Fig. 2-1):

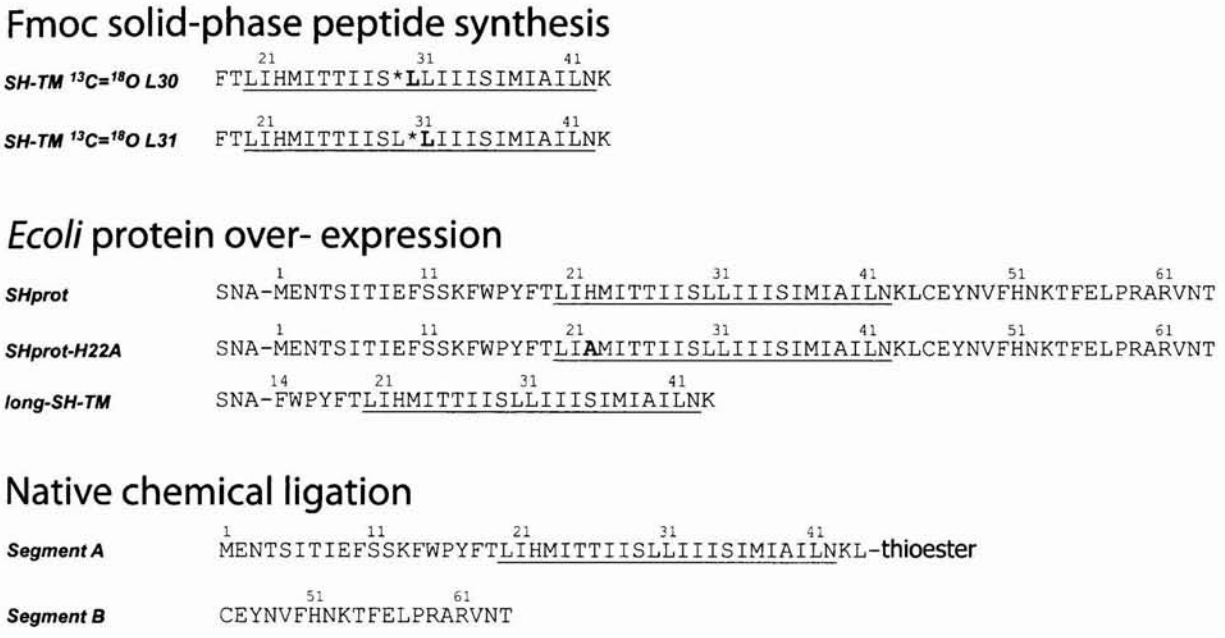


Figure 2-1. SH sequences used in this study and method to prepare them. The sequence corresponding to the transmembrane domain based on TMHMM prediction is underlined. ***L** is $^{13}C{=}^{18}O$ labelled amino acid.

2.1.1) Solid phase peptide synthesis of SH-TM and purification by RP-HPLC. Synthetic peptides corresponding to the transmembrane segment of SH, SH-TM (residues 18-43), were obtained by microwave assisted solid-phase fluorenylmethyloxycarbonyl (FMOC) chemistry using the Odyssey Microwave peptide synthesizer (CEM Corporation) available at the peptide synthesis core facility in the School of Biological Sciences. The peptides were prepared on a 0.025-mmol scale with acetylated N-terminus and amidated C-terminus,

cleaved from the resin with trifluoroacetic acid (TFA) and lyophilized. For the SSID experiment, two peptides were synthesized, each containing one $^{13}C=^{18}O$– labeled carbonyl at positions L30 or L31.

The lyophilized peptides were dissolved in a minimum amount of TFA (less than 10 µl) followed by dilution to 1 ml of acetonitrile to a final peptide concentration of ca. 5 mg/ml and immediately injected onto a Zorbax C18-300 Å column (Phenomenex, Cheshire, UK) connected to a high-performance liquid chromatography (HPLC) system (Shimadzu, Japan). The solvents used were: solvent A (water/TFA, 99.9:0.1, v/v) and solvent D (isopropanol/acetonitrile/TFA, 80:19.9:0.1, v/v). The column was previously equilibrated with a mixture of solvents A and D (7:3, v/v). The SH-TM was eluted with a linear gradient to a final solvent composition of 85% of solvent D. Pooled fractions were lyophilized and the purity of the samples was checked by MALDI mass analysis.

2.1.2) $^{16}O/^{18}O$ exchange. The two oxygen atoms in the carboxylic group of $^{13}C=^{16}O$-labeled leucine (Cambridge Isotopes Laboratories, Andover, MA) were exchanged to ^{18}O by incubating the amino acid at 100°C at acidic pH conditions (pH ~ 1) with a mixture of $H_2{}^{18}O$ and dioxane (3:1, v/v) for 1h. The extent of exchange was monitored using mass spectrometry. The solution was lyophilized and the amino acid was derivatized with FMOC using standard fluorenylmethylsuccinimidyl carbonate protocol (Wellings and Atherton 1997).

2.1.3) Synthesis of SH Protein by native chemical ligation (NCL). Two fragments were synthesized by SPPS with a 44 residues peptide representing the C-terminal thioester peptide (SH44-thioester) and a 20 residues peptide representing the N-terminal cysteine peptide (SH-C20). SH44-thioester was synthesized manually on a 4-methyl-benzhydrylamine (MBHA)

resin using standard *tert*-butoxycarbonyl (Boc) chemistry. The thioester group was incorporated in the resin before chain elongation. Double coupling with PyBop as an activator was employed to ensure efficient coupling. Deprotection was done in 30% trifluoroacetic acid (TFA). Ninhydrin test was performed to monitor efficiency at each coupling and deprotection steps. Final peptide cleavage from the resin was achieved by hydrogen fluoride (HF) cleaving at 0°C in the presence of *p*-cresol. SH-C20 was synthesized by microwave assisted solid phase FMOC chemistry using the Odyssey microwave peptide synthesizer (CEM Corporation). The SH44-thioester peptide was purified by RP-HPLC on a C4-semipreparative column with a linear gradient of solvent D (95% isopropanol/5% H_2O/0.1% TFA). SH-C20 was purified with C18-semipreparative column with a linear gradient of solvent B ((95% acetonitrile/5% H_2O/0.1% TFA).

Native chemical ligation of SH44-thioester and SH-C20 was performed in the NCL buffer containing 8M urea, 48 mM DPC, 20 mM TCEP, 90 mM MESNA, and 20 mM phosphate buffer at pH 8.0, heated at 40°C with stirring. The progress of ligation was monitored by analytical RP-HPLC over a certain time course.

2.1.4) Expression and purification of SH protein and the long-SH-TM.

2.1.4.1) Cloning of SH gene. The nucleotide sequence of SH was obtained from NCBI (NC_001803.1). The SH gene was synthesized and the purity was determined by agarose gel electrophoresis. The SH gene was cloned into pTBMalE vector with MBP as fusion partner carrying His-tag at N terminal. The pTBMalE plasmid was a gift from Ian Brett (Dr Steven Smith's lab at Stony Brook University, New York).

2.1.4.2) Over-expression of fusion proteins. The DNA plasmid containing the SH gene was transformed into an *E. coli* competent cell strain BL-21 (DE3) CodonPlus-RIL strain (Stratagene) for protein over-expression. Cells from a single colony were picked to inoculate 10 ml LB media with 100 ug/ml Ampicillin (Amp) and 34 ug/ml Chloramphenicol (Chl), and grown overnight at 37°C with shaking. 8 ml of the overnight culture was then transferred to 800 ml LB media with 1:100 dilutions and grown at 37°C with shaking to an OD_{600} of 0.6 – 0.7. For unlabelled sample, cells were induced with 0.4 mM isopropyl-β-thiogalactoside (IPTG) and grown at 23°C overnight with shaking. For labelled sample, cells were harvested by centrifugation when an OD_{600} of 0.6 – 0.7 was reached and washed with M9 minimal media once. Then the cells were transferred to M9 minimal media containing labeled chamicals, either ^{15}N ammonium chloride for nitrogen labeling or ^{13}C glucose for carbon labeling. The expression level in minimal media was enhanced by increasing cell densities through a 4:1 concentrating method described by (Marley et al. 2001). The concentrated culture in M9 media was induced with 0.4 mM IPTG and grown overnight at 23°C with shaking.

After induction, cells were harvested and resuspended in Ni^{2+}-NTA binding buffer containing 20 mM Tris-HCl, 500 mM NaCl, and 5 mM Imidazole, pH 8.0, and then kept frozen at -20°C overnight. Thawed cells were incubated with 0.2 mg/ml lysozyme and 0.02 mg/ml benzonase for 10 min. Then, β-OG was added to the sample to a final concentration of 50 mM. The cells was lysed with microfliudizer at 15 kPSI pressure and supernatant was collected after centrifugation at 20,000g for 30 min and loaded to Econo column (BioRad) packed with Ni^{2+}-NTA agarose resin (QIAGENE) pre-equilibrated with binding buffer. The fusion proteins were allowed to bind to the Nickel resin with gentle shaking at 4°C overnight.

2.1.4.3) Purification of fusion proteins by Ni^{2+}-NTA chromatography. The Nickel resin with bound fusion proteins were washed with 20 column volume of buffer containing 20 mM Tris-HCl, 500 mM NaCl, and 20 mM Imidazole, pH 8.0, to remove unbound proteins. The bound proteins were eluted with elution buffer containing 20 mM Tris-HCl, 500 mM NaCl, 500 mM Imidazole, pH 8.0, and 30 mM β-OG. All fraction collected was stored at 4°C while the fractions containing the fusion proteins were checked by SDS-PAGE.

2.1.4.4) Expression and purification of TEV protease. The DNA plasmid containing TEV gene with polyhistidine tag was a gift from Ian Brett (Dr Steven Smith's lab at Stony Brook University). Expression and purification of TEV protease was done in a similar way for SH protein except no β-OG was added to the sample as it is a soluble protein. TEV was stored at -20°C in buffer containing 50% glycerol, 10 mM Tris-HCl, 250 mM NaCl, 500 mM Imidazole, 1 mM EDTA, and 5 mM DTT, pH 8.0.

2.1.4.5) Cleaving of fusion proteins by TEV protease. Fractions containing Ni-NTA column purified fusion proteins were diluted twofold by a buffer containing 20 mM Tris-HCl, 500 mM NaCl, 1 mM EDTA, and 5 mM DTT, pH 8.0, to decrease the CMC of β-OG before mixing with TEV protease. TEV protease was added to the fusion proteins at a ratio of 1:5 (TEV: fusion protein, mg/ml). Cleaving reaction was performed at room temperature with gentle shaking and progress of reaction was monitored by SDS-PAGE. Usually, the reaction was completed in 16-24 hr. Cleaving reaction was stopped by addition of TCA at a final concentration of 6% and the precipitate was collected by centrifugation at 18,000g for 30 min. The pellet was washed with water twice followed by lyophilization.

2.1.4.6) Purification of SH proteins and SH-TM by organic solvent extraction and RP-HPLC. The cleaved proteins were purified by two step purification strategy. First, The proteins were extracted by methanol by mixing 10 ml methanol with 1L culture and mixed gently for 2 hours at room temperature. This approach assumed the target protein (a membrane protein) should dissolve in organic solvent while the MBP and TEV (soluble protein) should precipitate. The supernatant was collected by centrifugation at 18,000g for 30 min and checked by MALDI and SDS-PAGE. The proteins were further purified by injecting the supernatant onto a Zorbax C3-300 Å column connected to HPLC system. The proteins were eluted with a linear gradient of solvent A (water/TFA, 99.9:0.1, v/v) and solvent D (isopropanol/acetonitrile/TFA, 80:19.9:0.1, v/v). Pooled fractions were lyophilized and the purity of the samples was checked by MALDI.

2.2) Gel electrophoresis. The peptide sample was subjected to reduced SDS-PAGE or PFO-PAGE using precast tricine gradient gel (Bio-rad), with or without heat treatment. SDS or PFO sample buffer was added to the lyophilized peptide to a final concentration of 2 μg/ul. The sample was mixed with sample buffer for 1 min followed by heating at 95°C for 5 min before loading to the gel. The gel was run at constant voltage of 80 V for 3h at room temperature. The gel was stained with Coomassie blue

2.3) Analytical ultracentrifugation. Equilibrium sedimentation is a useful tool to determine the association state of proteins and the association constants. The experiments were performed in a Beckman XL-1 analytical Ultracentrifuge (Beckman Coulter) using six channel carbon-epoxy composite centerpieces equipped with quartz windows for absorbance experiment or sapphire windows for interference experiment. The experiment was performed

at 25°C on samples solubilized in DPC (Avanti Polar Lipids), C14SB (Sigma), or C8E5 (Sigma) at three different initial concentrations (0.3, 0.5, 0.8 absorbance at 280 nm) and three different rotor speeds. This is to provide enough data for global equilibrium data analysis. The software program Winmatch was used to monitor equilibrium at each speed every 3-4 hr. Usually, the samples achieved equilibrium in about 20 hr.

2.3.1) Determination of density-matching point for DPC and C14SB. The contribution of detergent to the buoyant molecular weight of the protein/detergent complexes were eliminated by using D_2O. The density-matching point for these detergents were carried out with sedimentation equilibrium experiments on 5 mM C14SB or 10 mM DPC in buffer containing 50 mM Tris-HCl, and 100 mM NaCl, pH 7.3 mixed with D_2O at various concentration ranged from 5% to 60%. Data was collected at 45,000 rpm and monitored with interference optic. The experiment was carried out at 25°C which is the temperature used in following experiment.

2.3.2) Sample preparation. The samples in methanol were dried in a stream of nitrogen gas and connected to lyophilizer for at least 3 hr to remove any residual organic solvent. Detergent solutions matched with D_2O was added to the dried protein samples and mixed by 5 min vortexing. In the cases where the sample was not dissolving, more vigorous mixing was performed by a cycle of sonication for 5 min followed by heating at 45°C for 5 min. Sample was centrifuged at 15,000 rpm for 5 min before loading into the centerpiece. 110 µl of sample in detergent solution and 120 µl of reference buffer (same buffer without sample) were loaded to the six channel centerpiece.

2.3.3) General data fitting strategy. Data obtained by UV absorption at 280 nm were analyzed with Sedphat by non-linear least squares curve-fitting of radial concentration profiles using the Levenberg-Marquardt algorithm. Different reversible association models were tried to fit the data to obtain the most suitable model that describe the data. AUC sedimentation equilibrium experiment measures concentration gradient along the radial position. The data can be fit to a monomer-n-mer equilibrium with an aggregation number *n*:

$$C_{total} = C_{monomer,r_0} exp\left[\frac{\omega^2}{2RT} M(1-\nu\rho)(r^2 - {r_0}^2)\right] + C_{n-mer,r_0} exp\left[\frac{\omega^2}{2RT} nM(1-\nu\rho)(r^2 - {r_0}^2)\right] + E$$

Equation 2-1

where C_{ro} is the concentration of monomer at the reference radius r_0, ω is angular velocity in radians per second, M is the monomer molecular weight, and E is baseline absorbance. For the monomer-*n*-mer equilibrium system, the association constant is:

$$K_{A,app} = \frac{C_{n-mer}}{(C_{monomer})^n}$$

Equation 2-2

Therefore, Equation 2-1 can be rewritten as :

$$C_{total} = C_{monomer,r_0} exp\left[\frac{\omega^2}{2RT} M(1-\nu\rho)(r^2 - {r_0}^2)\right] + K_{A.app}(C_{monomer,r_0})^n exp\left[\frac{\omega^2}{2RT} nM(1-\nu\rho)(r^2 - {r_0}^2)\right] + E$$

Equation 2-3

Equation 2-3 allows for determination of both the oligomeric size and the free energy of association for a monomer-*n*-mer species which is expressed as:

$$\Delta G_{A,app} = -RTlnK_{A,app}.$$

Equation 2-4

However, calculation of the equilibrium constant and free energy change for membrane protein association has to be treated in a different way. Since membrane proteins are solubilized in and surrounded by detergent micelles, association constant is expressed in mole fraction scale, where the $K_{A,app}$ is normalized for the amount of detergent:

$$K_X = K_{A,app}[micellar\ Det]_w$$

Equation 2-5

where [micellar Det]$_w$ is the concentration of detergent in micellar phase. Then, the standard free energy of association can be calculated according to Equation 2-5:

$$\Delta G_X = -RTln\left(K_{A,app}[micellar\ Det]_w\right)$$

Equation 2-6

This standard free energy is calculated assuming that the protein-micellar detergent phase has an ideal-dilute solution behavior. A series of micellar detergent concentration can be used to test for this condition. This has been tested on glycophorin A transmembrane dimer in micellar C8E5 and C14SB which prove the association is always in ideal-dilute solution condition (Fleming 2002; Fleming et al. 2004). Once the ideal-dilute solution behavior is observed, calculation of the standard free energy change is a straightforward matter.

Before analysis, data was trimmed by Sedfit program to extract the region to be analyzed, and to remove any bad points. The monomeric mass, extinction coefficient at 280 nm, and partial specific volumes for each samples ware calculated using the program Sednterp. These values were constant in fitting to various equilibrium models. The values of sample partial specific volume, and buffer conditions is listed in Table 2-1.

Sample	Partial specific volume	Detergent	Buffer	Buffer density	Buffer viscosity
SHprot	0.7590	33 mM C8E5	20 mM Na_3PO_4, 200 mM NaCl, pH7.3	1.0071	0.0090549
SHprot-H22A	0.7604	15 mM DPC	50 mM Tris-HCl, 100 mM NaCl, pH 7.3, 52.4% D_2O	1.0589	0.0089047
Long SH-TM	0.7913	5 mM C14SB	50 mM Tris-HCl, 100 mM NaCl, pH 7.3, 29.4% D_2O	1.0342	0.0089047

Table 2-1. Sample partial specific volume and buffer conditions used in the AUC sedimentation equilibrium experiment at 25°C.

2.4) Infrared spectroscopy. FTIR spectra were recorded on a Nicolet Nexus 560 spectrometer (Madison, USA) purged with N_2 and equipped with a MCT/A detector cooled with liquid nitrogen. Attenuated total reflection (ATR) spectra were measured with a 25-reflections ATR accessory from Graseby Specac (Kent, UK) and a wire grid polarizer (0.25 mM, Graseby Specac). Approximately 100 µl of sample in water with 20:1 lipid/peptide molar ratio were applied onto a trapezoidal (50 mm x 2 mm x 20 mm) Ge internal reflection element (IRE). The lipid used here was DMPC and POPC (Avanti Polar Lipids). A dry, or D_2O saturated, N_2 stream flowing through the ATR compartment was used to remove bulk water or to achieve D_2O exchange, respectively. After insertion of the plate in the ATR cell,

spectra were collected. A total of 200 interferograms collected at a resolution of 4 cm^{-1} were averaged for every sample and processed with one-point zero filling and Happ-Genzel apodisation. The area corresponding to the $^{13}C=^{18}O$ (isotope-labeled) carbonyl stretching vibration was obtained integrating the band at 1590 cm^{-1}. The area of the amide A (N-H stretching, centered at ~3300 cm^{-1}) was obtained by peak integration from 3200 cm^{-1} to 3400 cm^{-1}. No difference in band area was observed employing other means of peak size estimation such as peak fitting and Fourier self-deconvolution. The helix dichroism was measured from the amide A when the sample was hydrated in D_2O. The dichroic ratio of the band was calculated as the ratio between the integrated absorptions of the spectra collected with parallel and perpendicular polarized light. The data was analyzed according to the theory of site-specific infrared dichroism described previously.

2.5) Solution NMR study of SH protein in detergent micelles. $^{1}H/^{15}N$-HSQC spectra were acquired for 0.8 mM SHprot reconstituted into 170 mM SDS, 100 mM DHPC, or 150 mM DPC in 100 mM acetic acid, pH 4.5, and 2% D_2O. NMR experiments were performed at 37°C. (310K) using Bruker Avance-II 700 NMR spectrometers equipped with cryogenic probes (Bruker BioSpin), with three channels and triple resonance probes with shielded z-gradient coils. The carrier position was set to 4.70 ppm for ^{1}H and 119 ppm for ^{15}N. DSS was used as the internal reference for ^{1}H nuclei. The chemical shifts of ^{15}N nuclei were calculated from the ^{1}H chemical shifts. Spectra were processed using Bruker's XWinNMR and figures generated with Sparky.

2.6) Crystallization screening for SH protein and SH-TM. A sparse matrix screening kit, MemGold (Molecular Dimensions, USA), with 96 conditions for crystallization of α–helical

membrane proteins was used as an initial crystallization screening for SHprot and SH-TM. Samples were reconstituted into β-OG at ratio 1:50. Two protein concentrations were tried, 1 mM and 2 mM. Samples were prepared by mixing equal volumes of protein solution and well solution, 0.5 µl each. Crystals were grown by hanging drop vapour diffusion technique at 18°C on 24 well plate.

2.7) Global Search Molecular Dynamics (GSMD) simulation and evolutionary conserved data. Molecular dynamics is an important tool for understanding the physical basis of the structure and interaction of TMs. In GSMD, multiple symmetric helical bundles are generated, each differing from the other by the rotation of the helices about their axes and helix tilt. These are then used as starting conformations for the simulations (Adams et al. 1995). This method produces multiple low energy structures that can be clustered, because some simulations may converge to the same structure. The correct model can be selected using the evolutionary conserved data (Briggs et al. 2001). The rationale behind this is that, silent amino acid substitutions will not affect the native structure, but may destabilize non-native structures (Briggs et al. 2001). When simulations are performed using different homologous sequences, a set of similar structures appears in all of the searches, forming a "complete set" of common model (Briggs et al. 2001). The model obtained will provide information on the helix tilt, handedness and the oligomeric size independent from experimental data (Briggs et al. 2001).

2.7.1) GSMD protocol

The simulations were performed using a Compaq Alpha Cluster SC45 which contains 47 nodes. All calculations were carried out using the parallel version of the molecular modeling and manipulation program the *Crystallography and NMR System* (CNS Version 3.3), and the Parallel Crystallography and NMR System (PCNS) (Brunger et al. 1998). Using the OPLS force field (Optimized Potential for Liquid Simulations force field) (Jorgensen and Tirado-Rives 1988), the global search was carried out *in vacuo* with united atom topology, explicitly describing only polar and aromatic hydrogen atoms as described elsewhere and using CHI 1.1 (*CNS Helical Interactions*) to obtain the initial configurations. The homooligomeric interaction between the helices was assumed to be symmetrical.

Trials were carried out starting from either left or right crossing angle configurations. The initial helix tilt, β, was restrained to 5° and the helices were rotated about their long helical axes in 10° increments until the rotation angle reached 350°. Henceforth, the simulation was repeated by increasing the helix tilt in discrete steps of 10°, up to 45°. We must note that the restraint for the helix tilt is not completely strict, i.e., at the end of the simulation a drift of up to ±5° from the initial restrained value could be observed in some cases. Three trials were carried out for each starting configuration using different initial random velocities, for right and left-handed configurations, and only for pentamers. Hence, a total of 5 (tilt) × 35 (rotation) × 3 (repeats) × 2 (handedness) = 1050 structures were produced and analyzed for each sequence, for all 13 sequences, totaling 13,600 structures. Clusters were identified for each tilt with a minimum number of similar structures. Any structure belonging to a certain cluster was within 1.0 Å RMSD (*root mean square deviation*) from any other structure within the same cluster. Finally, the structures belonging to each cluster were averaged and subjected to energy minimization. These final averaged structures were taken as the representatives of the respective clusters. The tilt angle of the models, β, was taken as the

average of the angles between each helix axis in the bundle and the bundle axis. The bundle axis, coincident with the normal to the bilayer, was calculated by CHI. The helix axis was calculated as a vector with starting and end points above and below a defined residue, where the points correspond to the geometric mean of the coordinates of the five α carbons N-terminal and the five α carbons C-terminal to the defined residue. The rotational orientation angle ω of a residue is defined by the angle between a vector perpendicular to the helix axis, oriented towards the middle of the peptidic C=O bond of the residue, and a plane that contains both the helical axis and the normal to the bilayer. In this work, to compare the models, a residue was chosen arbitrarily, and the ω angle is always given for residue 30 (corresponding to the first labeled residue). Intersequence comparisons between low-energy clusters were performed by calculating the RMSD between their α-carbon backbone using the program ProFit (*http://www.bioinf.org.uk/software/profit*). The structure preserved in sequences was taken as the likely model of interaction. The energies calculated correspond to the total energy of the system, including both bonded, for example, bond, angle, dihedral, improper, and nonbonded, that is, van der Waals and electrostatic terms. The interaction energy for the residues was calculated with the function chi_interaction implemented in CHI.

2.7.2) Homologous sequences for SH.

A total of 13 sequences were used for the simulations (Fig. 2). Homologous sequences were obtained using *ncbi homoloGene search (http://www.ncbi.nlm.nih.gov/)*. The assignment of the transmembrane domain for these sequences was based on the hydrophilicity/surface probability plots (Krogh et al. 2001) and the transmembrane predictions from the TMHMM server. According to these predictors, the transmembrane region of these sequences spans 23

residues. The alignment of these sequences in the TM domain is shown in Fig. 2-2, with their NCBI accession numbers shown in Fig. 2-2, left column.

	20					25					30					35					40		
AAG28111	L	I	H	M	I	L	T	L	I	S	L	L	I	I	I	T	I	M	I	A	V	L	N
AAG28084	L	I	H	M	I	T	T	I	I	S	L	I	I	I	I	S	I	M	I	A	I	L	N
VSHHRSV1	L	I	H	M	I	L	T	P	I	S	L	L	I	I	I	T	I	M	I	A	I	L	N
NP044594	L	I	H	M	I	T	T	I	I	S	L	L	I	I	I	S	I	M	I	A	I	L	N
AAG28086	L	T	H	M	I	T	T	I	I	S	L	I	I	T	I	S	I	M	I	A	I	L	N
AAG28139	L	T	H	M	I	L	T	L	I	S	F	L	I	I	I	T	I	M	I	A	I	L	N
AAG28127	L	T	H	M	I	L	T	L	I	S	L	L	I	I	I	T	I	M	I	A	I	L	N
AAG28107	L	I	H	M	I	L	T	L	I	S	L	L	I	I	I	T	I	M	I	A	I	L	N
AAX23992	L	I	H	M	I	T	T	I	I	S	L	L	I	I	I	S	I	M	T	A	I	L	N
BAA00813	L	V	F	M	M	L	I	I	G	F	F	F	V	I	T	S	L	V	A	A	I	L	N
AAG32979	L	V	F	M	M	L	T	I	G	F	F	F	I	V	T	S	L	V	A	A	I	L	N
AAG32970	L	V	F	M	M	L	I	I	G	F	F	F	I	V	T	S	L	V	A	A	I	L	N
AAG32982	L	V	F	M	M	L	I	I	G	F	F	V	I	V	T	S	L	V	A	A	I	L	N

Figure 2-2. Sequences corresponding to SH-TM homologues used in molecular dynamics simulations. The NCBI entries are indicated on the left column. The first 9 sequences are from various subtypes of hRSV, and the last 4 sequences are from bovine RSV variants. The numbering shown at the top corresponds to the full length protein. All the conserved polar residues are highlighted in grey.

2.8) SSID data analysis. Site specific dichroism is a technique based on the fact that the measured dichroism, R of a particular transition dipole moment is a function of the sample fractional order, f and the spatial orientation of the dipole, which is defined by the parameters: β, the helix-tilt, α which relates the transition dipole moment to the helix director and ω, the rotational pitch angle (Fig. 2-3). The rotational pitch angle ω is arbitrarily defined as 0° when the C=O transition dipole moment, the helix director and the z-axis all reside in the same plane. The following residue is assumed to be 100° away as in the canonical helix. The angle a is known from fibre diffraction studies, and is 39° for transition dipole moment of the peptidic C=O bond and 29° for the N-H bond (Tsuboi 1962).

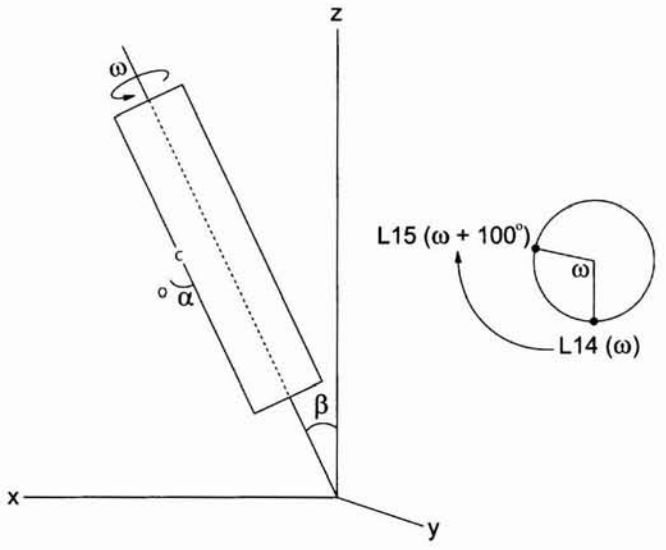


Figure 2-3. Left: schematic representation of the geometric parameters that define the amide I (C=O) transition dipole moment orientation in a transmembrane helix. The helix tilt β and the rotational orientation ω are derived from the experimental data. Right: schematic vertical view for the relative positions of labels L14 and L15 showing their distribution around the helical axis. The helical segment has been drawn as a perfect cylinder for simplicity.

From each measurement, two different dichroisms are obtained. The first is R_{Helix}, the dichroism that corresponds to the $^{12}C{=}O$ dipoles, or N-H in the case of amide A, involved in the helical structure. We note that when the dichroic ratio of the helix is obtained from the amide A dichroism, the dichroic ratio should reflect more accurately the tilt of the transmembrane domain. In fact, the amide A band in these conditions, i.e. the sample being exposed to D_2O, originates only from the transmembrane α-helix that has not exchanged. This dichroism arises from residues distributed around the helical axis, (i.e. one every 100° for a standard α-helix). Therefore, this dichroism is independent of ω, and dependent only on β and f_i:

$$R_{Helix_i}(\beta, f_i) = \frac{e_z^2(f_i K_z + \frac{1-f_i}{3}) + e_x^2(f_i K_x + \frac{1-f_i}{3})}{e_y^2(f_i K_y + \frac{1-f_i}{3})}$$

Equation 2-7

where $\kappa_{x,\ y \text{ or } z}(\langle\omega\rangle)$ are the rotationally averaged integrated absorption coefficients, f_i represents the fractional order of preparation i. The parameter f is 1 if the sample is completely ordered and zero if completely random. Finally, e_x, e_y and e_z are the electric field components for each axis given by Harrick (1967) according to a thick film approximation. The thickness of the film was calculated as being more than 30 μm, whereas the amplitude of the evanescent wave decays (at 1 / e of its initial value), 1 μm in a germanium.

The second dichroism, R_{site}, corresponds to the $^{13}C{=}^{18}O$ i label, consequently it will be dependent on the ω angle for this particular label:

$$R_{Sites_i}(\beta, f_i, \omega) = \frac{e_z^2(f_i K_z(\omega) + \frac{1-f_i}{3}) + e_x^2(f_i K_x(\omega) + \frac{1-f_i}{3})}{e_y^2(f_i K_y(\omega) + \frac{1-f_i}{3})}$$

Equation 2-8

These two equations are not sufficient to obtain β, ω and f_i (three unknowns), therefore a second label is inserted with a different ω. For example, if the label is inserted one residue above or below the first label (for a canonical a-helix there are 3.6 residues per turn) the increment in ω is 100°. Thus, two additional equations can be obtained. One is R_{Helix}, dependent on β and f_j and the other is R_{site}, dependent on β, ω + 100° and f_j . Solving these four equations for each i and j pair, will yield β_{ij}, ω_{ij}, f_i and f_j, where β_{ij} and ω_{ij} are the results obtained from the combinations of sample i and sample j.

The non-linear equations were solved with Newton's method as implemented in the Find Root function in Mathematica 3.0 (Wolfram Research, Champaign, USA). The final values of β and ω were obtained by averaging β_{ij} and ω_{ij}, respectively:

$$\beta = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij}$$

Equation 2-9

$$\omega = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_{ij}$$

Equation 2-10

Note that the maximum dichroism R_{max} for C=O groups in an α-helix was obtained from equation 2-7 ($\beta = 0, f = 1, \alpha = 39°$) and is 4.34 for germanium.

2.9) Black Lipid Membrane. Ion channel formation of SH-TM was measured by single channel recording. A bilayer membrane was formed on a 250 μm (or 150 μm) septum of a Delrin cup with a working volume of 1 ml in each cis and trans chambers. Single channel experiments were carried out in asymmetric ionic conditions. The cis and trans chambers consist of 500 mM NaCl, 50 mM HEPES, pH 7.2, and 50 mM NaCl and 50 mM of HEPES, pH 7.2 respectively. A lipid cocktail of POPE:POPS:POPC (5:3:2, v/v/v) and a lipid concentration of 50 mg/ml was used to form the artificial lipid bilayer. The lipid cocktail was painted onto the hole of the Delrin cup. Usually the bilayer could be formed quickly in the presence of an asymmetric solution. The bilayer formation was examined electrically by the current amplitude when a voltage pulse was applied.

The whole assembly was shielded from electrical and vibration interference. After the formation of the bilayer across the aperture of the septum, small aliquots of SH-TM (about 5-10 μg) reconstituted in liposome were added to the cis chamber with continuous stirring to facilitate the insertion of the peptide into the planar lipid bilayer. Stirring of the solution was

stopped when channel activity was detected. Electrical currents were recorded using Bilayer Clamp BC-525D amplifier (Warner Instruments) in capacitive feedback configuration via Ag/AgCl electrodes linked via agar salt bridges (2 % of agarose in 1 M NaCl). The trans chamber was set as reference and the cis chamber was held at different potentials ranging from -100 mV to +20 mV in 20 mV increments. Data were filtered at 50 Hz with an 8-pole Bessel filter, and analogue output signal was digitized at a sampling rate of 1 kHz by using an A/D converter (Digidata 1322A, Axon Instruments). Data processing was performed using pClamp 9.2 software (Axon Instruments). Single channel conductance was calculated from the corresponding Gaussian fits using SigmaPlot 9.0 software (Systat Software, Inc.) to current histograms by using data from segments of continuous recordings lasting longer than 10 s. Openings shorter than 0.5 ms were ignored. To avoid electrostatic interference during recording, the recording cells were placed in a Faraday cage set on a mechanically isolated table to obtain low noise recording of single channel currents. Data was recorded at the room temperature. Voltages were corrected for calculated liquid junction potential using pClamp 9.2 software (Axon Instruments).

2.10) Electrophysiological recordings on mammalian cell line and data analysis. The full-length SH gene, wild type and H22A mutant were cloned into the pIRES-AcGFP1 vector (Clonetech) by using restriction enzyme BgIII and PstI. The plasmid contains a green fluorescent protein gene to monitor expression. The cDNA was transiently transfected into HEK293 cells using the standard calcium phosphate method (Peterson et al. 1999). The vector pIRES-AcGFP1 was also transiently transfected in separate experiments as a control.

Whole-cell current was recorded at room temperature using the standard patch clamp technique, 48-72 hr after transfection. The batch contained the following (mM): 124.0 NaCl,

3.5 KCl, .01 NaH_2PO_4, 26.2 $NaHCO_3$, 1.3 $MgSO_4$, 2.5 $CaCl_2$, and 10.0 D (+)-glucose; gassed with a mixture of 95% O_2 and 5% CO_2; pH 7.4 and an osmolarity of 300 mOsmkg. The internal solution (pipette solution) contained the following (mM): 135.0 potassium gluconate, 10.0 KCl, 10.0 HEPES buffer, 0.5 EGTA, 2.0 Mg-ATP (pH adjusted to 7.3 with KOH; and an osmolarity of 275-285 mOsmkg). The voltages were uncorrected for a -9 mV junction potential, and actual voltage is obtained by subtracting 9 mV from the reported values. Whole-cell currents, obtained under voltage clamp with an Axonpatch Multiclamp 700B amplifier (Axon Instruments), were filtered at 1-5 kHz and sampled at 5-50 kHz. The access resistance Ra (usually less than 20 MΩ) and the capacitive transients were not compensated. To determined if SH is sensitive to pH change, the external solution was changed to pH 5.5, after a stable conductance in neutral pH were detected.

3) RESULTS & DISCUSSIONS.

3.1) Synthesis of SH-TM & SH protein

3.1.1) Solid phase peptide synthesis (SSPS) of SH-TM and full-length SH protein, and purification by RP-HPLC. Microwave-assisted peptide synthesis is a powerful tool to produce peptide with high yield and purity, especially for difficult peptides that are very hydrophobic and have high tendency for aggregation during chain elongation. Microwave irradiation is a form of energy that falls between infrared and radio waves in the electromagnetic spectrum. The microwave aids in peptide synthesis through dipolar polarization mechanism. The electric field from microwave radiation interacts with any molecule that has dipole moment and causes rotation of that molecule. Since the N-terminal amine group and peptide backbone are polar, they will constantly try to align with the alternating electric field of the microwave. In this way, applying microwave radiation in peptide synthesis prevents formation of chain aggregation during peptide chain elongation. We applied this synthesis strategy to produce the SH-TM. The peptide can be separated from the truncated sequences by RP-HPLC with solvent D gradient despite the fact that the yield is low (Fig. 3-1). The SH-TM was eluted with a linear gradient to a final solvent composition of 85% of solvent D. The purity of the samples was checked by MALDI mass spectroscopy, which did not show the presence of adducts. Pooled fractions from RP-HPLC was lyophilized and stored at -20°C in powder form.

Since development of SSPS, much advancement on the instrument and strategies has been made. The technology enables synthesis of peptide as long as 100 amino acids. The p7 viroporin with 63 residues was chemically synthesized by SSPS (Chew et al. 2009). Therefore it is tempting to synthesize the full-length SH protein by SSPS as it is just 64 amino acids long. However, attempts to synthesize the full-length SH protein were not

successful even with double coupling strategy. The RP-HPLC profile shows a very broad peak indicating many deletions, truncated peptides, or aggregation (Fig. 3-2). Also, the crude sample did not show any significant peak in MALDI mass spectroscopy.

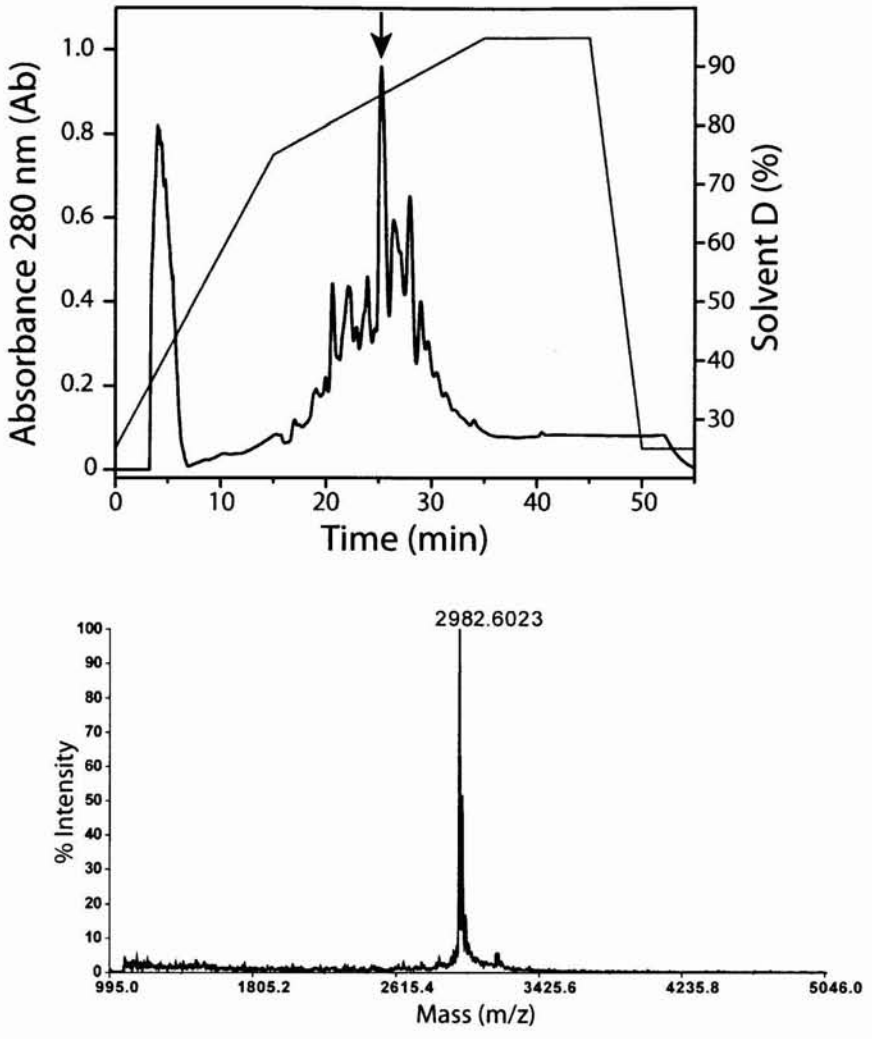


Figure 3-1. RP-HPLC purification of SH-TM. A C18 RP-HPLC chromatography column was used to purify SH-TM using linear isopropanol gradient, monitored at 220nm. The peak corresponding to purified SH-TM is indicated by an arrow and the mass spectra analysis is shown at the lower panel.

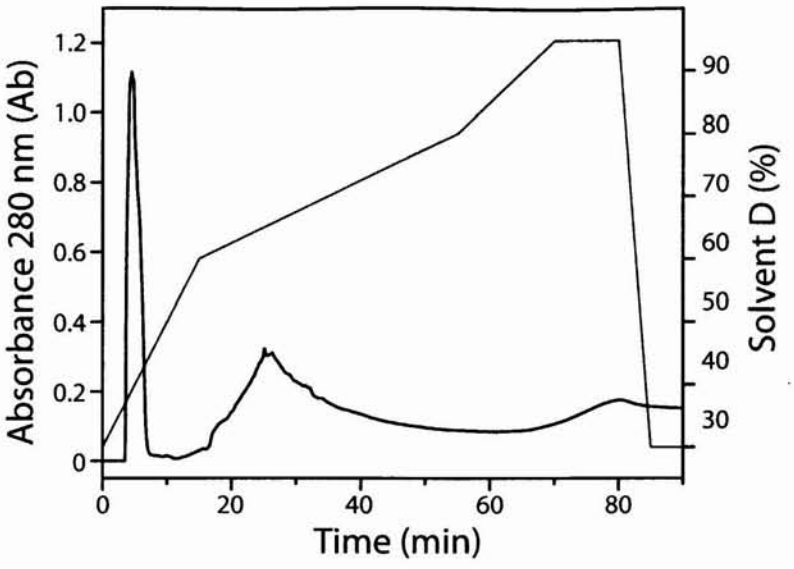


Figure 3-2. RP-HPLC purification of full-legth SH protein. A C4 RP-HPLC chromatography column was used to purify SHprot from major contaminants on the basis of hydrophobicity using linear isopropanol gradient, monitored at 220nm.

3.1.2) Synthesis of SH Protein by native chemical ligation (NCL). Attempts to produce full-length SH protein continued with the NCL strategy. This method was introduced by Kent & Tam in 1994 (Dawson et al. 1994; Liu and Tam 1994). A schematic of the NCL mechanism is presented in Figure 3-3. Briefly, two unprotected peptides are used in this reaction, one containing a C-terminal thioester group and the other containing an N-terminal cysteine. The peptides react chemoselectively under neutral conditions to form a native peptide bond at the ligation site.

This approach has been employed successfully in water soluble proteins ranging from 100 to 200 amino acids long. Several successful cases also reported for small membrane proteins such as the HIV-1 Vpu protein (Kochendoerfer et al. 2004), Influenza A virus M2 protein (Kochendoerfer et al. 1999), the mechanosensitive channel from *E. coli* and mycobacterium tuberculosis (Clayton et al. 2004), and the F_1F_0 ATP synthase subunit c (Sato et al. 2002). Unlike in water soluble proteins, detergent or organic solvent is needed during the ligation

process to solubilize peptide fragments. The progress of ligation was monitored by RP-HPLC. .

The major advantage of NCL is the possibility to generate selectively or segmentally labeled membrane proteins for biophysical studies. Unnatural amino acids, fluorescent label, EPR spin-labels, and amino acids with $^{13}C=^{18}O$– labeled carbonyls can be incorporated site-specifically. The latter strategy can be used to study transmembrane helix orientation and tilt of the full length protein with FTIR dichroism experiments. For structural studies of membrane proteins by NMR, segmentally labeled proteins can resolve the problem of severe spectral broadening that result in peak overlap commonly found in membrane protein NMR spectra.

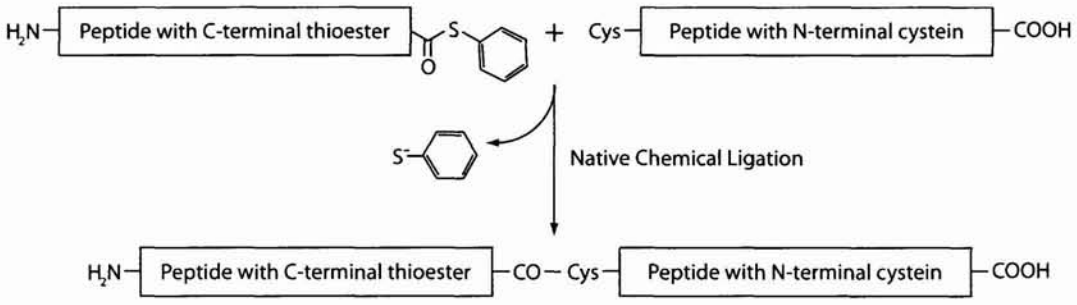


Figure 3-3. Schematic illustration of Native Chemical Ligation. Two fragments were synthesized by SPPS with a peptide representing the C-terminal thioester and a peptide representing the N-terminal cysteine peptide. The step involves nucleophilic attack of the thioester group of the by the thiol group of the N-terminal cysteine peptide. The intermediate undergoes a spontaneous rearrangement to form a natural peptide bond at the ligation site.

The SH protein is 64 amino acids long and contains a cysteine at position 45, right after the transmembrane domain (see Fig.2-1). This makes it a suitable candidate for NCL as cysteine is required for ligation. Therefore Cys45 was selected as the ligation site. The protein can be

synthesized by NCL of the 44 residues C-terminal thioester peptide (SH44-thioester) and 20 residues N-terminal cysteine peptide (SH-C20).

NCL for SH protein failed, as purification step for segment A was not possible by RP-HPLC. Although the SH-C20 can be synthesized and purified easily (Figure 3-4), SH44-thioester which contains the transmembrane domain, shows the same profile as the synthetic full-length SH protein (Figure 3-5). MALDI mass spectroscopy was used as an alternative way to monitor the ligation progress but no significant peak was detected.

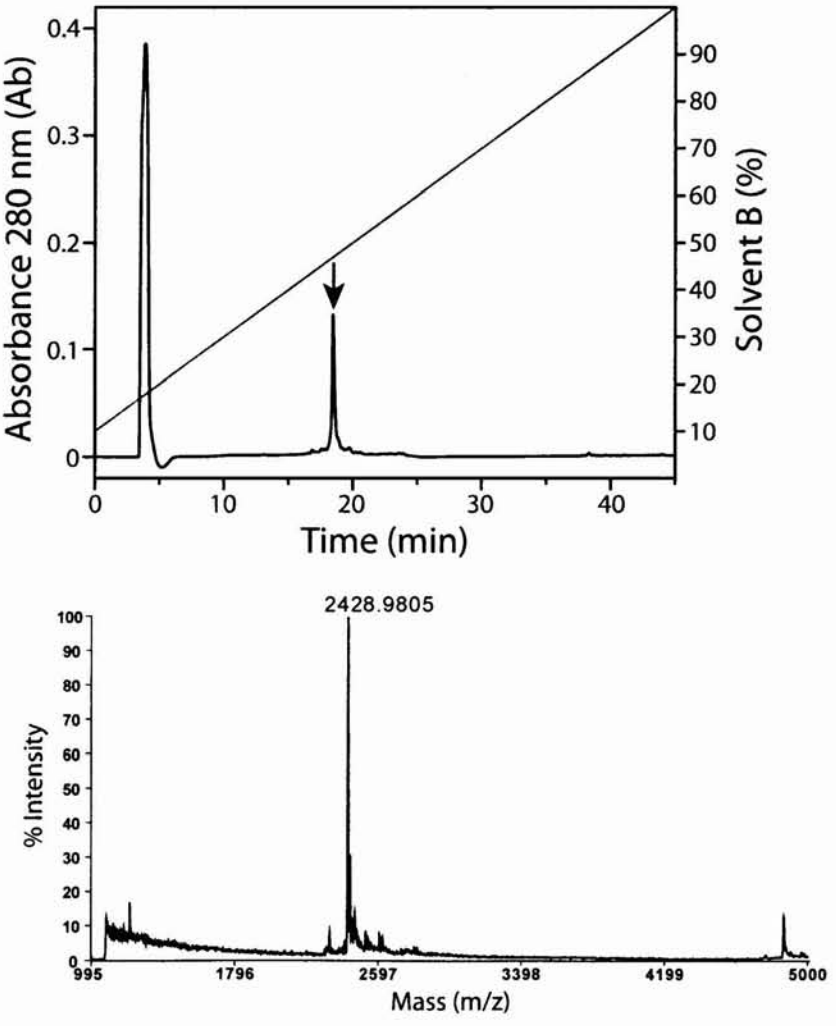


Figure 3-4. RP-HPLC purification of SH-C20. A C18 RP-HPLC chromatography column was used to purify SH-C20 using linear acetonitrile gradient, monitored at 220nm. The peak

corresponding to purified SH-C20 is indicated by an arrow and the mass spectra analysis is shown at the lower panel.

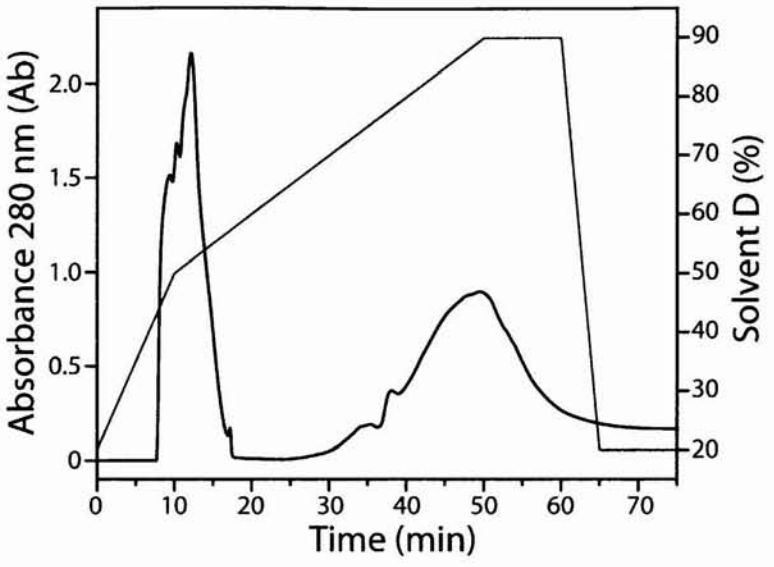


Figure 3-5. RP-HPLC purification of SH44 thioester. A C4 RP-HPLC chromatography column was used to purify SH44 thioester from major contaminants on the basis of hydrophobicity using linear isopropanol gradient, monitored at 220nm.

3.1.3) Over-expression and purification of SH protein and the long-SH-TM. Protein over-expression in *E. coli* has been a popular and effective way to produce protein in large scale. However, over-expression of membrane protein in *E. coli* is still a challenging task. The major challenge is the hydrophobicity nature of membrane protein, and it needs to insert into the membrane. Fusion of membrane proteins to big water soluble proteins such as maltose binding protein (MBP) or glutathione-S-transferase (GST) can facilitate solubility, reduce toxicity, and decrease proteolytic degradation during expression. Applying either of both fusion partners has different influences and advantages during expression. Successful cases have been reported for both systems. This protein dependent, and one has to try to check out the most suitable fusion partner for their target protein. In many cases over-expression of membrane protein in *E. coli* result in inclusion body formation although fusion

partner strategies are employed. In this way, bacteria are able to produce target protein at large scale without toxicity. The drawback of this is the refolding problem, as protein expressed into inclusion bodies is in an insoluble aggregate form, and denaturants such as GndHCl or urea must be used to extract out the protein. Refolding is required for functional studies and, most importantly, to remove a fusion partner by enzymatic cleavage that recognizes only specific sequences.

In our lab, we have tried GST, ketosteroid isomerase (KSI) and beta barrel platform (BBP) fusion proteins for expression of a small membrane protein, the E protein of SARS-CoV. The BBP proved to be the most powerful fusion partner for effective production of the E protein. Chemical cleavage by cyanogen bromide is used to release the product from the BBP. However, this is not applicable to proteins that contain methionine as cynogen bromide cleaves at methionine. SH protein has two highly conserved methionines. Moreover, cyanogen bromide cleaving reaction is a harsh condition and this could modify some amino acid side chains and complicate the downstream protein characterization work.

Recently, an effective construct of MBP fusion expression system has been developed for expression and purification of membrane proteins (Hu et al. 2007b). Briefly, this system uses MBP as the carrier protein and TEV as the cleavage enzyme (see Fig. 3-6). His-tag was incorporated at the N-terminal of MBP to aid in the purification step. This system has proved to be effective in enhancing protein expression levels, greater stability, and solubility of some small membrane protein and transmembrane domain such as M2, CorA, and KdpF (Hu et al. 2007b).

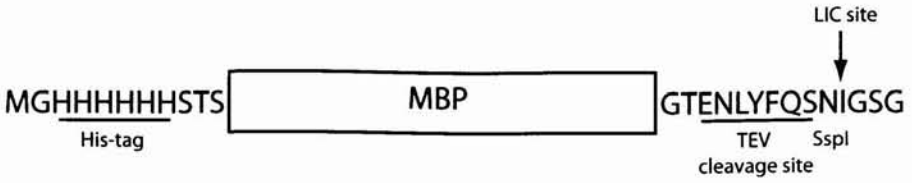


Figure 3-6. Construction of the fusion protein expression system. Both the His-tag and TEV cleavage site are underlined. TEV cleavage occurs between Gln and Ser residues and the cleave proteins have an additional three residues (SNA) at their N-terminus. The proteins are cloned into the plasmid by LIC (*shaded box*).

We have used this expression construct for SH protein and the transmembrane domain. The fusion proteins were successfully expressed in *E. coli* cells with the BL21 codon plus strain. Initial attempts to express the SH protein in the BL21 pLys strain, which has the advantage to express toxic proteins, showed very low level of expression. Later, rare codons were found in the sequence of SH protein. BL21 codon plus strain contains extra copies of rare *E. coli* codon bias and SH protein levels were significant, although just moderate (Fig. 3-7). The optimum expression condition for SH protein was found at 23°C and expressed overnight with 0.4 mM IPTG (Fig. 3-7). The long-SH-TM was expressed under the same condition. The expression levels of SH-TM was higher, about 3-fold than that of SH protein (Fig. 3-8).

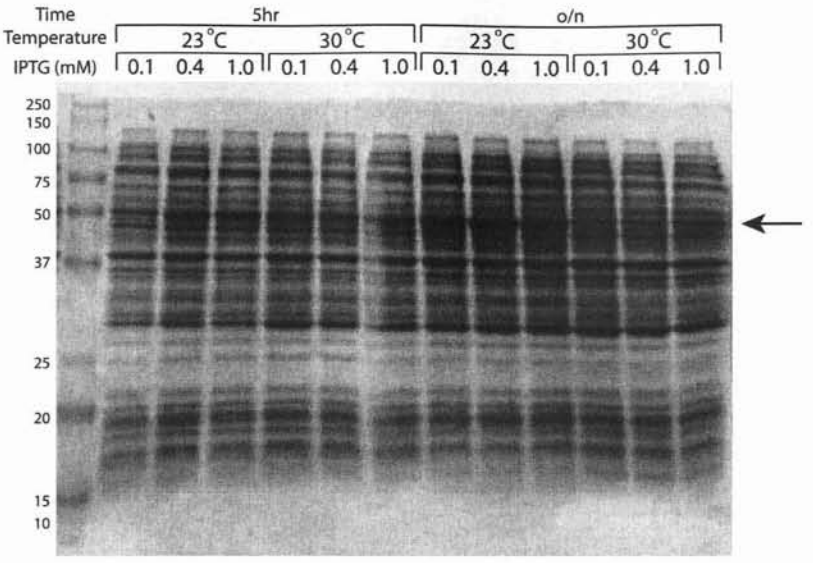


Figure 3-7. Expression of SHprot-MBP fusion protein in LB media under various conditions. The first lane is the molecular weight marker. The position of expressed SHprot-MBP fusion protein is indicated with an arrow.

Interestingly, the fusion proteins were found in the membrane fraction (Fig 3-8). For long-SH-TM, the expressed fusion protein was also found in the cytoplasmic fraction (Fig. 3-8B). Only a small amount of fusion proteins were forming inclusion bodies. Similar observations have been reported for cytochrome b6, M2, CorA, and KdpF fused with MBP (Hu et al. 2007b). The expression of SHprot and long-SH-TM in the membrane suggest that the expressed proteins fused with MBP are properly folded and potentially in a native-like conformation. It is impressive that MBP could help over-expression of membrane protein and even proper folding of membrane protein. It has been proposed that MBP may act as a chaperone to prevent aggregation of hydrophobic proteins (Kapust and Waugh 1999; Fox et al. 2001).

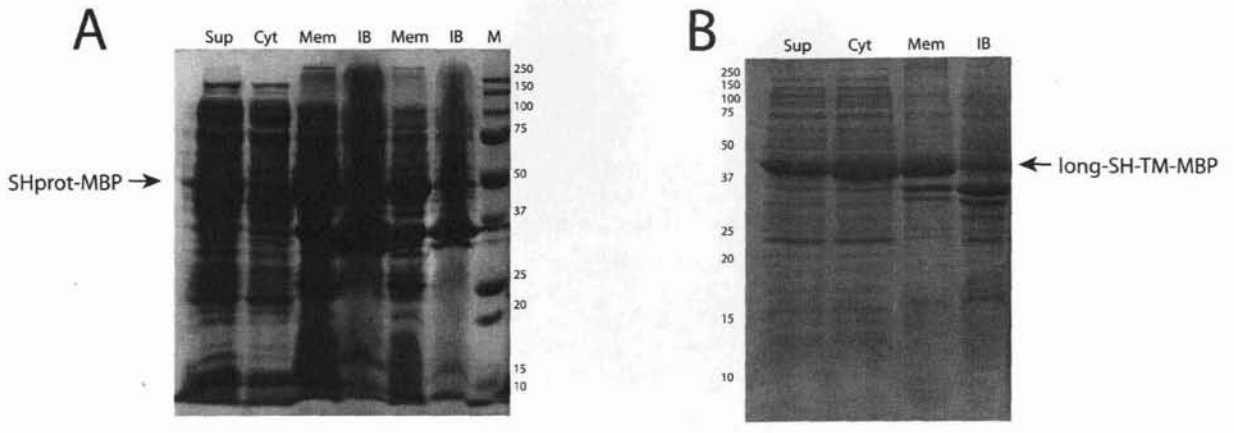


Figure 3-8. Expression and distribution of SHprot-MBP (A) and long-SH-TM-MBP (B) fusion proteins expressed at 23°C in an overnight culture. Cells were collected by centrifugation and then sonicated (Sup). After centrifugation at 8000 g for 15 min, the inclusion body fraction (IB) was collected. The supernatant was ultracentrifuged at 100,000 g for 30 min, and the new supernatant (cytoplasmic fraction, Cyt) and the pellet (membrane fraction, Mem) were collected. The position of expressed fusion protein is indicated with an arrow.

SHprot-MBP protein was successfully purified by Ni-NTA column (Fig. 3-9A). TEV cleaving was stopped in 20 hr, about 90% of SHprot-MBP was cleaved (Fig. 3-10A, *Lane 2*). Attempts to purify SHprot by methanol extraction was complicated by contaminants including TEV and MBP (Fig. 3-10A, *Lane 4*). Thus the sample was subjected to further purification by RP-HPLC. SHprot was successfully purified by RP-HPLC, the peaks were well resolved and separated from those of TEV and MBP (Fig. 3-11). MALDI mass spectroscopy shows a major peak corresponding to SHprot (Fig. 3-11). The yield of SHprot is 1.2 mg/L culture.

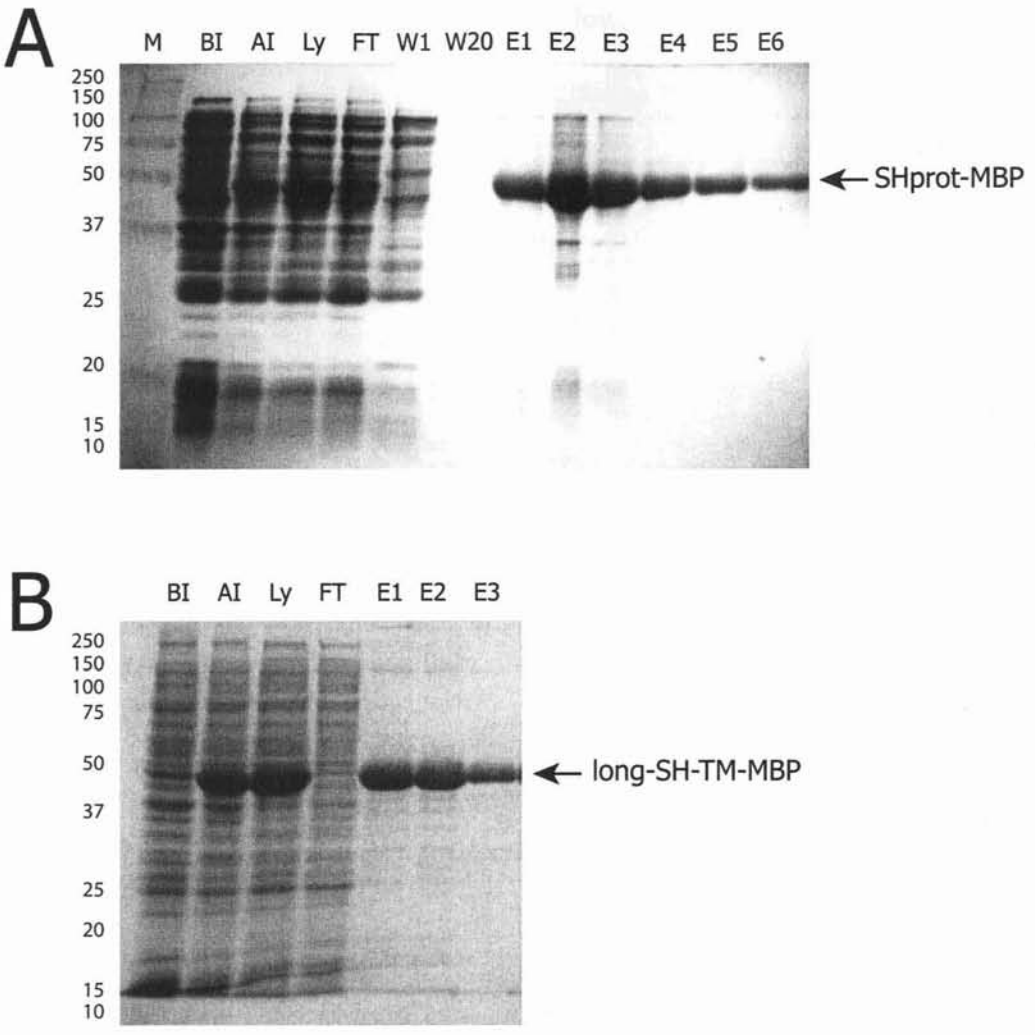


Figure 3-9. Ni-NTA purification of SHprot-MBP (A) and long-SH-TM-MBP (B) fusion proteins. Expression of fusion proteins were compared before induction (BI) and after induction with 0.4 mM IPTG (AI) and expressed at 23°C in an overnight culture. Cells were lysed with microfluidizer then centrifuged at 20,000g for 30 min, the supernantant (Ly) was collected and incubated with Ni-NTA agarose resin for overnight. The unbound proteins were removed from flowthrough (FT). The column was washed with washing buffer (W1, W20) and the bound proteins were eluted with elution buffer. The position of expressed fusion protein is indicated with an arrow.

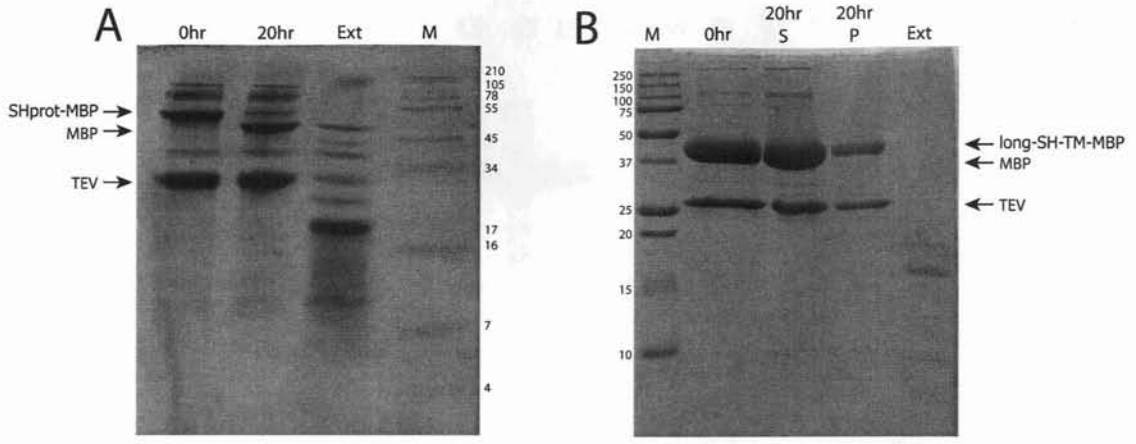


Figure 3-10. TEV cleaving and purification of SHprot (A) and long-SH-TM (B) with organic solvent extraction. TEV and fusion proteins were mixed together for cleaving reaction (0hr) with gentle shaking for 20hr at room temperature. The proteins were precipitated with TCA and lyophilized. The proteins were extracted with methanol and the supernatant was collected by centrifugation at 12,000 g (Ext).

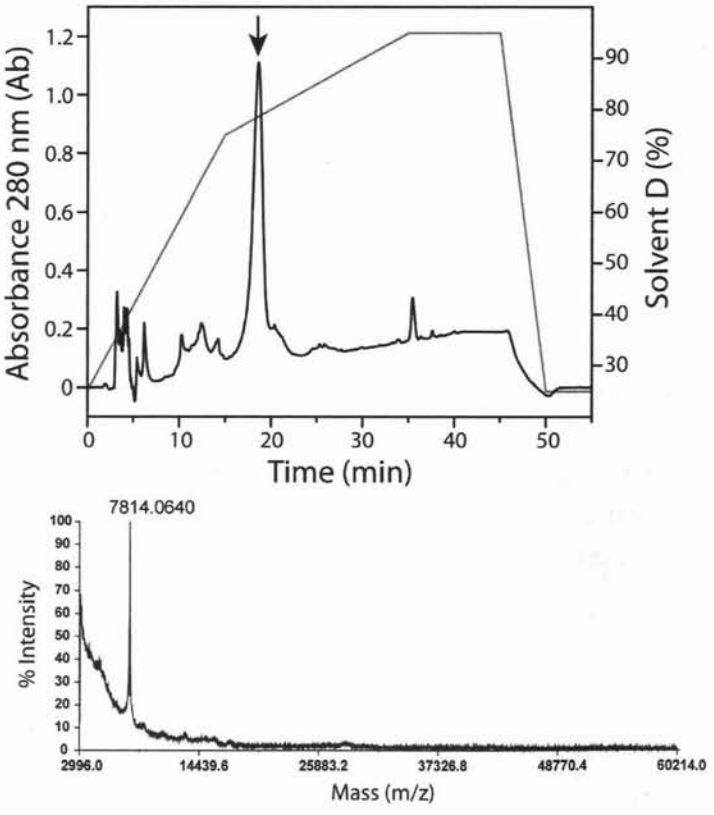


Figure 3-11. RP-HPLC purification of SHprot. A C4 RP-HPLC chromatography column was used to purify SHprot from major contaminants on the basis of hydrophobicity using linear isopropanol gradient, monitored at 220nm. The peak corresponding to purified SHprot is indicated by an arrow and the mass spectra analysis is shown at the lower panel.

In contrast, the yield for long SH-TM was very low, only 0.3 mg/L culture although the expression level of long SH-TM was better and higher than that of SH protein (Fig. 3-8). This is due to sample precipitation during elution step of Ni-NTA column chromatography. Increasing the concentration of β-OG did not improve the situation. Therefore, only a small amount of sample was recovered from the eluted fractions (Fig. 3-9B). This may be due to high hydrophobic nature of the transmembrane domain and that β-OG is not a suitable detergent for this sample. In future, a range of detergents should be tried for the elution step. The expressed SH-TM can be extracted by methanol and purified by RP-HPLC (Fig. 3-10B and 3-12). The purity was confirmed by MALDI mass spec (Fig. 3-12).

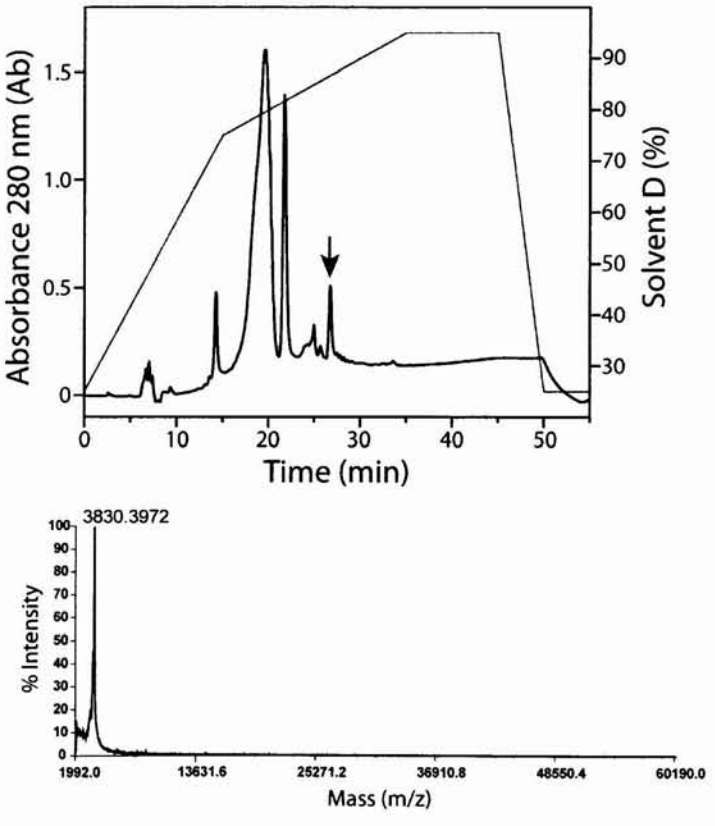


Figure 3-12. RP-HPLC purification of long SH-TM. A C4 RP-HPLC chromatography column was used to purify long SH-TM from major contaminants on the basis of hydrophobicity using linear isopropanol gradient, monitored at 220nm. The peak corresponding to purified long SH-TM is indicated by an arrow and the mass spectra analysis is shown at the lower panel.

3.1.4) Over-expression of SH protein in M9 minimal media to produce ^{15}N or $^{13}C/^{15}N$ uniformly labeled sample for solution NMR experiments. The growth of *E. coli* cells in minimal media that is deficient in nutrients is often low compared to that of rich LB media. A cost effective and efficient way to enhance *E. coli* growth and protein expression in minimal media was employed. This method utilizes protein expression in high cell densities (Marley et al. 2001). The cell densities were increased through a 4:1 concentration method by growing 2 L of cell culture to OD_{600} ~0.7 in LB media followed by centrifugation and resuspension into 0.5 L M9 minimal media. Induction was initiated after a short period of growth in minimal media to allow for recovery. By increasing the cell mass per unit volume, higher expression level was achieved (see Fig. 3-13). The yield of purified protein is 1.5 mg/ L culture. MALDI mass spec shows a good labeling level, with 97% ^{15}N incorporation and 93% $^{15}N/^{13}C$ incorporation (Fig. 3-14).

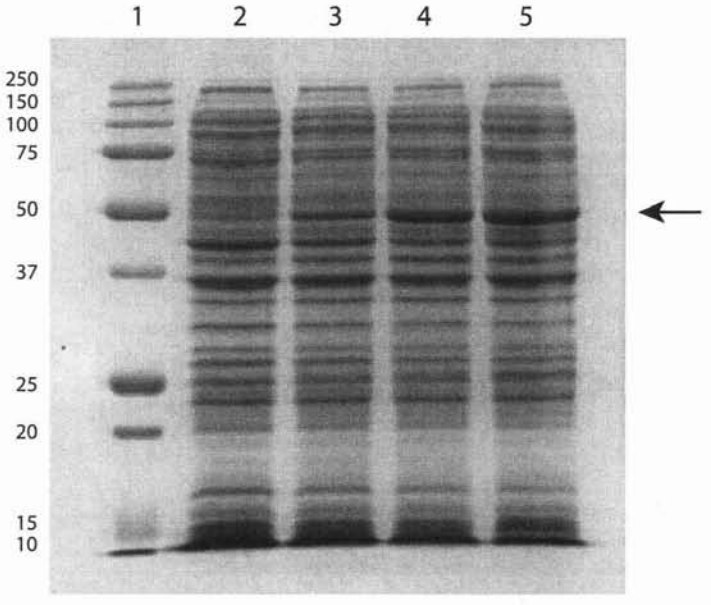


Figure 3-13. Expression of SHprot in minimal media. Lane 1 is molecular marker, lane 2 is before induction, lane 3-5 are 3 hours, 5 hours, and overnight expression after induction with 0.4 mM IPTG. The position of expressed SHprot fused with MBP is indicated with an arrow.

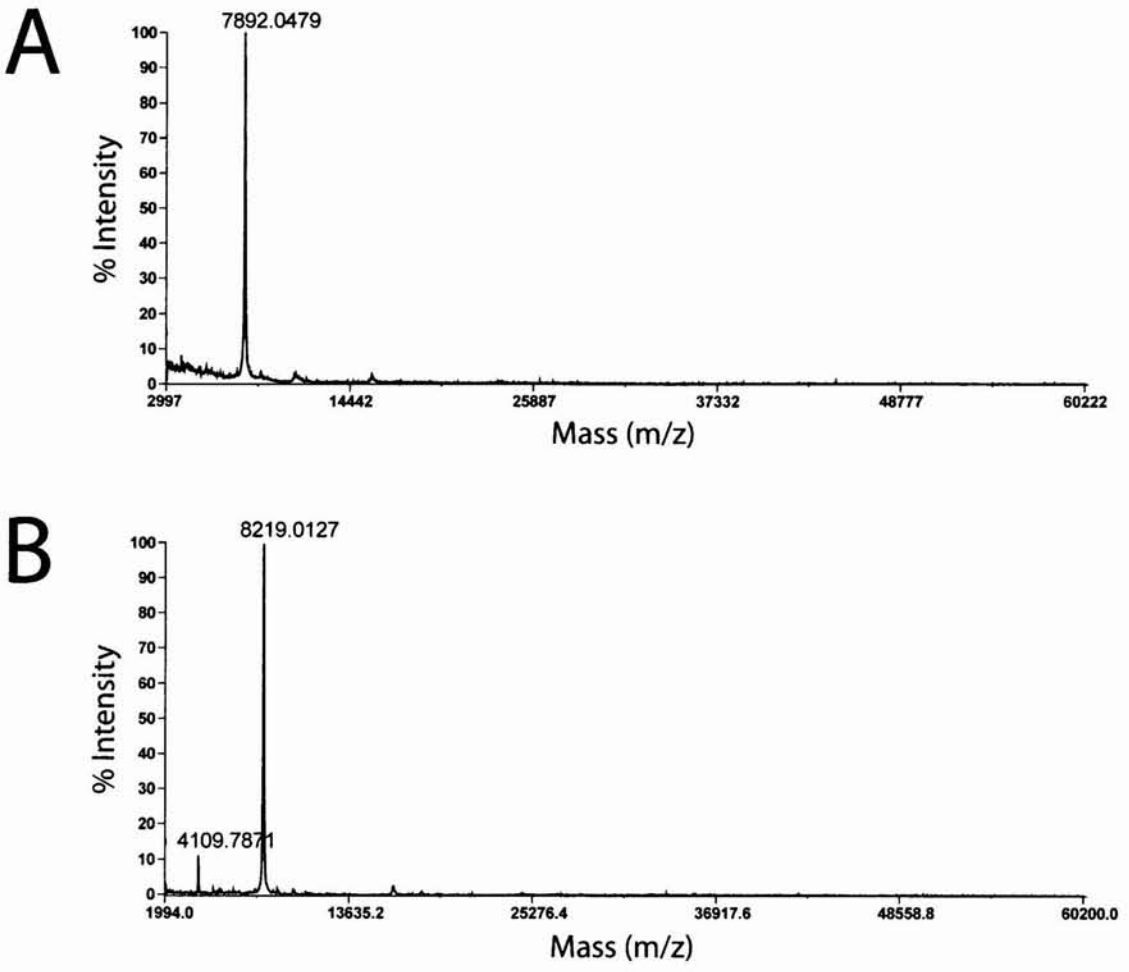


Figure 3-14. ^{15}N (A) and $^{15}N/^{13}C$ (B) labeled SHprot MALDI spectral analysis. The expected molecular weight for fully labeled ^{15}N and $^{15}N/^{13}C$ SHprot are 7895 and 8254, respectively.

3.2) Determination of SH Protein Oligomers

It has been shown that the SH protein can be cross-linked to higher oligomers, from dimers to pentamers with disuccinimidyl suberate (linker distance of 11.4 Å) and dithiobis(succinimidyl) propionate (linker distance of 12 Å) (Collins and Mottet 1993; Rixon et al. 2005). To determine the oligomeric size of SH protein, we performed SDS-PAGE, PFO-PAGE, and AUC-SE on SH-TM, long-SH-TM, SHprot, and SHprot-H22A in a detergent environment. The long-SH-TM was generated to include the tryptophan residue in the native sequence for AUC-SE experiment because there is no strong chromophore, such as tryptophan or tyrosine residue, in the sequence of SH-TM. Mutant SHprot-H22A was also generated, as His22 is the only charged residue in the transmembrane domain of SH, and is possibly related to the function of SH as an ion channel. We wish to characterize the effect of this mutation on the oligomer stability.

3.2.1) Assessment of oligomeric state of SH by SDS-PAGE & PFO-PAGE. Figure 3-15 shows the result of gel electrophoresis analysis of the oligomeric state of SHprot, SHprot-H22A, long SH-TM, and SH-TM in the presence of SDS or PFO. SH-TM forms only monomers in the SDS-PAGE whereas the SH protein, SH-H22A protein and the long SH-TM show the presence of higher molecular species (Fig. 3-15A). Both the SHprot and SHprot-H22A proteins show a similar migration pattern but the bands are smeared indicating some non-specific aggregation induced by SDS. Nevertheless, some dimers and trimers can be seen in both samples in SDS-PAGE. This is contrast with the previous studies, where SH protein expressed in mammalian cell lines only forms monomers in SDS-PAGE. Interestingly, the long SH-TM with extended 7 amino acid residues before the TM is aggregating to form tetramers and trimers in SDS-PAGE. It is possible these additional residues before the TM destabilize the oligomer in SDS micelles. In short, no stable oligomeric structure was detected by SDS-PAGE.

In contrast, higher oligomers were seen when SDS was replaced by PFO (Fig. 3-15B&C). This pentamer structure was detected only in the SH-TM (Fig. 3-15B, lane 5). The oligomers for long SH-TM were not stable even in the presence of PFO; the sample smeared during electrophoresis, indicating non-specific aggregation. In the case of full-length protein, both the SHprot and SHprot-H22A showed the presence of octamers (Fig. 3-15C). Traces of pentamer can be seen in the SHprot whereas higher aggregates were detected in the SHprot-H22A sample. Both full-length proteins showed some degree of instability in the presence of PFO as evidenced from smearing during electrophoresis. Together, it seems that the extracellular domain destabilizes the pentameric structure in the PFO micelles.

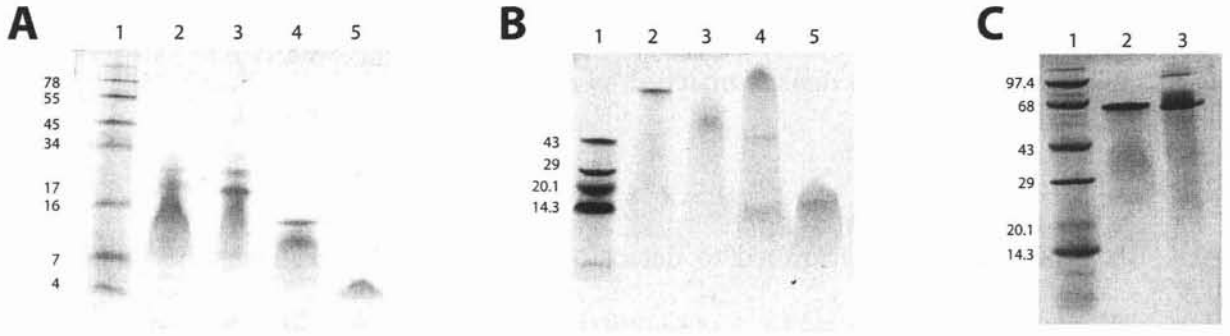


Figure 3-15. Gel electrophoresis analysis of SHprot, SHprot-H22A, long-SH-TM and SH-TM and mutants in SDS (A) and PFO (B &C). Lane 1 is protein markers; lane 2 is SHprot (expected mw 7808), lane 3 is SHprot-H22A (expected mw 7742), lane 4 is long-SH-TM (expected mw 3806), and lane 5 is SH-TM (expected mw 2983).

In mammalian cells infected with RSV, SH protein is modified by tyrosine phosphorylation that influences its cellular distribution (Rixon et al. 2005). However, pentamerization is not affected by this modification (Rixon et al. 2005). The sequence of SH contains two tyrosine residues; one is located at the N-terminus domain, and the other at the C-terminus part. It is not known which of these residues is phosphorylated. The long-SH-TM sequence contains

the N-terminus tyrosine and it is possible this unmodified tryrosine destabilizes the pentamer in the SDS and PFO environment.

The results of PFO-PAGE suggest that the transmembrane domain of SH is the main driving force for the pentamerization as seen in many membrane proteins.

3.2.2) Analytical ultracentrifugation studies of SHprot, SHprot-H22A and long-SH-TM homopentamerization equilibrium in detergent solutions.

The structure, function, and dynamics of membrane proteins are strongly dependent on the surrounding environment. Table 3-1 lists the detergents used throughout this project from purification to characterization. As only detergents with partial specific volume between 0.9 and 1 ml/g can be effectively density matched by of D_2O, only C8E5, DPC, and C14SB were used for the AUC experiment.

AUC-SE experiment was performed to detect oligomer and to investigate the interaction energetics of SHprot, SHprot-H22A, and long-SH-TM in these detergents. The sedimentation equilibrium conditions, speed, temperature, and detergent concentrations were identical for each of these samples.

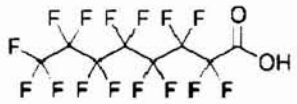
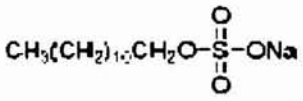
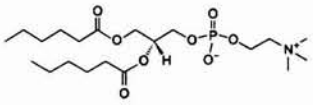
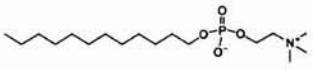
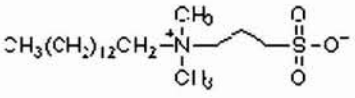
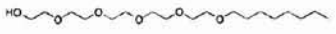
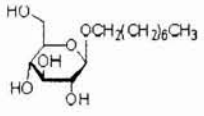
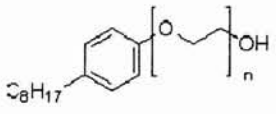
Detergents	Charge	Monomer mass (M_r)	CMC (mM)	Aggregation number	V_D (cm^3/g)
PFO 	Anionic	414	9.7	-	0.588
SDS 	Anionic	288	1.2-7.1	62-101	0.863
DHPC 	Zwitterionic	454	15	19-35	0.865
DPC 	Zwitterionic	352	1.1	50-60	0.937
C14SB 	Zwitterionic	364	0.1-0.4	83	-
C8E5 	Non-ionic	350	4.3-9.2	82	0.993
β-OG 	Non-ionic	292	25	84	0.859
Triton X-100 	Non-ionic	625	0.25	75-165	0.908

Table 3-1. List of detergents used in this study.

3.2.2.1) Determination of density-matching point for DPC and C14SB. Determination of the matching point of D_2O concentration was done by measuring the slope of the curve obtained from interference scans. The slope of the line indicates relative densities of the micelle and the solution. A positive slope indicates the detergent micelles are more dense than the solution and the micelles sediment; a negative slope indicates the detergent micelles are less dense than the solution and the micelles float. A zero slope indicates the micelles are evenly distributed and a matching D_2O concentration is found.

Fig. 3-16A shows an initial trial to match DPC with D_2O at various concentrations. The slope reverses between 50% and 55%. Plotting the slopes of the micelles distribution as a function of D_2O percentage gave an initial guess of matching point around 52% of D_2O (Fig. 3-16B). Performing the experiment with D_2O concentration around this value gave a zero slope at 52.4% D_2O (Fig.3-16C). This experiment shows that density matched solutions are sensitive to slight changes of D_2O concentration. Our value is in good agreement with the reported values under the same buffer conditions, 52.4% and 52.5% (Kochendoerfer et al. 1999; Stouffer 2006).

The density matching experiment for C14SB was done in a similar way as DPC (Fig 3-17). The slope reverses between 20% and 30% (Fig. 3-17A) and the micelles distribution versus D_2O concentration plot estimates a value around 29% (Fig. 3-17B). Final density matching experiment gave a zero slope at 29.4% D_2O. The value is 0.4% higher than the reported value, 29% (Burgess et al. 2008). However, many factors can contribute to the difference in density matching value and individual lab should perform this experiment to obtain the correct D_2O matching value.

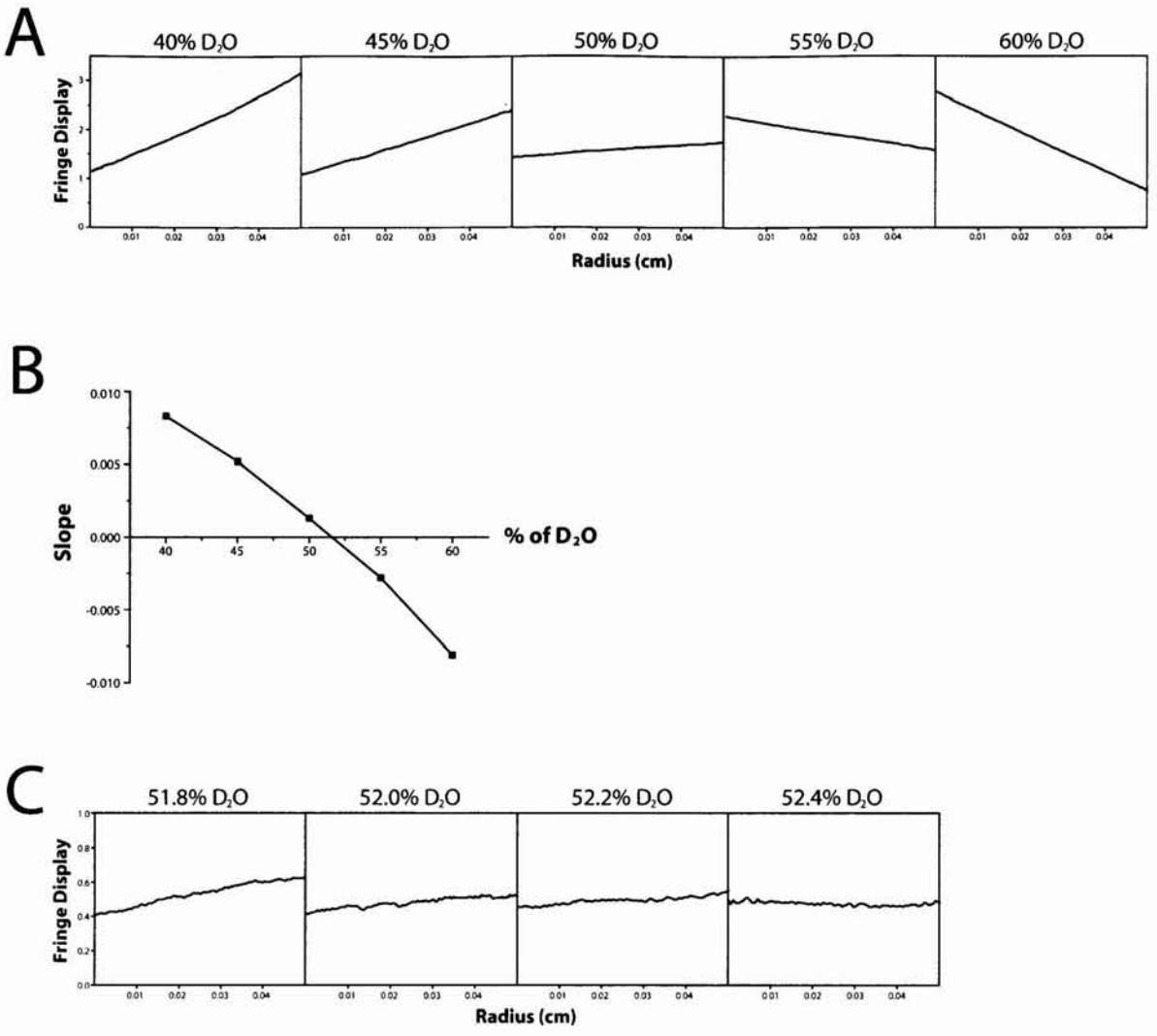


Figure 3-16. Density matching of 10 mM DPC with D_2O in buffer containing 50 mM Tris-HCl, and 100 mM NaCl, pH 7.3. (A) Interference scans of various concentration of D_2O. (B) The plot of slope of micelle distribution as a function of D_2O percentage. (C) Final density matching of 10 mM DPC based on the plot at (B).

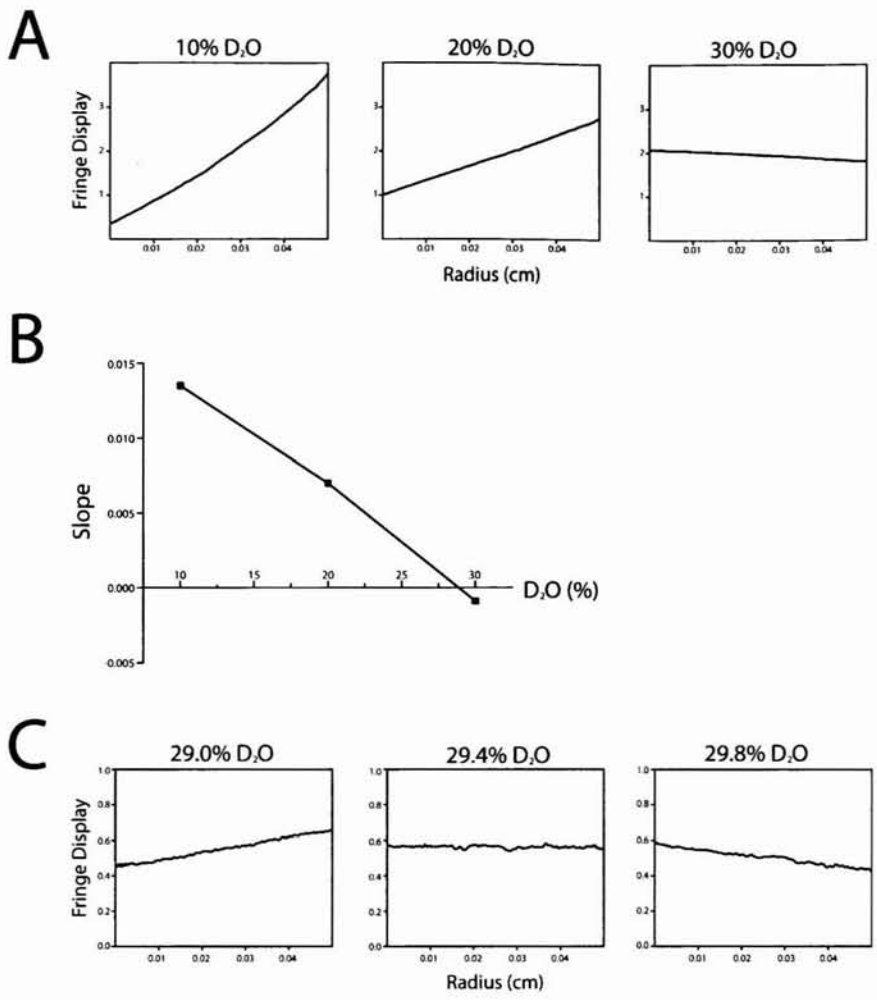


Figure 3-17. Density matching of 5 mM C14SB with D_2O in buffer containing 50 mM Tris-HCl, and 100 mM NaCl, pH 7.3. (A) Interference scans of various concentration of D_2O. (B) The plot of slope of micelle distribution as a function of D_2O percentage. (C) Final density matching for 5 mM C14SB based on the plot at (B).

3.2.3) SH protein forms reversible mono-pentamer equilibrium in C8E5, DPC, or C14SB. To explore the best model to describe the behaviour of SH protein in these detergents, different reversible association models from mono-dimer to mono-heptamer were tried to fit the data. Finally, a monomer-pentamer equilibrium model gave the best fit. The fitting results for SH protein in these detergents are summarized in Table 3-2. Two parameters were used to evaluate the fitting result, the square root of the variance (rmsd) and global chi square (χ^2). The χ^2 value should approach 1 and the the rmsd value should be

below 0.01 (the expected noise level from the optical detectors) if the mathematical model accurately describes the data (Laue 1995).

Among these detergents, global fitting of SH protein using a reversible mono-pentamer equilibrium in C14SB detergents gave the best fitting result (see Fig. 3-18). The residuals are all randomly distributed around zero, and the χ^2 and rmsd values are in the expected range, 1.16 and 0.0071, respectively, suggesting that the monomer-pentamer equation with a single equilibrium constant is a good description of the data. The calculated standard free energy of SH protein pentamerization in C14SB micelles is -15.54 kcal/mol.

The fitting for SH protein mono-pentamer equilibrium in C8E5 micelles is also a good fit (see Fig. 3-19). The χ^2 and rmsd values are 1.46 and 0.0073, respectively. The calculated standard free energy of SH protein pentamerization in C8E5 micelles is -24.89 kcal/mol.

In the case of DPC (see Fig. 3-20), the χ^2 value is 1.83, far from a reasonable range. The rmsd value is 0.0095, just slightly below the accepted value. The residuals are randomly distributed around zero, except the high and mid concentrations at 28 krpm. It is very likely that at the lowest speed, 28 krpm, less mono-pentamer equilibrium is observed in SH protein solubilized in DPC micelles. It is well known in sedimentation equilibrium experiments that lower speeds will produce data defining the larger oligomers. The standard free energy of SH protein pentamerization in DPC micelles is -18.69 kcal/mol.

The values of standard free energy of SH protein pentamerization in these detergents suggest that the process is strongly affected by detergent environment. The pentameric structure is most favorable in non-ionic detergents, the C8E5. Comparison of the pentamer equilibrium distribution for SH protein in all three detergents is shown in Figure 3-21. Under our experimental conditions, more than 95% of SH protein are forming pentamer in C8E5. In

DPC, 82% to 90% of the protein is forming pentamer, whereas 78% to 88% of SH protein is forming pentamer in C14SB.

The association of membrane proteins in different detergent environments has been measured for glycophorin A and MS1. The transmembrane helix of glycophorin A fused to the C-terminal of *Staphylococcal* nuclease (SN-GpA-TM) dimerization has been investigated in the C8E5 and C14SB and the dimerization is more favorable in C8E5 compared to C14SB (Fleming 2002; Fleming et al. 2004). The MS1 is a designed peptide for transmembrane version of the GCN4-P1 peptide. And the oligomerization is described by mono-dimer-trimer model. While C14SB promotes trimerization MS1 peptide, DPC promotes dimerization of MS1 peptide (Gratkowski et al. 2002).

Conditions	**Global Chi^2**	**Rmsd**	**$K_{A,app}$ (M^{-1})**	**$\Delta G_{A,app}$ (kcal/mol)**	**K_x**	**ΔG_x (kcal/mol)**	**ΔG_x (kcal/mol) (per monomer)**
5 mM C14SB	1.163	0.00709	4.044×10^{20}	-28.11	2.53×10^{11}	-15.54	-3.11
15 mM DPC	1.826	0.00954	1.017×10^{21}	-28.66	5.15×10^{13}	-18.69	-3.74
33 mM C8E5	1.455	0.00728	1.540×10^{24}	-33.00	1.83×10^{18}	-24.89	-4.98

Table 3-2. Summary of sedimentation equilibrium analysis of SH protein pentamer in various detergent environments.

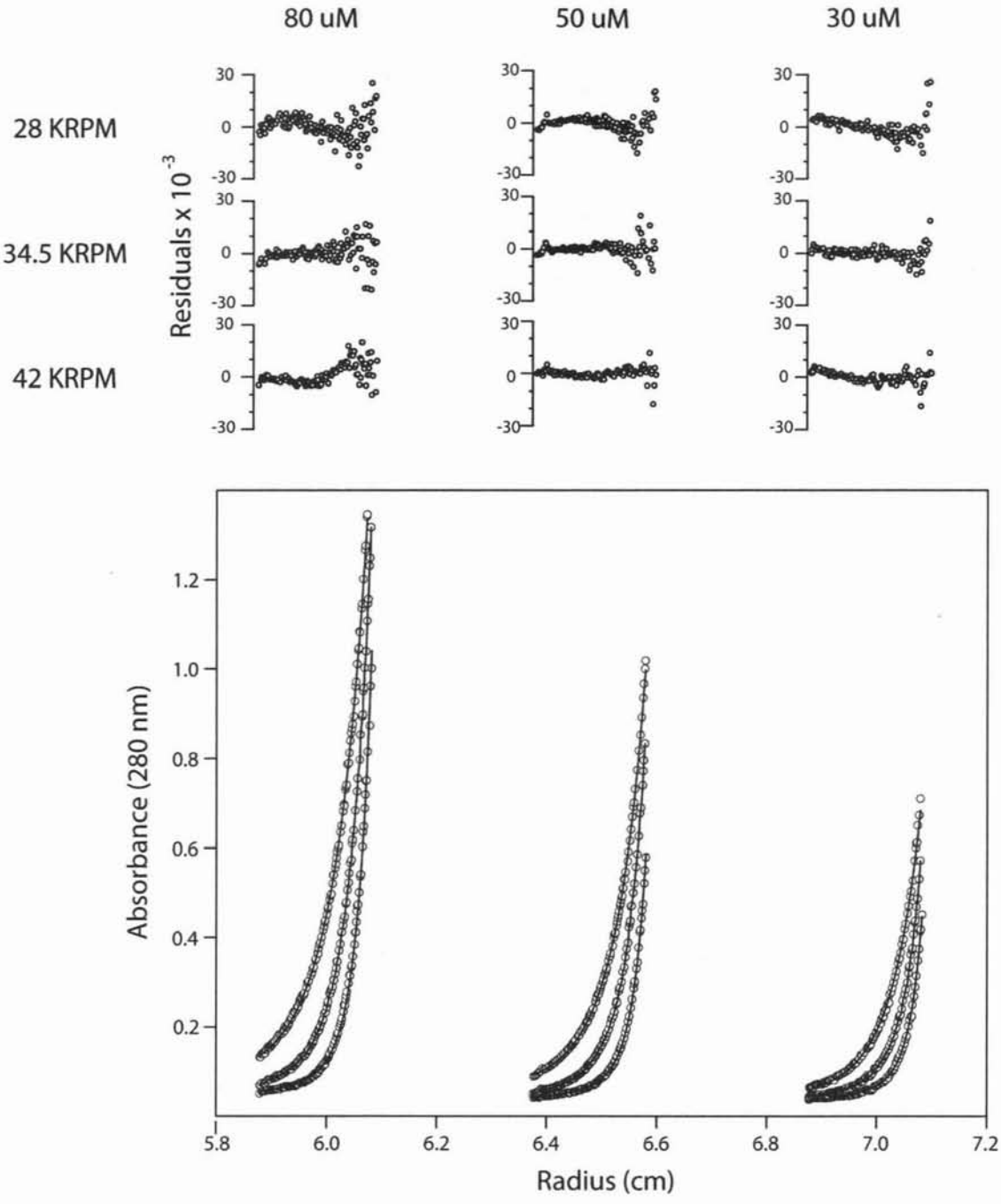


Figure 3-18. Global fit analysis of SH protein in 5 mM C14SB micelles to a monomer-pentamer equilibrium function. The open circles in the lower panel are the data collected at three different concentrations (80, 50, and 30 uM, from left to right) and at three different speeds (28, 34.5, and 42 krpm, more shallow to less shallow exponential for each concentration). The lines in the lower panel represent the global fit to all data. The open circles in the upper panels represent the residuals of each fit.

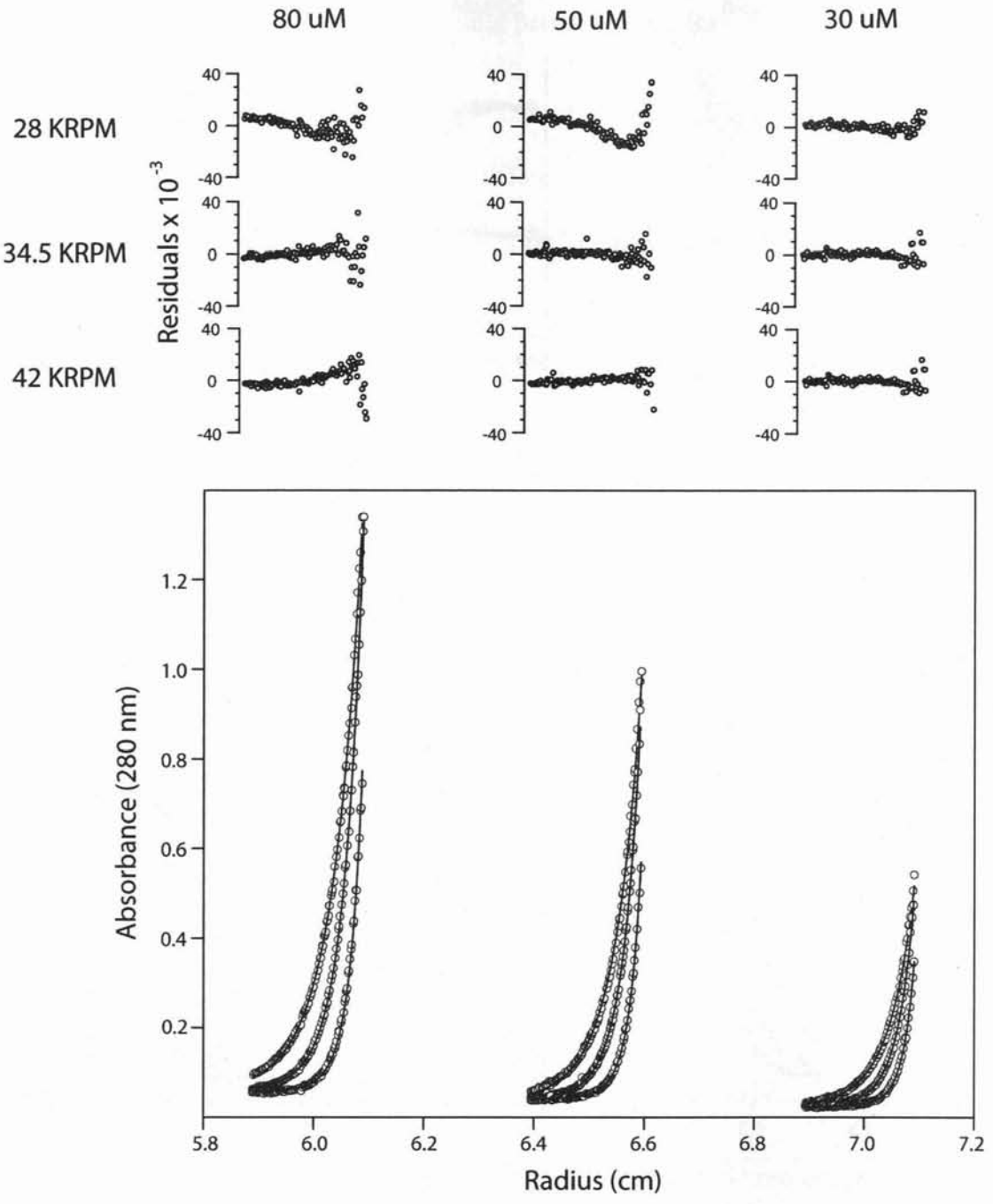


Figure 3-19. Global fit analysis of SH protein in 33 mM C8E5 micelles to a monomer-pentamer equilibrium function. The open circles in the lower panel are the data collected at three different concentrations (80, 50, and 30 uM, from left to right) and at three different speeds (28, 34.5, and 42 krpm, more shallow to less shallow exponential for each concentration). The lines in the lower panel represent the global fit to all data. The open circles in the upper panels represent the residuals of each fit.

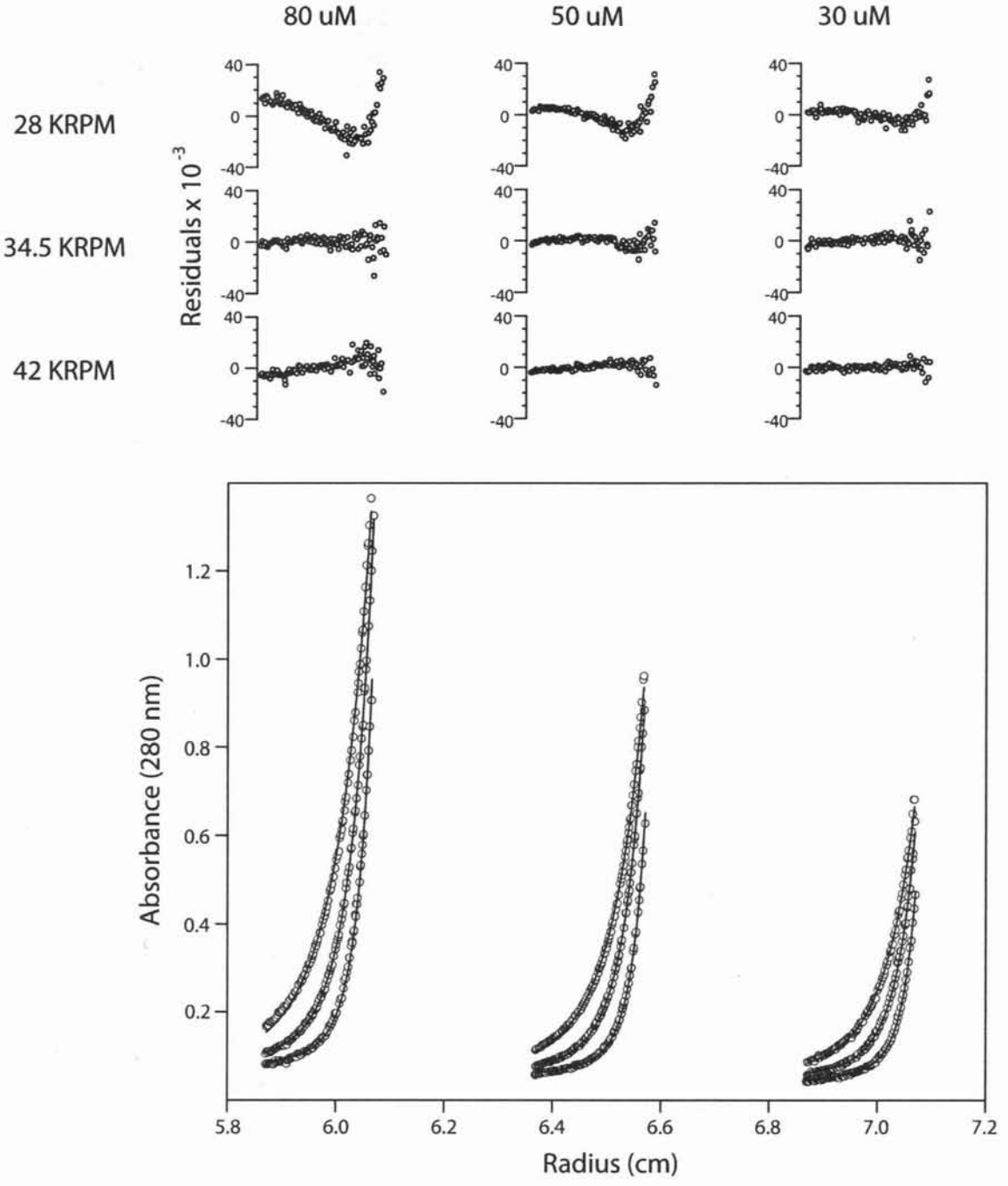


Figure 3-20. Global fit analysis of SH protein in 15 mM DPC micelles to a monomer-pentamer equilibrium function. The open circles in the lower panel are the data collected at three different concentrations (80, 50, and 30 uM, from left to right) and at three different speeds (28, 34.5, and 42 krpm, more shallow to less shallow exponential for each concentration). The lines in the lower panel represent the global fit to all data. The open circles in the upper panels represent the residuals of each fit.

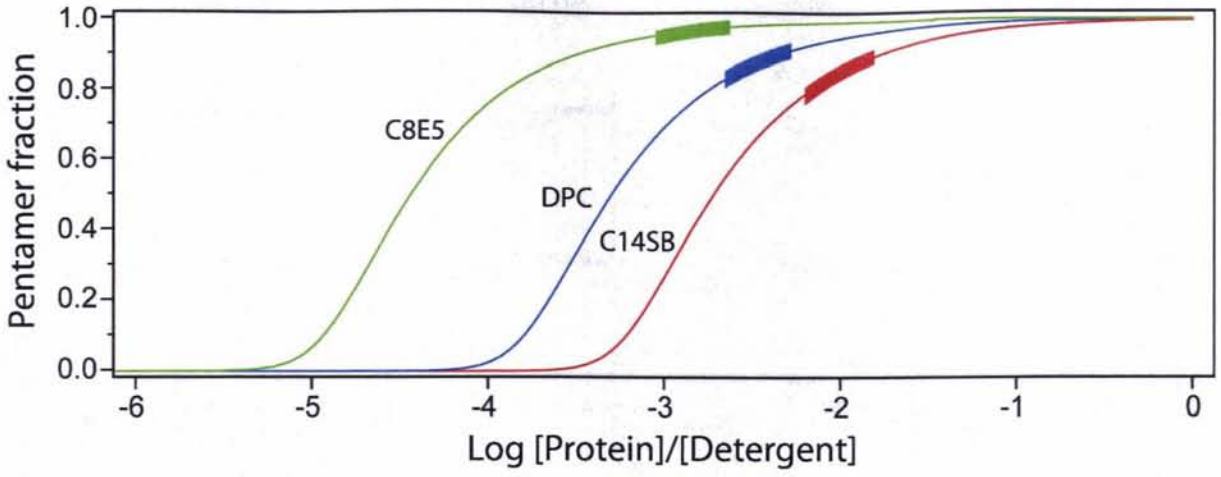


Figure 3-21. A comparison of the SH protein pentamer species equilibrium distribution in C8E5 (green), DPC (blue), and C14SB (red) micelles. The thick bar on the curve represents the observed pentameric populations in the range of protein to detergent ratio used in the experiment.

3.2.4) His22 is involved in SH protein pentamerization in detergent micelles. The His22 is located near the N-terminus of the transmembrane domain and it is the only charged residue that would help in regulating the SH channel activity or association of SH pentamer. The AUC-SE of SHprot-H22A demonstrates that SH protein pentamerization was destabilized by the His22 mutation. None of the data can be fitted to a monomer-pentamer equilibrium model and the protein behaves differently in each detergent (Table 3-3). In C14SB, a monomer-tetramer model gave the best fit with standard free energy of -10.26 kcal/mol (Figure 3-22). In DPC, the His22 mutation causes the protein to display a monomer-hexamer equilibrium, with standard free energy of -25.24 kcal/mol (Fig. 3-23). In the case of C8E5, the detergent which shows the greatest propensity for SH protein pentamer formation, the His22 mutation causes the protein to behave as forming part of a monomer-heptamer equilibrium, with standard free energy of -31.22 kcal/mol (Fig. 3-24).

These results suggest that the His22 is involved in pentamerization of SH protein, at least in the detergent micelles environment. Histidine is a good candidate that can mediate the

association of transmembrane helices. The polar Nδ and Nε atoms of the imidazole ring are capable of being both hydrogen bond donor and acceptor (Zhou et al. 2001). The His37 residue in the transmembrane helix of the Influenza virus M2 proton channel is important for the channel selectivity but also for stabilizing the tetramer structure (Howard et al. 2002). A detailed high resolution structure of SH will help in elucidate the role of His22 for SH pentamerization.

Conditions	GlobalChi^2	Rmsd	$K_{A,app}$	$\Delta G_{A,app}$ (kcal/mol)	K_x	ΔG_x (kcal/mol)	ΔG_x (kcal/mol) (per monomer)
mM C14SB ono-tetramer)	0.843	0.00596	2.696×10^{14} M^{-3}	-19.69	3.37×10^{7}	-10.26	-2.57
5 mM DPC ono-hexamer)	1.067	0.00748	4.342×10^{27} M^{-5}	-37.72	3.30×10^{18}	-25.24	-4.21
3 mM C8E5 ono-heptamer)	0.459	0.00451	6.158×10^{31} M^{-6}	-43.37	7.95×10^{22}	-31.22	-4.46

Table 3-3. Summary of sedimentation equilibrium analysis of SHprot-H22A oligomers in various detergent environments.

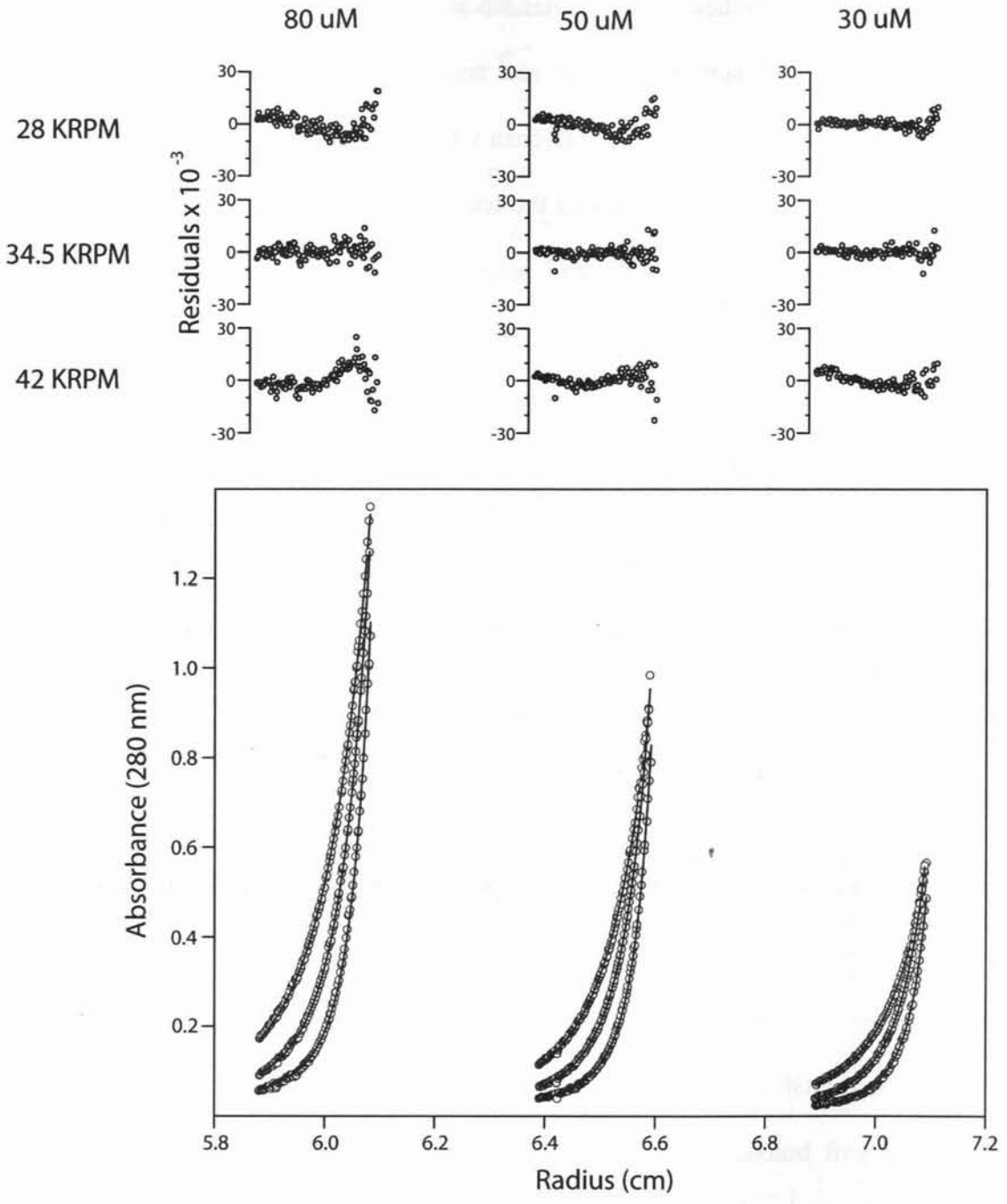


Figure 3-22. Global fit analysis of SHprot-H22A in 5 mM C14SB micelles to a monomer-tetramer equilibrium function. The open circles in the lower panel are the data collected at three different concentrations (80, 50, and 30 uM, from left to right) and at three different speeds (28, 34.5, and 42 krpm, more shallow to less shallow exponential for each concentration). The lines in the lower panel represent the global fit to all data. The open circles in the upper panels represent the residuals of each fit.

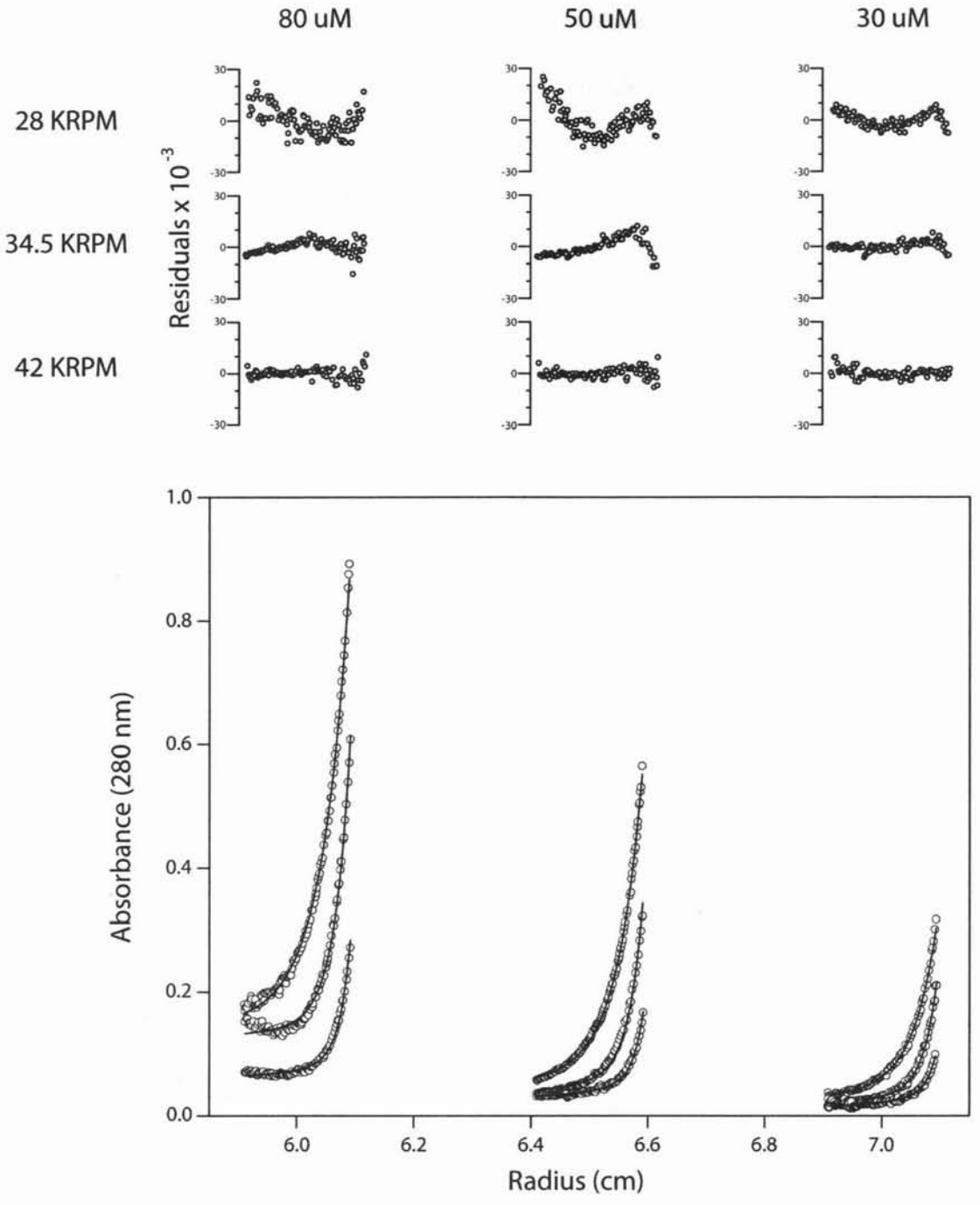


Figure 3-23. Global fit analysis of SHprot-H22A in 15 mM DPC micelles to a monomer-hexamer equilibrium function. The open circles in the lower panel are the data collected at three different concentrations (80, 50, and 30 uM, from left to right) and at three different speeds (28, 34.5, and 42 krpm, more shallow to less shallow exponential for each concentration). The lines in the lower panel represent the global fit to all data. The open circles in the upper panels represent the residuals of each fit.

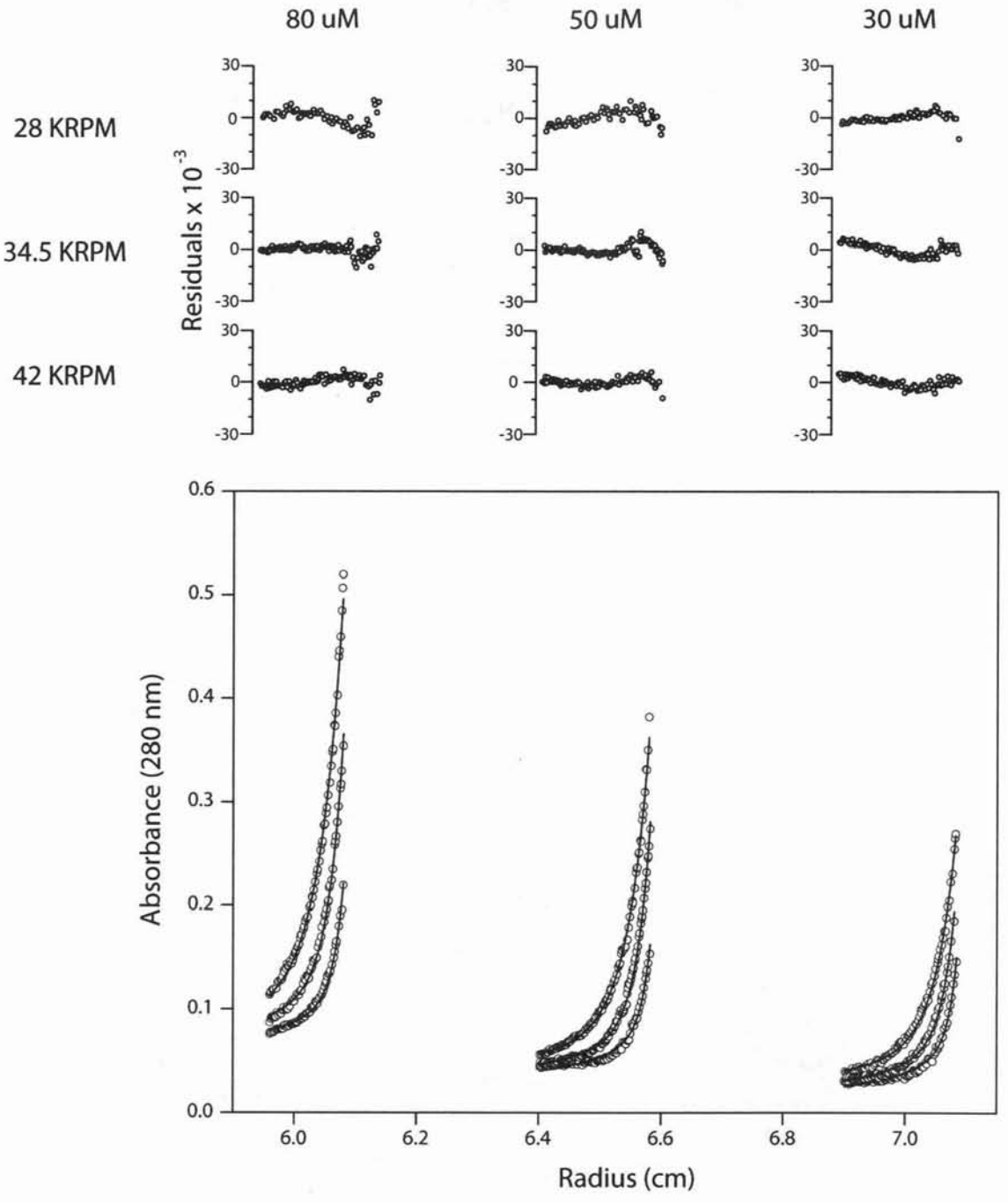


Figure 3-24. Global fit analysis of SHprot-H22A in 33 mM C8E5 micelles to a monomer-heptamer equilibrium function. The open circles in the lower panel are the data collected at three different concentrations (80, 50, and 30 uM, from left to right) and at three different speeds (28, 34.5, and 42 krpm, more shallow to less shallow exponential for each concentration). The lines in the lower panel represent the global fit to all data. The open circles in the upper panels represent the residuals of each fit.

3.2.5) The long-SH-TM shows a mixture of oligomers in C14SB or C8E5 micelles. The long-SH-TM forms non-specific aggregation in the presence of SDS or PFO. The investigation for long-SH-TM pentamer stability was continued with AUC-SE experiment in zwitterionic (C14SB) and non-ionic (C8E5) detergents. DPC was not used for the long-SH-TM because the sample was not solubilized in this detergent.

The data fits well for the long-SH-TM monomer-pentamer equilibrium model in both detergents. The rmsd values are 0.00643 and 0.00627 for C14SB and C8E5, respectively. However, the data can be fit to monomer-tetramer and monomer-hexamer models with good fitting statistics (Table 3-4). AUC-SE fitting results for the long-SH-TM monomer-tetramer, monomer-pentamer, and monomer-hexamer equilibrium models in C14SB and C8E5 micelles are shown in Figure 3-25 and Figure 3-26, respectively. The determination of the best equilibrium model is tricky for the long-SH-TM, as the residuals are all randomly distributed around zero. Examination of the residuals of the fits for all three models does not allow for unambiguous distinction between them especially at the higher protein concentration. However, at high speed and low protein concentration, the difference between these models become more apparent (Figure 3-25C and 3-26C). Judging from the closeness of the residuals distribution to zero, a monomer-tetramer equilibrium model is the best fit for the long-SH-TM in C14SB micelles, whereas a monomer-pentamer equilibrium model is the best fit for the long-SH-TM in C8E5 micelles.

Overall, a mixture of monomer, tetramer, pentamer, and hexamer population co-exist in the long-SH-TM in C14SB and C8E5 micelles and it is concentration dependent. This is not uncommon for AUC-SE experiment for transmembrane helices in detergent micelles. AUC-SE data fitting to mixture of oligomers have been shown for the transmembrane helix of erythropoietin receptor fused to the C-terminal of *Staphylococcal* nuclease (SN-EpoR-TM) and the MS1 peptide solubilized in C14SB (Choma et al. 2000; Ebie and Fleming 2007). The

usage of AUC-SE to monitor oligomerization is limited by weak association, and the presence of low amount of oligomer allows data fitting to other models. In fact, the percentage of long-SH-TM pentamer in C14SB is less than 22% under the experimental condition used (Figure 3-27). The long-SH-TM pentamerization is more favorable in C8E5 compared to C14SB, consistent with the SHprot result. However, only 12% to 47% of long-SH-TM pentamer was presence in C8E5 under the experimental condition (Figure 3-27).

In the future, different approaches for sample design should be explored to obtain the free energy of association data for SH-TM pentamer. The *Staphylococcal* nuclease chimeric construct is an attracting system and it is well-characterized for AUC-SE experiment (Fleming and Engelman 2001).

Conditions	GlobalChi^2	Rmsd	$K_{A,app}$	$\Delta G_{A,app}$ (kcal/mol)	K_x	ΔG_x (kcal/mol)	ΔG_x (kcal/mol) (per monomer)
5 mM C14SB							
mono-tetramer	0.715	0.00645	4.459×10^{12} M^{-3}	-17.26	5.57×10^{5}	-7.83	-1.96
mono-pentamer	0.823	0.00643	1.544×10^{16} M^{-4}	-22.09	9.65×10^{6}	-9.52	-1.90
mono-hexamer	0.995	0.00654	7.626×10^{19} M^{-5}	-27.12	2.38×10^{8}	-11.42	-1.90
33 mM C8E5							
mono-tetramer	1.037	0.00737	9.545×10^{13} M^{-3}	-19.07	3.43×10^{9}	-13.00	-3.25
mono-pentamer	0.810	0.00627	1.408×10^{17} M^{-4}	-23.40	1.67×10^{11}	-15.30	-3.06
mono-hexamer	0.923	0.00622	6.240×10^{20} M^{-5}	-28.37	2.44×10^{13}	-18.25	-3.04

Table 3-4. Summary of sedimentation equilibrium analysis of long SH-TM oligomers in various detergents.

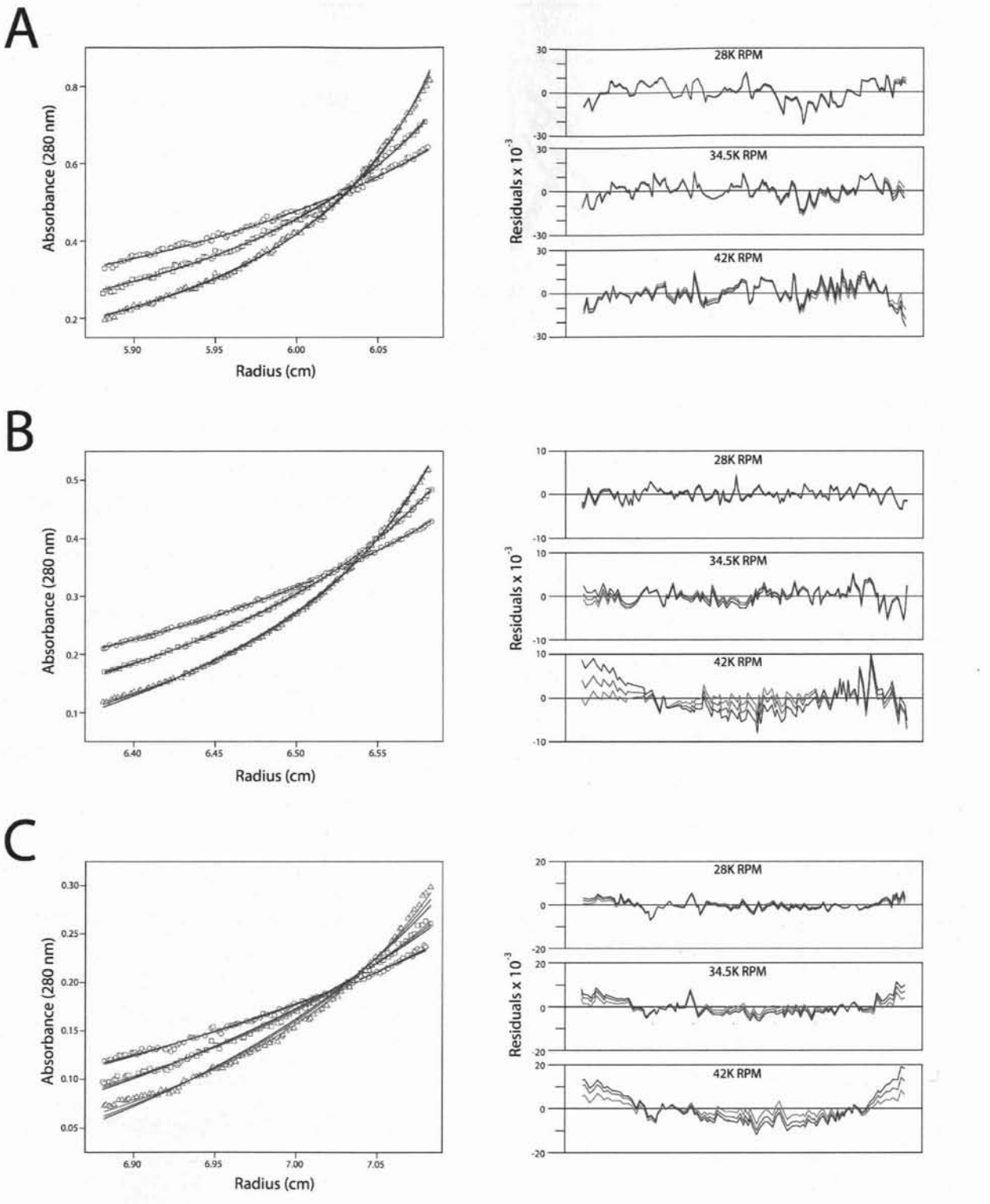


Figure 3-21. Analytical ultracentrifugation of long SH-TM in C14SB micelles. Peptide concentrations are 60 uM (panel A), 36 uM (panel B), and 18 uM (panel C) with 5 mM C14SB. Data at 28 (open circles), 34.5 (open squares), and 42 (open triangles) krpm were globally fit to a function describing sedimentation equilibrium of a monomer in equilibrium with a tetramer (red lines), or pentamer (green lines), or hexamer (blue lines). Residuals of curve fits for each individual data are shown in the right panels. The fit to a mono-tetramer equilibrium gave a good fit to the data, particularly at high speed, 42 krpm.

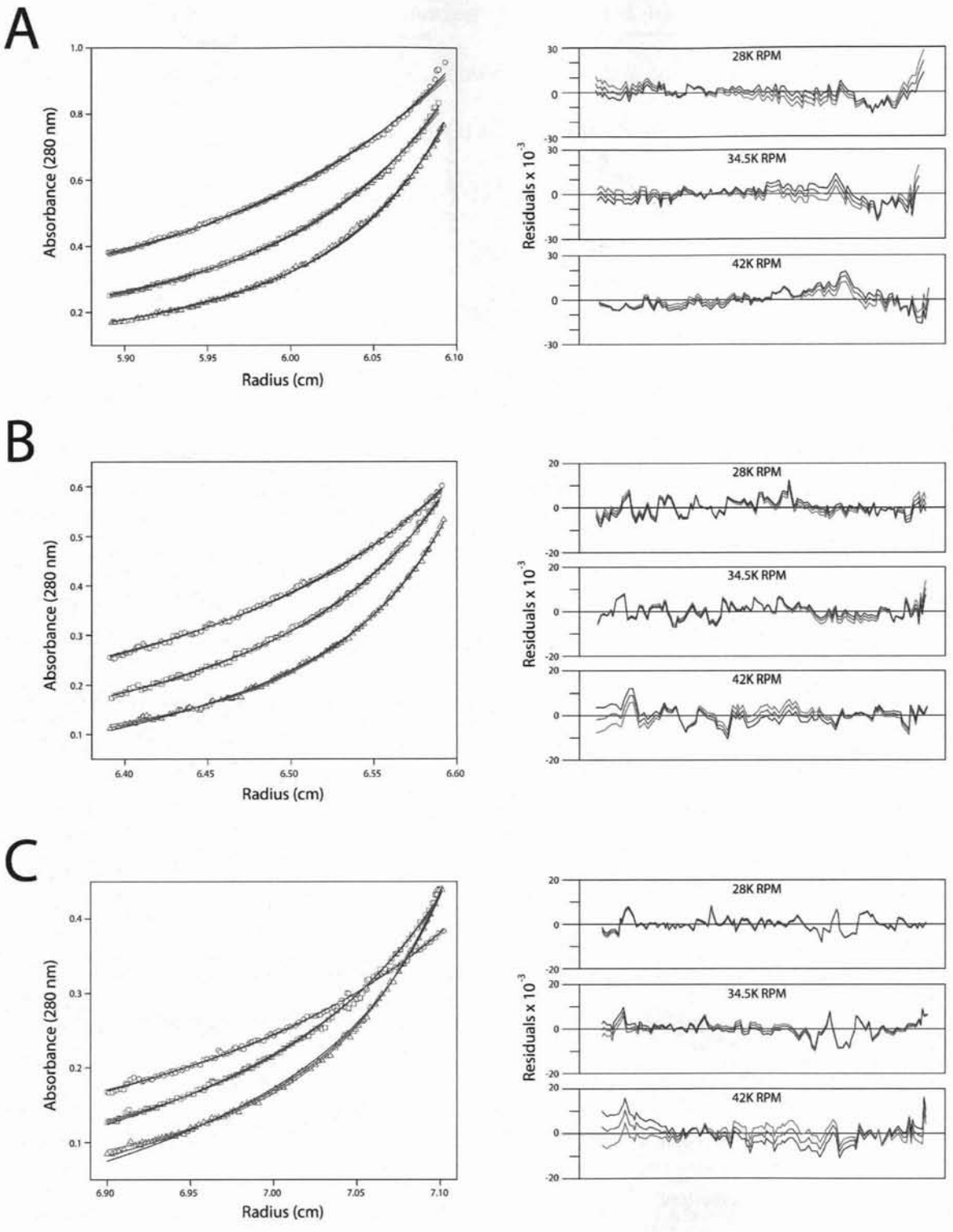


Figure 3-26. Analytical ultracentrifugation of long SH-TM in C8E5 micelles. Peptide concentrations are 72 uM (panel A), 48 uM (panel B), and 24 uM (panel C) with 33 mM C8E5. Data at 28 (open circles), 34.5 (open squares), and 42 (open triangles) krpm were globally fit to a function describing sedimentation equilibrium of a monomer in equilibrium with a tetramer (red lines), or pentamer (green lines), or hexamer (blue lines). Residuals of curve fits for each individual data are shown in the right panels. The fit to a mono-pentamer equilibrium gave a good fit to the data, particularly at high speed, 42 krpm.

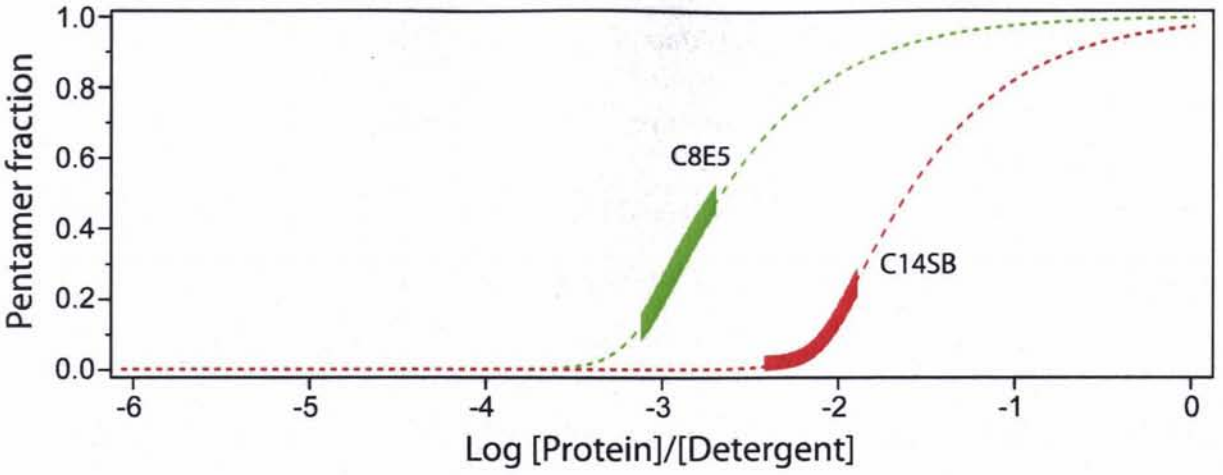


Figure 3-27. A comparison of the long-SH-TM pentamer species equilibrium distribution in C8E5 (green) and C14SB (red) micelles. The thick bar on the curve represents the observed pentameric populations in the range of protein to detergent ratio used in the experiment.

3.2.6) Association energetics of membrane proteins. The progress in understanding the self-association of membrane proteins has always been limited by the available thermodynamic data. Table 3-5 shows the standard free energy of association for several transmembrane α-helices and β-barrel proteins in detergent micelles. The SN-GpA-TM, SN-EpoRh-TM, SN-ErbB-TM, Mp0-TM-K6, MS1, M2-TM, and E-TM are representation of free energy of association for transmembrane helix-helix interactions, from dimer to pentamer. The OMPLA and OmpF are β barrel proteins.

The GpA transmembrane helix dimerization is the most favourable helix-helix association system, with the standard free energy of association per interface equal to -3.50 kcal/mol. The packing of GpA-TM is optimized through van der Waals interactions. A remarkable helix-helix association motif, GxxxG is found in the helix interface. The lack of side-chain in glycine provides a flat surface that allows for other side-chains to pack tightly (Russ and Engelman 2000). The GxxxG motif is the most common motif for helix-helix interaction. To date, several motifs have been identified for helix-helix association, e.g. the leucine-zipper

(Simmerman et al. 1996), QxxS (Sal-Man et al. 2005), and WxxW (Sal-Man et al. 2007) motifs.

To date, no standard free energy of association for pentameric membrane proteins have been documented. In our lab, we have measured the standard free energy of association for the SARS-CoV E protein pentamerization to be -1.89 kcal/mol per interface. The SARS-CoV E protein is a small membrane protein that consists of 76 amino acids with single α-helical transmembrane domain at the N-terminal. The SH protein has higher propensity of pentamerization compared to SARS-CoV E protein.

Proteins	SE models	Detergents	ΔG_x (kcal/mol)	ΔG_x (kcal/mol) (per monomer)	Ref.
SN-GpA-TM	mono-dimer	C8E5	-7.0	-3.50	(Fleming 2002; Fleming et al. 2004)
		C14SB	-5.7	-2.85	
SN-EpoRh-TM	mono-dimer	C14SB	-3.0	-1.50	(Ebie and Fleming 2007)
SN-ErbB-TM	mono-dimer	C8E5	-0.5	-0.25	(Stanley and Fleming 2005)
MP0-TM-K_6	mono-dimer	C8E5	-3.4	-1.69	(Plotkowski et al. 2007; MacKenzie and Fleming 2008)
	mono-tetramer	C8E5	-9.8	-2.45	
MS1	mono-trimer	C14SB	-6.8	-2.27	(Choma et al. 2000; Fleming et al. 2004)
M2-TM (22-46)	mono-tetramer	DPC	-8.3	-2.08	(Stouffer et al. 2005)
E protein	mono-pentamer	C14SB	-9.45	-1.89	Unpublished data from our lab
E-TM (7-42)	mono-pentamer	C14SB	-6.95	-1.39	Unpublished data from our lab
OMPLA	mono-dimer	C14SB	-4.8	-2.10	(Stanley et al. 2006)
OmpF	mono-trimer	C14SB	-26.0	-8.60	(Burgess et al. 2008)

Table 3-5. Summary of reported ΔG_x (kcal/mol) value for various oligomeric systems in detergent micelles measured with AUC-SE. All data were collected at 25°C.

3.3) Structural Determination of SH Protein

3.3.1) Determination of secondary Structure of SHprot, SH-TM and SH-C20 reconstituted in model lipid bilayers by ATR-FTIR. Amide I band is contributed mainly by peptide backbone C=O stretching vibration, which is sensitive to hydrogen bonding. The ATR-FTIR absorbance spectra of SHprot, SH-TM, and SH-C20 in the amide I regions are shown in Figure 3-28. We have assigned secondary structure content for SHprot, SH-TM, and SH-C20 based on experimental data listed in Table 1-1. The full length SH protein (SHprot) shows a major peak centered at 1653 cm^{-1} and a shoulder centered at 1632 cm^{-1} indicating a mixture of secondary structure, α-helix and β-strand (Figure 3-28A, left panel). This is consistent with the result from some secondary structure predictor programs in which α-helical structure in the transmembrane domain and β-strand and unordered structure at both termini at predicted (Fig. 3-29). While amide I of SH-TM show a narrow band centered at 1654 cm^{-1} indicating a large fraction of α-helix (Figure 3-28A, middle panel), the amide I region of SH-20 is centered at 1635 cm^{-1} indicating majority of β-strand structure (Figure 3-28A, right panel).

To determine the relative contents of different secondary structures in these samples, the amide I band of these samples were analyzed using Fourier self-deconvolution (Fig.3-28B). All three samples display distinctive peaks that can be assigned to a particular secondary stucture, except SH-C20, show a broad shoulder from 1690 to 1645 cm^{-1}. Thus, peak assignment for SH-C20 is limited to β-strands only. The quantitative results of secondary structure components for SHprot, SH-TM and SH-C20 reconstituted in DMPC and POPC are summarized in Table 3-6. In the case of DMPC, the relative area of α-helical component is 60% for SHprot, and 74% for SH-TM. There maybe a small percentage of α-helical structure in SH-C20, but less prominent. Secondary structure assignment for SH-C20 is dominated by

β-structure, about 60%. In addition, Fourier self-deconvoluted spectra also shows a small fraction of turns structure in SHprot and SH-TM.

Although these FTIR measurements were performed at room temperature, the frequency of the lipid methylene C-H stretching bands of DMPC (T_m = 23°C) indicated that the membranes were in the gel phase. Therefore, spectra for SHprot and SH-TM were recorded in supported membrane of POPC (T_m = -2°C), which should form a fluid liquid crystal phase. This was indeed observed by a shift of the lipid symmetric stretching vibration, from 2849 to 2852 cm^{-1}, and the anti-symmetric methylene stretching vibration, from 2917 to 2922 cm^{-1} (Tamm and Tatulian 1997). Comparison of the data obtained in DMPC and POPC shows that the number of α-helical residues in SHprot decrease from 40 (DMPC) to 38 (POPC). In contrast, SH-TM is more stable, no changes in secondary structure contents was observed for SH-TM reconstituted in DMPC or POPC.

Taken together, in the full-legth SH protein, the α-helical region is extended beyond the transmembrane domain. Therefore, we propose a long α-helical region for SH protein from Phe14 to Asn52, based on the measured α-helical content by ATR-FTIR and secondary structure prediction.

	SHprot		SH-TM		SH-C20	
assignment	**ν (cm^{-1})**	**relative area (%)**	**ν (cm^{-1})**	**relative area (%)**	**ν (cm^{-1})**	**relative area (%)**
DMPC						
β-strand	1632	34.93 ± 1.25	1636	18.72 ± 0.50	1635	59.10 ± 0.15
α-helix	1653	59.51 ± 0.85	1654	74.01 ± 1.77		
					others	40.90 ± 0.15
Turns	1672	5.56 ± 0.40	1670	7.28 ± 1.57		
POPC						
β-strand	1630	32.84 ± 0.48	1636	21.23 ± 1.41	-	-
α-helix	1653	57.42 ± 0.39	1654	71.46 ± 1.25	-	-
Turns	1674	9.74 ± 0.30	1670	7.31 ± 0.68	-	-

Table 3-6. Assignment, wavenumber (ν), and relative areas of the component peaks deduced from the decomposition of the amide I absorbance bands of SH protein, SH-TM, and SH-C20 reconstituted in DMPC and POPC bilayers supported on germanium plate.

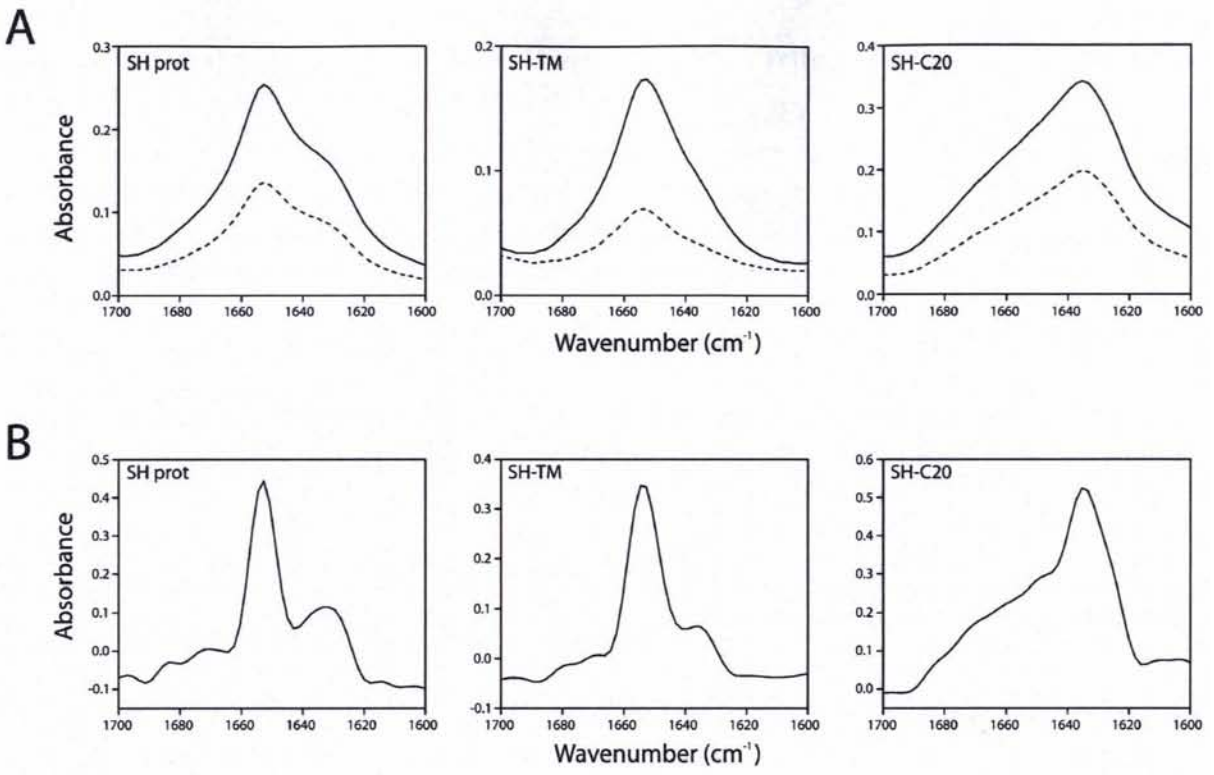


Figure 3-28. ATR-FTIR spectra and Fourier self deconvolution of SHprot, SH-TM and SH-C20 reconstituted in DMPC bilayers supported on germanium plate in D_2O. (A) Polarized ATR-FTIR of SHprot, SH-TM and SH-C20 obtained at parallel (solid line) or perpendicular (dotted line) polarized light. (B) Fourier self deconvolution of these samples obtained with a bandwitdth of 20 cm^{-1} and an enhancement factor of 2.0.

```
                1           11          21          31          41          51          61
                MENTSITIEFSSKFWPYFTLIHMITTIISLLIIISIMIAILNKLCEYNVFHNKTFELPRARVNT
PsiPred         CCCCEEEEEECCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCCCCCCHHHCCC
SSPro           CCCCEEEEEECCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCCCCCCCCEECC
PredProt        CCCCEEEEEECCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCCCCCCCCCCCCCC
SUM             CCCCEEEEEECCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCCCCCCCC.CCC
```

Figure 3-29. Secondary structure prediction of SHprot with PsiPred, SSPro, and Predict Protein online available free program (*http://bioinf.cs.ucl.ac.uk/psipred*, *http://scratch.proteomics.ics.uci.edu*, *http://www.predictprotein.org*). Summary of the predicted secondary structure is in bold. H is refering to α-helical structure, E is β-strand, C is coil structure, and . is unasigned structure. In summary, 55% of α-helical struture (35 aa), 9% of β-strand (6 aa), and 34% of coil structure (22 aa) are predicted.

3.3.2) Amide hydrogen-deuterium exchange of SHprot, SH-TM and SH-C20 reconstituted in model lipid bilayers. The amide II band is used to monitor protein hydrogen-deuterium exchange kinetics. The frequency of the amide II band is due primarily to the peptide backbone N-H bending vibration. Upon hydrogen-deuterium exchange, the peptide N-H become N-D and the frequency of amide II downshifts from near 1545 to 1450 cm^{-1} (~100 cm^{-1}) (Figure 3-30A). In a typical experiment, the amide exchange can be measured on the basis of the decrease in the intensity of the unexchanged amide II.

In the case of membrane protein reconstituted into model lipid bilayers, hydrogen-deuterium exchange experiment can be used to determine the number of residues embedded in the bilayer, such as transmembrane domain. In addition, slow exchange or even resistance to hydrogen-deuterium exchange, is related to protein conformation. For instance, very tight packing of secondary or tertiary structure would prevent exposure of protein to the aqueous environment (Byler and Susi 1986).

Figure 3-30B shows the spectra of amide II band of SHprot, SH-TM, and SH-C20 recorded in H_2O and after one hour exposure to D_2O. The fraction of unexchanged residues was calculated using Equation 1-5. The results of hydrogen-deuterium exchange for SHprot, SH-TM, and SH-C20 reconstituted in DMPC and POPC were listed in Table 3-7. For SHprot, more than half of the protein is accessible to hydrogen-deuterium exchange, only 45% of the protein is protected. In contrast, SH-TM shows a highest fraction of residues resistant to hydrogen-deuterium exchange, only 20% of the protein is accesible to D_2O. This is consistent with a transmembrane domain in which the peptide is embedded in the lipid bilayers and not accessible by D_2O. No difference of hydrogen-deuterium exchange for SHprot and SH-TM reconstituted in DMPC or POPC. The SH-C20 shows a highest number of exchanged residues upon hydrogen-deuterium exchange, 89% of the peptide was exchanged. This is consistent with the predicted topology of SH protein that the protein is evenly divided into

three parts with the transmembrane located in the middle, flanked by two termini (Fig. 1-5). The SH-C20 is not inserted into the membrane and therefore is exposed to D_2O exchange.

Taken together, the results of hydrogen-deuterium exchange for SHprot, SH-TM, and SH-C20 reconstituted in lipid bilayers suggest that the SH protein crosses the lipid bilayer once with a single transmembrane domain. Both N- and C-termini are accessible to D_2O.

	Non-exchange (%)	Number of residues
DPMC		
SHprot	45.23 ± 5.82	30.31 ± 3.90
SH-TM	77.53 ± 7.74	20.16 ± 2.02
SH-C20	10.78 ± 1.49	2.16 ± 0.30
POPC		
SHprot	44.20 ± 5.18	29.61 ± 3.47
SH-TM	79.15 ± 3.81	20.58 ± 0.99

Table 3-7. Results of hydrogen/deuterium exchange for SHprot, SH-TM, and SH-C20 reconstituted in DMPC or POPC bilayers.

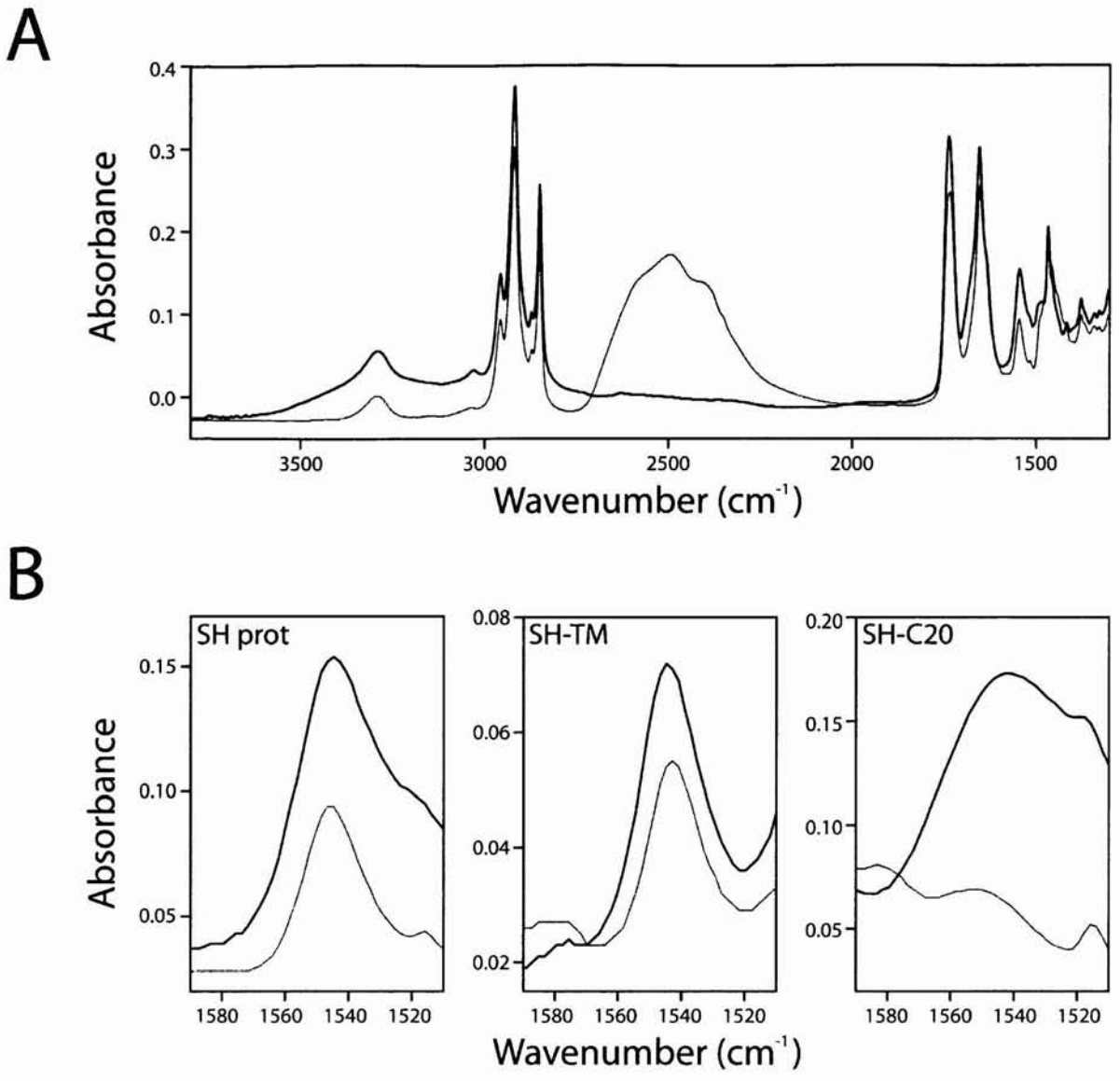


Figure 3-30. Hydrogen-deuterium exchange of SHprot, SH-TM, and SH-C20 reconstituted in supported DMPC bilayers. Dark and grey lines correspond to spectra recorded in H_2O and D_2O, respectively. (A) Overview of SHprot hydrogen-deuterium exchange, the spectra recorded in D_2O has a broad band centered at 2500 cm-1. (B) Amide II band of SHprot, SH-TM, and SH-C20 in H_2O and after I hour exposure to D_2O.

3.3.3) Structural determination of SH protein in detergent micelles by solution NMR. A high resolution structure is required to further characterize the SH protein. Structural study of membrane protein by NMR has confronted many obstacles, from sample preparation to structural determination. The major challenge is to provide a most suitable membrane-mimicking environment, which is the detergent used to solubilize the membrane protein. The environment surrounding membrane proteins stongly affects their function, structure, and dynamics. ^{15}N labelled SHprot was tested in three detergents, namely DPC (medium-chain, zwitterionic), DHPC (short-chain, zwitterionic), and SDS (anionic), to test for the best detergent condition (see Fig. 3-31). These detergents are widely used in solution NMR for membrane proteins. Although SDS is a strong detergent that denature many proteins, well-resolved spectra of membrane proteins have been recorded in this detergent. For instance, the MerF, human FXYD1 and FXYD4 structure was solved in SDS micelles (Howell et al. 2005; Franzin et al. 2007; Teriete et al. 2007). DPC and DHPC have a head group that closely mimics that of the phosphotidylcholine, the most abundant phospholipid in native membrane. To date, many membrane protein structures have been solved in DPC micelles, e.g. KcsA (Yu et al. 2005), human phospholamban (Oxenoid and Chou 2005), Diacylglycerol kinase (Van Horn et al. 2009), and Rv1761c from Mycobacterium tuberculosis (Page et al. 2009). The M2 (Schnell and Chou 2008) and Vpu (Park et al. 2003) structures were solved in DHPC micelles.

In the case of SH protein, the ^{1}H/^{15}N-HSQC spectra shows that the dispersion of the peaks is limited, many resonances are not well-resolved, and peaks are overlapping (Fig. 3-31). Only about 50% of peaks can be observed in SDS and DHPC samples. DPC appears to be the best detergent for SH protein compared to SDS and DHPC, but only about 75% of peaks are observed. Obtaining a good quality ^{1}H/^{15}N-HSQC spectrum is a prerequisite for NMR based structural study. In the future, more conditions should be screened for SH protein

reconstituted into DPC by varying temperature, protein concentration, pH, and salt concentration. Another detergent worth trying for SH protein is the lyso-phosphatidylglycerol group including LPPG and LMPG. They have negatively charged headgroups and long hydrophobic chain. The advantage of this detergent is the negatively charged headgroup that may keep the micelles separated from each other and thereby prevent unfavourable micelle collisions and detergent/protein exchange. The structure of human KCNE1 was solved in LMPG micelles (Kang et al. 2008) and the GPCR double TM fragment was solved in LPPG micelles (Neumoin et al. 2009).

In addition, sample heterogeneity was observed for SHprot-micelles complex in all three detergents tried here. This is evidenced by the resonance of tryptophan indole side chain, Nε-Hε, at region around 10.0-10.5 ^{1}H ppm. There is only one tryptophan residue in the sequence of SHprot, therefore if a single structure is present, only one peak should be observed. However, always two resonances were observed at the lower left region in the 2D ^{1}H/^{15}N-HSQC spectra of SH protein in all three detergents. This indicates two backbone conformations or two different rotameric states of the tryptophan indole side chain.

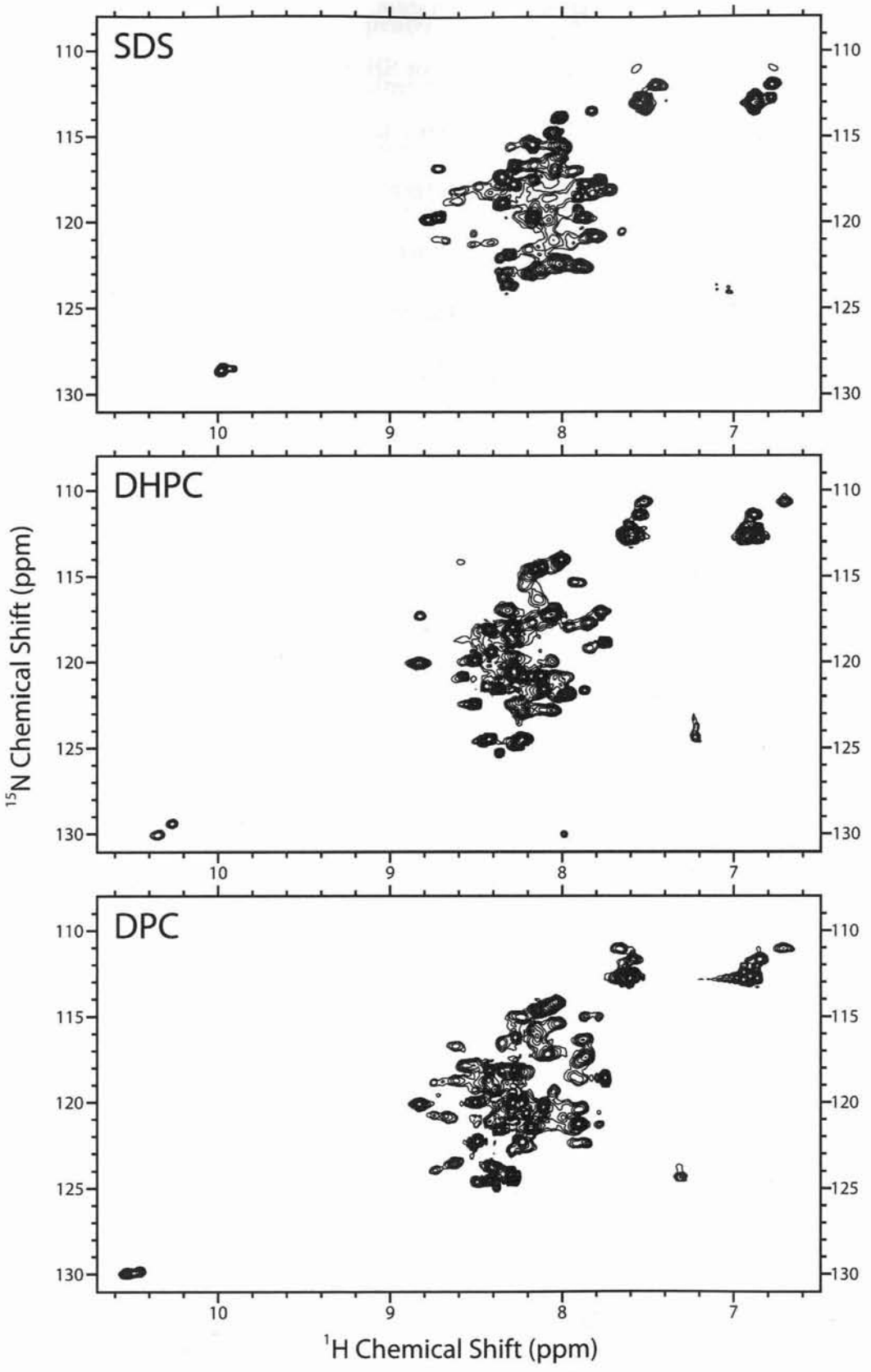


Figure 3-31. The effect of different micellar environments on SH protein illustrated with $^{1}H/^{15}N$-HSQC spectra. The detergents SDS (Top), DHPC (Middle), and DPC (Bottom) were used to obtain the spectra.

3.3.4) SH-TM can be crystallized in β-OG and diffracted to 2.3 Å. We have shown that SH-TM forms a symmetrical pentameric α-helical bundle in lipid bilayers. This makes SH-TM an interesting candidate for crystallization as seen in the successful crystallization of M2-TM, a tetrameric α-helical bundle (Stouffer et al. 2008). As an initial approach to crystallize SH-TM, the MemGold sparse matrix α-helical membrane proteins crystallization screening kit was used. This crystallization screening kit was designed based on 121 successful cases of α-helical membrane proteins crystallization conditions (Newstead et al. 2008).

We have tried crystallization screening of SH-TM reconstituted in β-OG because the M2-TM was successfully crystallized with this detergent (Stouffer et al. 2008). The M2-TM crystal structure was the only α-helical transmembrane peptide ever reported and SH-TM shares similarity with the M2-TM. Both are α-helical transmembrane ion channel and forming stable oligomer, tetrameric bundle for M2-TM, and pentameric bundle for SH-TM.

Plates like crystals were observed in target 1-11, 2-1, 2-11, 2-25, 2-28, and 2-32 (Figure 3-32). The crystallization conditions are listed in Table 3-8. The crystals formed in condition 1-11, 2-11 and 2-25 are bigger and more separated from each other compared to the small clusters of plates like crystal in condition 2-1, 2-28, and 2-32. Therefore, only crystals formed in condition 1-11, 2-11 and 2-25 were collected for diffraction experiments. In collaboration with Dr Julien Lescar (School of Biological Sciences, NTU, Singapore), the 1-11 and 2-11 crystals were diffracted up to 2.2Å resolution whereas 2-25 were diffracted to 3Å. However, these experiments were not yet satisfactory due to phasing problem. The effort to solve SH-TM crystal structure is in progress. Currently, we are trying the isomorphous replacement method (heavy-atom approach) by soaking the crystal in solutions containing heavy-metal such as platinum. Alternatively, multiple wavelength anomalous dispersion method can be employed to solve the SH-TM crystal structure by replacing one of the methionine residue in SH-TM to selenomethionine.

We have applied the same crystallization strategy to full-length SH protein. However, no crystal was observed until now. In future, more detergents should be screened for SH protein crystallization.

Number	Salt	Buffer	pH	Precipitant
1-11	None	0.15 M potassium phosphate	6.5	3.3 M ammonium sulfate
2-1	0.22 M sodium citrate	0.1 M Tris	8.5	35 % v/v PEG 400
2-11	0.05 M sodium sulfate/ 0.05 M lithium sulfate	0.05 M Tris	8.0	35 % v/v PEG 400
2-25	0.2 M ammonium sulfate	0.1 M sodium acetate	4.6	28 % v/v PEG 550 MME
2-28	0.1 M lithium sulfate/ 0.05 M di-sodium hydrogen phosphate	0.05 M citric acid	none	19 % w/v PEG 1000
2-32	0.1 M sodium chloride/4% v/v ethylene glycol	0.1 M MES	6.5	33 % v/v PEG 400

Table 3-8. Crystallization condition for SH-TM reconstituted into β-OG.

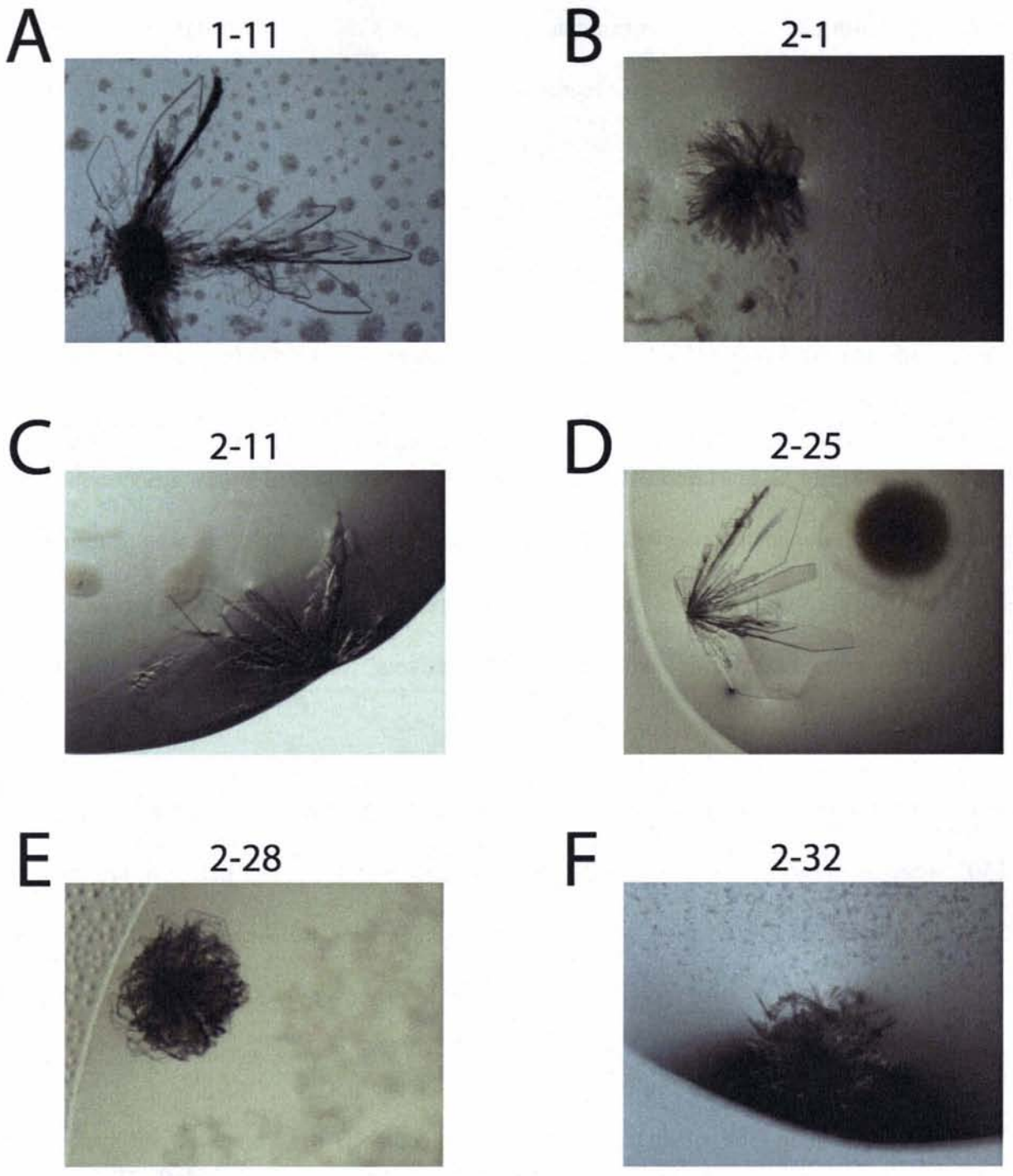


Figure 3-32. Crystal formed by SH-TM reconstituted into β-OG using the MemGold sparse matrix screening kit. Plates like crystals are found in condition 1-11 (A), 2-11 (C), and 2-25 (D). Small cluster of plates like crystals are found in condition 2-1 (B), 2-28 (E), and 2-32 (F).

3.3.5) Modelling of SH-TM pentameric structure by GSMD and SSID. The results of oligomeric size determination for SHprot and SH-TM access by AUC-SE and PFO gel electrophoresis indicates a pentameric structure. Next, this pentameric structure of SH-TM was modelled by combining evolutionary conservation data in GSMD simulation and the orientation restrains of TM α-helices obtained from SSID experiment.

First, conformational search for RSV TM homo-pentamers were done by performing GSMD simulation using the predicted transmembrane domain of SH protein and its homologous sequences (see Fig. 2-2, Materials and Methods). Structures shared by all RSV sequences, and hence evolutionarily conserved, were only observed when the helix tilt was restrained to 25°. Only two complete sets were found, which were left-handed structures. The Cα RMSD between any pair of structures within these 'complete sets' was never higher than 1.2 Å. No other complete sets were found for other helix tilts.

In model 1, the rotational orientation ω (see Materials and Methods) for residue 30 (ω_{30}) was 130°, whereas in model 2, ω_{30} was 240°, i.e., a difference of 110°. The helix tilt was 25° and 20°, respectively. Representative structures for these two sets are shown in Figure 3-33. In fact, the model 1 is consistent with a model for SH-TM proposed by Kochva et al in a similar computational study (Kochva et al. 2003).

To determine which one of these two models is correct, the rotational orientation of isotopically labeled SH-TM α-helices in model lipid bilayers (DMPC) was measured using SSID. In the peptide labeled at L30, the band corresponding to the isotopic $^{13}C=^{18}O$ label is centered at 1592 cm^{-1} (Fig. 3-34A, arrow), as expected for a residue in an α-helical conformation. For the peptide labeled at residue L31, however, this band appeared at a lower frequency, 1576 cm^{-1} (Fig. 3-34B, arrow). As this band is not obscured by either amide I or II, and is well resolved (see Fig. 3-34B), this shift did not affect the measurement.

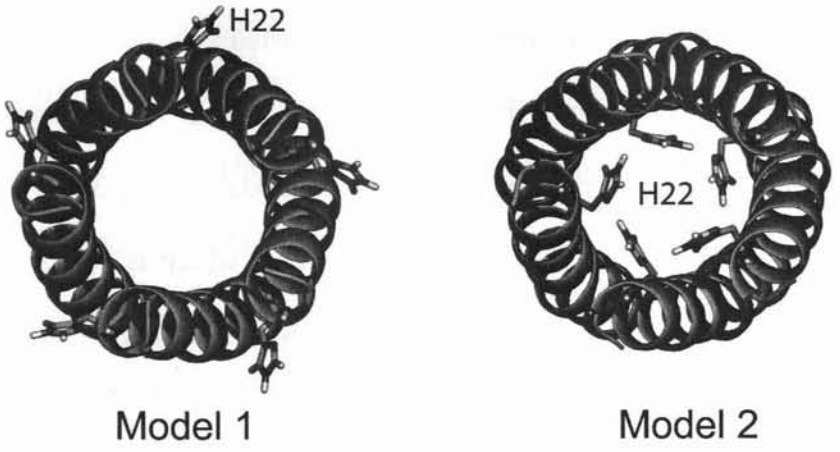


Figure 3-33. Top view of the two pentameric evolutionarily conserved models obtained from the global search. Residue H22 is shown in a stick representation.

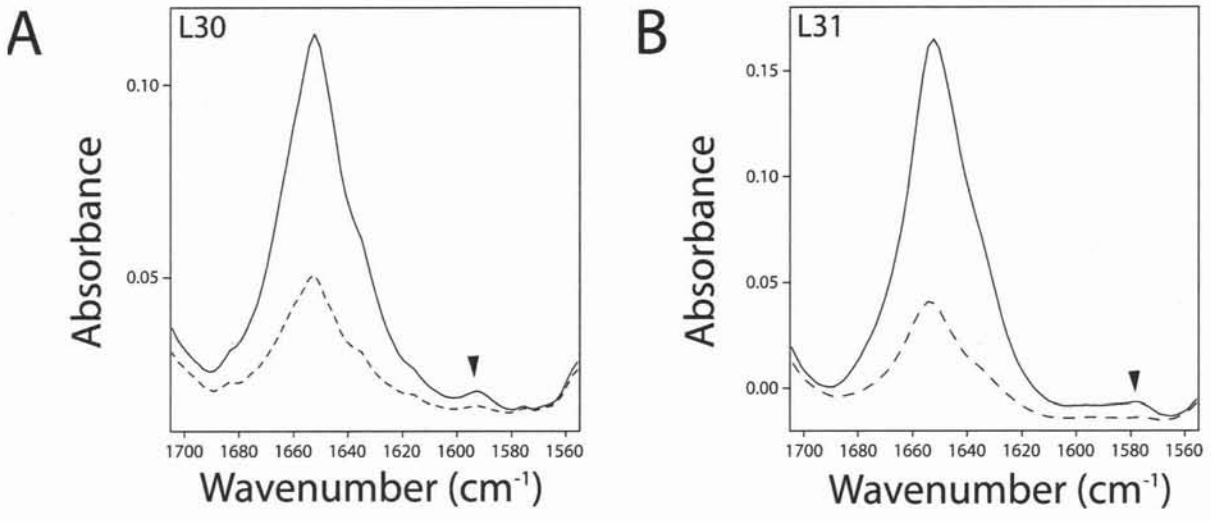


Figure 3-34. Polarized ATR-FTIR absorbance spectra of SH-TM reconstituted in DMPC bilayers in the amide I region of L30 (A) and L31 (B). The position of the labeled C=O stretching band is indicated with an arrow. Solid line and broken line indicate spectra obtained at parallel or perpendicular polarization, respectively.

The dichroic ratios obtained for each label in DMPC are listed in Table 3-9. The orientational parameters for SH-TM was obtained using these dichroic ratios,. The average ω_{30} was 258° (±7.3°). This orientation is only consistent with that of model 2 (Fig. 3-33, right), which is 240°. The helix tilt obtained from SSID was 21.5° (±1.3°), which is also similar to that of model 2. The fractional order parameter was around 0.5. Slices through the pentameric model 2 are presented in Figure 3-35.

	L30		L31	
Measurements	R_{helix}	R_{site}	R_{helix}	R_{site}
DMPC				
a	3.0	2.2	3.2	4.3
b	3.4	2.5	3.5	5.2
c	3.0	2.4	3.6	6.2
POPC				
a	2.7	2.5	3.1	7.5
b	2.7	2.4	3.0	4.5
c	2.6	2.0	2.7	4.3

Table 3-9. Dichroic ratios measured by SSID using the amide A and the isotopic label bands for SH-TM incorporated in DMPC and POPC lipid bilayers.

The infrared measurement was repeated in POPC, which should form a fluid liquid crystal phase. With the data obtained in POPC (Table-3-9), ω_{30} was calculated to be 271° (±16.0°) and the helix tilt β of 30° (±3.4°), in good agreement with the result obtained in DMPC.

The discrepancy between our computational prediction and a similar one performed by Kochva et al is probably due to the different simulation strategies. In Kochva's work, the helix tilt was not restrained, whereas our simulations were performed after restraining the helix tilt every 10 degrees, from 5 to 45 degrees. Restraining the helix tilt ensures that all conformational space is explored, as we have shown previously for influenza A M2 (Torres et al. 2001), where an evolutionarily conserved model showing the correct helix-helix interactions and helix tilt was only found when the tilt was restrained to 30 degrees, but not when the helix tilt was not restrained . In addition, only model 2 is consistent both with our experimental SSID results.

None of the conserved polar residues of SH-TM: T26, S29, T/S35, and N42 (highlighted in the sequence alignment in Fig. 2-2, Materials and Methods), are facing the acyl chains of the lipids (Fig. 35, see boxed residues). Future mutagenesis work will determine the precise role of these residues in inter-helical interactions. Notably, the arrangement of the His22 in this model is consistent with the AUC-SE result of SHprot-H22A in which mutation at this residue destabilizes the pentamer structure. In model 2, the polar Nδ and Nε atoms of the imidazole ring are forming a serial of hydrogen bonding that hold the pentamer together.

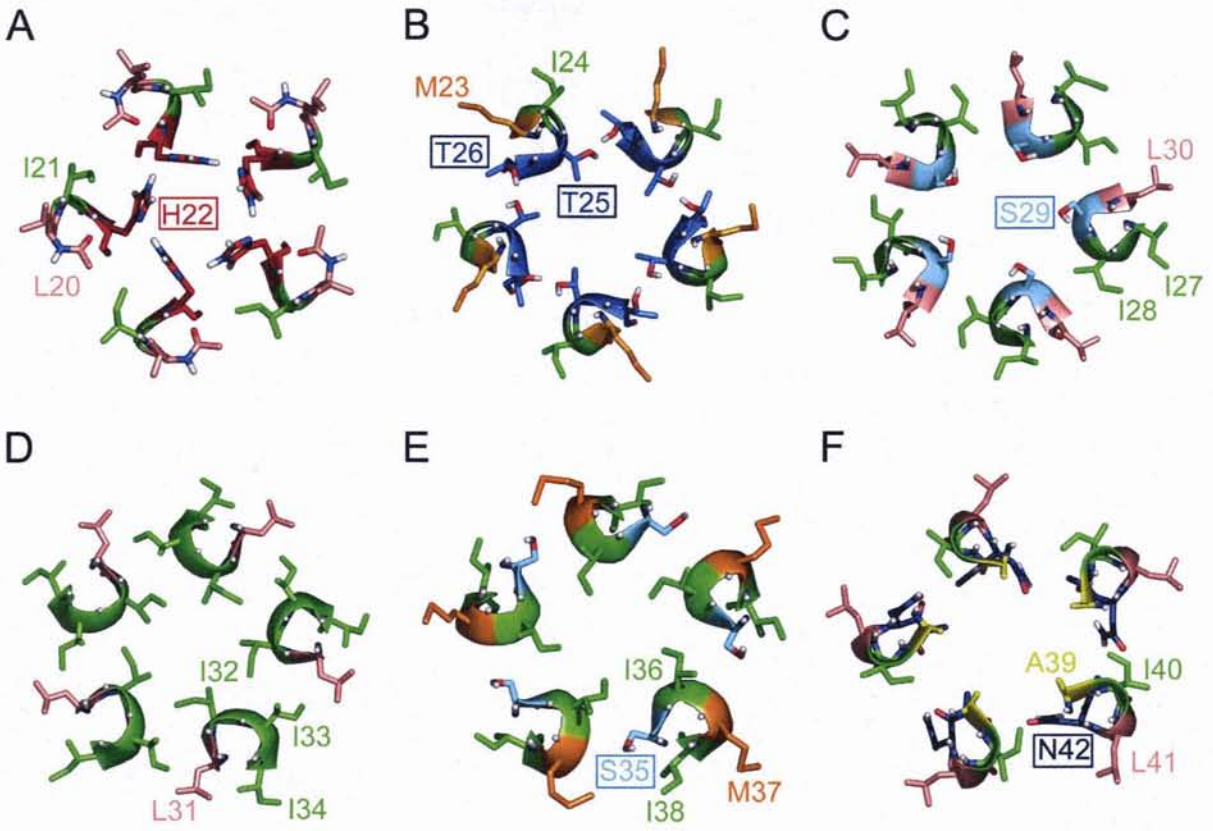


Figure 3-35. Consecutive slices corresponding to model 2 in Fig. 3-33. Residues 20-22 (*A*), 23-26 (*B*), 27-30 (*C*), 31-34 (*D*), 35-38 (*E*) and 39-42 (*F*). The residue numbers are indicated and polar residues are boxed. Color code: I, green; L, salmon; T, marine; S, cyan; M, orange; H, red; A, yellow; and N, blue.

3.3.6) Detection of an intra-helix hydrogen bond in SH-TM by site specific isotopic labeling infrared spectroscopy. As seen in Fig. 3-34B, the $^{13}C{=}^{18}O$ isotope label for L31 is shifted to a lower frequency from 1592 cm^{-1} to 1576 cm^{-1} biased from what is normally observed. According to the harmonic oscillator model, downshift of a vibrational frequency can occur if the reduced mass is increased or the strength of a bond is weakened. Indeed, the simulation results revealed an intrahelical hydrogen bond between the side chain of Ser35 and the backbone carbonyl oxygen from Leu31 (see Fig. 3-36B). By using a peptide with the Ser35 replaced by alanine, we observed a normal shift for $^{13}C{=}^{18}O$ isotope label at L31,

centered at 1589 cm^{-1} (Fig. 3-36A). This proves that an intra-helix hydrogen bond exists in the transmembrane domain of SH protein.

There are high tendency for the serine or threonine residues in the helices to form hydrogen bond to carbonyl oxygen at position *i*-4 in its own helix (Baker and Hubbard 1984). This additional hydrogen bond weakens the strength of carbonyl stretching vibration of amide I and result in the downshift of the vibrational frequency. Statistical studies of high resolution structure of several membrane proteins and soluble proteins have indicated that the intrahelical hydrogen bonding between the side chain of serine or threonine could induce a kink in the α-helix and this may be important for the functionality of the helix (Ballesteros et al. 2000). Some viral ion channels display certain degree of structural flexibility as seen in the influenza A virus M2 protein (Li et al. 2007) and SARS-Cov E protein (Parthasarathy et al. 2008).

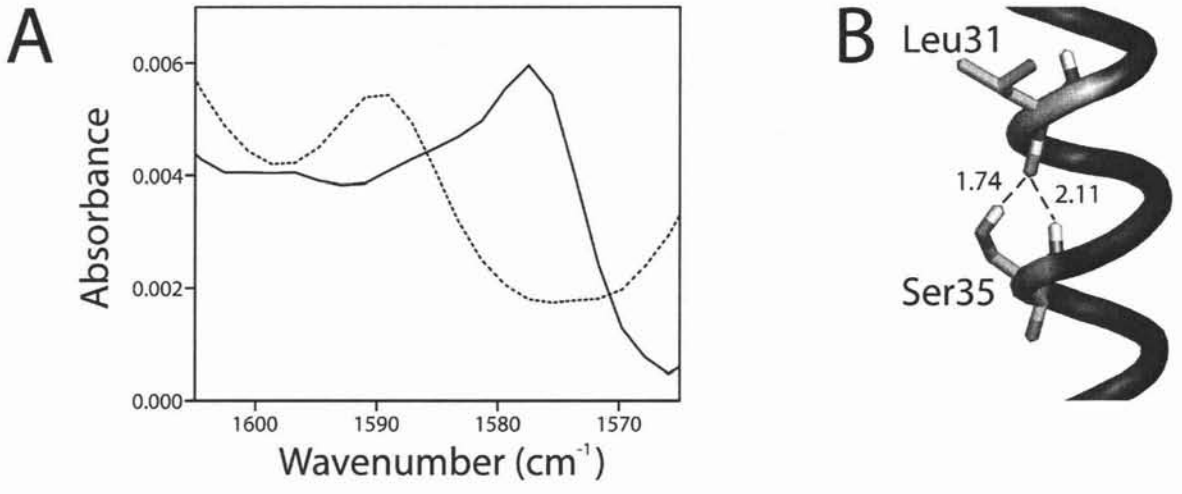


Figure 3-36. (A) ATR-FTIR absorbance spectra of L31 $^{13}C=^{18}O$ isotopic label for wildtype SH-TM (solid line) and S35A mutant (broken line). (B) The hydrogen bonding between Ser35 and Leu31.

3.4) Conductance Studies of SH-TM and SH Protein

3.4.1) Investigation of SH-TM channel activities with BLM. To test the hypothesis that SH protein has ion channel activity, we have collaborated with Dr Alex Gong (School of Biological Sciences, NTU, Singapore) to test the channel activity of SH-TM reconstituted SH-TM in black lipid membranes. The current traces are shown in Fig. 3-37A, which show current reversal at approximately –60 mV. From the linear current-voltage plots shown in Fig. 3-37B, we calculated a conductance of ca. 35 pS, and -68.1 mV for the reversal potential, close to the sodium Nernst potential in our experiment (-59.1 mV), indicating that the channel is cation selective. At voltages below and above the reversal potential, the channel activities were similar (Fig. 3-37C), suggesting a voltage-independent opening probability.

Histidine is the only protonatable residue in the SH-TM, and in our model for SH-TM this histidine is facing the lumen of the pore, which is reminiscent of the orientation of His37 in influenza A M2 proton channel (Kovacs and Cross 1997). Thus, we measured the ion channel activity at a pH below (4.0) and above (7.4) the expected pK_a of histidine in these conditions to test the effect of pH on SH-TM conductance. In the bulk, the expected pK_a of histidine imidazole is 6.4 (Cymes et al. 2005), whereas in the tetrameric pore of M2 from influenza A, the pK_a for the lumen-exposed histidine was determined to be ~ 5.8 (Okada et al. 2001). Fig. 3-37D shows a much lower conductance of SH-TM at pH 4.0, and the current has opposite sign. This strong pH dependence confirms that His22 must be exposed to the lumen of the pore, as in model 2, which is reminiscent to the orientation of His37 in Influenza A M2 proton channel. However, whereas M2 proton channel becomes open at low pH, below the pK_a of histidine (Okada et al. 2001), we observed that conductance of SH-TM is reduced drastically from pH 7.4 to 4.0. This suggests that histidine protonation stops cation transport due to electrostatic effect. A small opposite conductance is observed at pH 4.0, which could indicate chloride transport.

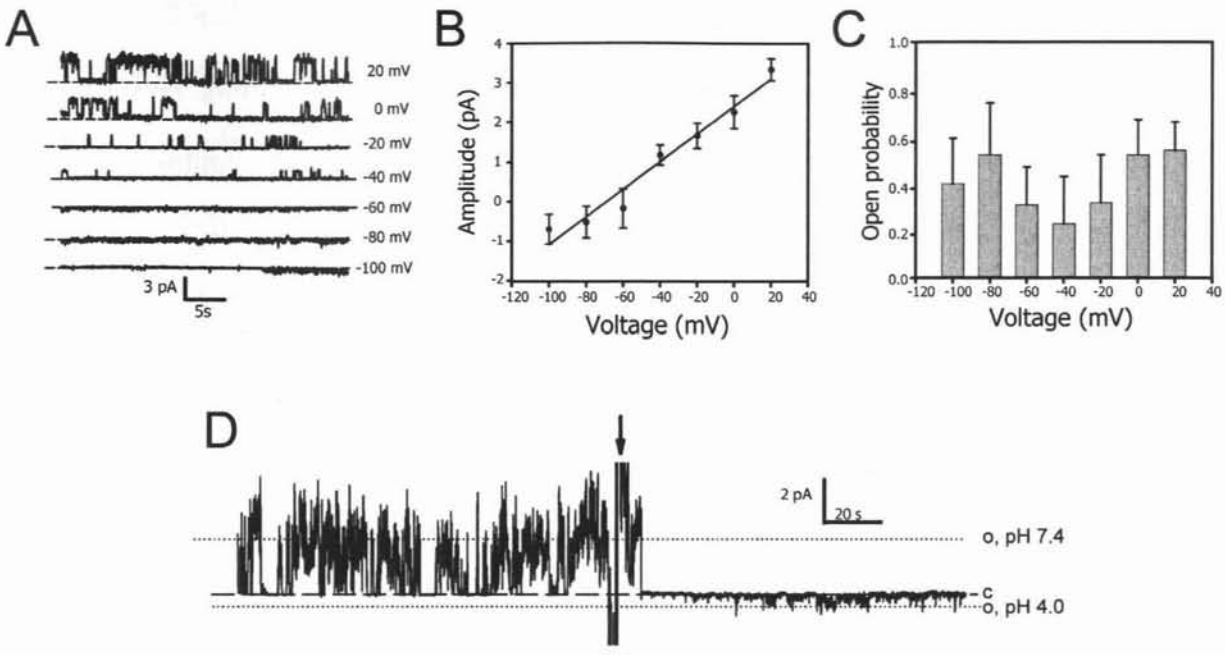


Figure 3-37. Single channel activity observed for SH-TM. (A) Single channel currents recorded over a range of membrane potentials indicated on the right, with openings shown as deviations from the baseline (dashed line). (B) Current-voltage relationship for open state of the channel. (C) The variation of open probability of the channel at different potential. (B, C) Mean of data from 3-5 recordings. (D) Single channel currents at pH 7.4 and pH 4.0 when the bilayer membrane potential was held at +60 mV. The arrow indicates the change from pH 7.4 to pH 4.0.

3.4.2) Patch clamp of HEK293 cells expressing SH protein at two different pH. We have shown that SH-TM is able to form ion channel in artificial bilayers. To provide direct evidence for the full-length SH protein having ion channel activity, the SH protein was expressed in mammalian cells. In collaboration with Dr Soong Tuck Wah (Center for Life Sciences, National University of Singapore), we have examined the channel activity of SH protein and H22A mutant transfected in HEK293 cell line by whole-cell patch clamp experiment. The expression of full-length SH gene, wild type and H22A mutant, were monitored by GFP. The intensity of the fluorescent is correlated with the expression level of functional ion channel (Marshall et al. 1995).

The full-length SH protein displayed channel activity when transiently expressed in HEK293 cells (Figure 3-38A). However, the expression level of SH protein in the HEK293 cells was low as the cells were showing weak fluorescent and the signal was not strong. Nevertheless, the transfected cells produced significant higher channel activity than the controls, the cells transfected with vector alone (Fig. 3-38C). The whole-cell patch clamp recording of SH channel reveals relatively more outward conductance (positive current) than inward (negative current) (Fig. 3-38A). The bath solution used in this experiment, the ACSF solution, contains a high concentration of NaCl (124 mM), whereas the internal solution contain contains a high concentration of potassium ion (145 mM), to mimic the physiological conditions. Under our recording conditions, the estimated equilibrium potentials, E_{Na} and E_K were 65 mV and -87 mV, respectively. Strong selectivity for either of these cations would produce a reversal potential near their corresponding equilibrium potential. If the channel was poorly selective, the reversal potential would have a value somewhere in between E_{Na} and E_K , whereas no selectivity would produce a reversal potential in the mid-point between these values, ~-10 mV. The observed value of reversal potential at 0 mV (Fig. 3-38A) indicates low selectivity between sodium and potassium for SH channel. However, as the expression level of SH protein in HEK293 cells was low, it is difficult to draw the real picture if the signal is weak. In future, it is worth trying to transfect SH protein in other cell lines such as green monkey kidney cells (CV-1) or mouse erythroleukemia (MEL) cells which were used in M2 channel recording (Wang et al. 1994; Chizhmakov et al. 1996).

Molecular modelling of the α-helical region of SH protein shows several polar residues lining the lumen of the pore, including charged residues: His22, Lys43, and His51. The pK_a for lysine and histidine are about 11.1 and 6, respectively. Therefore, histidines are most likely contributing to changes in SH channel activity if the pH of the bath solution was changed from neutral to acidic pH. We have shown in the BLM experiment of SH-TM which

contains His22 was acid sensitive. To test the effect of acidification on channel activity for the full-length SH protein, the bath solution were changed to pH 5.5 after a stable conductance were recorded in neutral pH. In contrast to the control, in which only slight changes were observed upon pH changes (Fig.3-38C), the SH channel responded more actively in acidic solution (Fig.3-38A). Larger outward current was detected upon exposure of SH channel to acidic solution (Fig. 3-38A)

To examine the role played by His22 which lines the pore of SH channel, a H22A mutant was constructed, in which His22 was changed to alanine. At neutral pH, whole-cell patch clamp recording of H22A mutant shows similar channel activity to the wild type (Fig. 3-38B). Also, the H22A mutant channel activity became more active upon exposure to an acidic bath solution, similar to the wildtype SH protein. This suggests that His51 is responsible for acid sensitivity in the full-length SH protein.

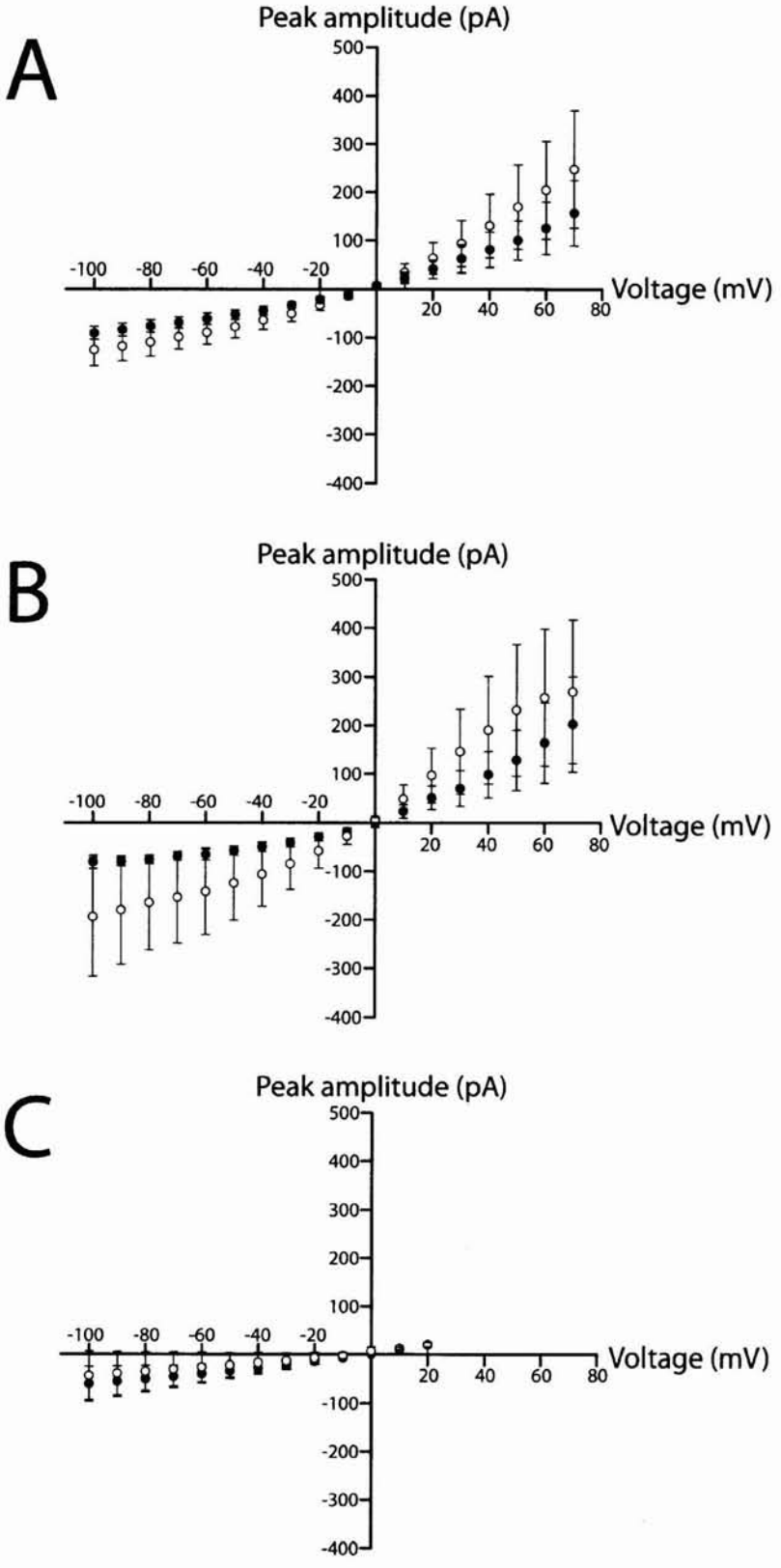


Figure 3-38. Whole-cell current-voltage plot of SH protein (A), H22A mutant (B), and vector alone-transfected HEK293 cells (C) at neutral (pH 7.2) (closed circle), and acidic (pH.5.5) (opened circle) solution. The cells were held at 0 mV and stepped to various potentials from 100 to 70 mV (in steps of 10 mV). Three independent data sets were used to plot the curve.

3.4.3) SH is a viroporin. Viroporin is a general term for ion channels encoded by viruses (Gonzalez and Carrasco 2003). The findings of changes in membrane permeability in cells infected by viruses lead to the classification of this group of proteins. Generally, viroporins are small integral membrane proteins comprised of 60-120 amino acids with α-helical transmembrane domain. They are able to oligomerize in membrane to form pore that allows for passage of ions or small molecules to cross the cells, modifying the membrane permeability. There are two stages of infection in which animal viruses permeabilize the cell membrane; at the early stage during entry of virus to the cells and during expression of the virus genome (Carrasco 1995). While some viral glycoproteins are involved in membrane permeability during virus entry by membrane fusion. viroporins are involved in late membrane permeabilization activity (Carrasco 1995).

It has been suggested that viroporins acts as a virulence factor during infection. They are not essential for virus replication but their presence enhances virus growth. Complicated interactions between viruses and host cells taking place during infection, pathogenicity and virulence are multifactorial traits. Forming an ion channel is one of the strategies for viruses to survive in the host system. The channel activity of viroporins leads to the dissipation of the membrane potential and disruption of cells homeostasis, leading to gradually damage of cells as infection progress.

To date, more than ten viroporins have been identified (Table 3-10). The influenza A virus M2 proton channel is the well studied and elegant example. The atomic structure of M2 has been solved by x-ray crystallography and NMR (Schnell and Chou 2008; Stouffer et al. 2008). The protein is a tetrameric proton channel activated by low pH (Holsinger and Lamb 1991; Sugrue and Hay 1991; Wang et al. 1994; Chizhmakov et al. 1996). The influenza virus enters the cells by endocytic pathway, and virus uncoating taking place in the endosomes. Membrane fusion of the viral envelope and the endosomal membrane is mediated by acid

triggered conformation change of haemagglutinin glycoprotein (Kostolansky et al. 1988). In the acidic environment of endosomes, M2 permits the flux of proton into the virion interior. This acidification causes dissociation between the matrix protein and the ribonucleoprotein core leading to release of the ribonucleoproteins into the nucleus (Martin and Helenius 1991). In addition, the M2 channel activity also modulates the pH of the *trans* Golgi network by equilibrating the acidic pH of the lumen of the Golgi with the cytoplasm, a function important for proper maturation of the acid sensitive haemagglutinin glycoprotein (Ciampor et al. 1992). The channel activity of M2 can be blocked by anti-influenza drugs, amantadine and rimantadine (Wang et al. 1994; Chizhmakov et al. 1996).

The structural information and *in vivo* electrophysiological studies for viroporins are lacking, partially due to the difficulty in handling highly hydrophobic membrane proteins. Nonetheless, structures of the SARS-CoV E protein (Pervushin et al. 2009), the HCV p7 protein(Cook and Opella 2009), and the HIV-1 Vpu proteins (Park et al. 2003; Sharpe et al. 2006) have been solved by NMR. All three channel activities have been confirmed by BLM and the channels can be blocked by hexamethylene amiloride (HMA). The SARS-CoV E protein also displayed channel activity in mammalian whole-cell patch clamp.

In this study, we showed that SH is a viroporin for hRSV. Further studies on hRSV infected cells should gain insight into the significant of SH viroporin activity in the hRSV life cycle. Recently, studies have indicated that SH protein can inhibit apoptosis in several mammalian cell lines by blocking the tumor necrosis factor alpha (TNF-α)-mediated apoptotic signaling pathway (Fuentes et al. 2007). However, ion channels may also control apoptosis in cells (Szabo et al. 2004; Lang et al. 2005; Burg et al. 2006; Madan et al. 2008). Disruption of cells homeostasis is a common sign of apoptosis, leading to plasma membrane depolarization associated with intracellular cation overload and cell volume decreases due to anion and

water efflux (Burg et al. 2006). In fact, the viroporin of Sindbis virus 6K, murine hepatitisvirus E protein, Influenza A M2 protein, HCV p7 protein, poliovirus 2b and 3A protein have been reported to manipulate apoptosis of infected cells (Neznanov et al. 2001; Campanella et al. 2004; Madan et al. 2008). While promotion of apoptosis helps to release the virus, inhibition of apoptosis in host cells during infection gives an advantage to the virus to replicate. In future, drugs that block ion channel of several viroporins, such as amantadine, rimantadine, and HMA could be tested on SH ion channel to obtain further understanding of the channel properties of SH protein.

Virus family	Viroporin	AA residues	Oligomer	Evidence
Picornaviridae	Poliovirus 2B	97	tetramer (Agirre et al. 2002)	Liposome permeability assay (Agirre et al. 2002)
	Coxsackievirus 2B	99	-	Mammalian cell membrane permeabilization assay (van Kuppeveld et al. 1997)
	Poliovirus 3A	87	-	*E. coli* cell membrane permeabilization assay (Lama and Carrasco 1992)
Togaviridae	Semliki forest virus 6K	60	-	*E. coli* cell membrane permeabilization assay (Sanz et al. 1994)
	Sindbis virus 6K	55	-	Mammalian cell membrane permeabilization assay (Sanz et al. 2003)
	Ross River virus 6K	62	-	BLM (Melton et al. 2002)
Retroviridae	HIV-1 Vpu	81	pentamer (Park et al. 2003)	BLM (Ewart et al. 1996)
Paramyxoviridae	HRSV SH	64	pentamer (Gan et al. 2008)	BLM (Gan et al. 2008)
Orthomyxoviridae	Influenza A M2	97	tetramer (Holsinger and Lamb 1991)	Patch clamp (Wang et al. 1994)
	Influenza B NB	100	-	BLM (Sunstrom et al. 1996)
	Influenza B BM2	108	tetramer (Balannik et al. 2008)	Patch clamp (Mould et al. 2003)
Reoviridae	Avian reovirus p10	98	-	Mammalian cell membrane permeabilization assay (Bodelon et al. 2002)
Flaviviridae	HCV p7	63	hexamer/ heptamer (Griffin et al. 2003; Clarke et al. 2006)	BLM (Griffin et al. 2003)
Phycodnaviridae	*Paramecium bursaria* chlorella virus Kcv	94	tetramer (Shim et al. 2007)	Patch clamp (Plugge et al. 2000)
Coronaviridae	SARS-CoV E protein	76	pentamer (Torres et al. 2006)	Patch clamp (Pervushin et al. 2009)
	Murine hepatitis virus E protein	83	-	*E. coli* cell membrane permeabilization assay (Madan et al. 2005)

Table 3-10. List of viroporins. The Table is adapted from Gonzalez and Carrasco 2003 (Gonzalez and Carrasco 2003).

4) CONCLUSIONS

4.1) General remarks.

Viral respiratory diseases pose a serious threat to human health. They are the most widespread infections and effective antiviral agents are limited. The discovery of viroporins in many viruses has shed a light for antiviral therapy. The hRSV is one of the main culprits for viral respiratory diseases causing millions of clinical cases each year. hRSV infection is the prime cause of bronchiolitis. As a consequence, there is an urgent need for antiviral drugs and vaccines. SH has been proposed to be a viroporin for hRSV and a possible target for antiviral therapy. The results presented in this thesis have provided several aspects of hRSV SH protein structure and function.

We have provided evidence that SH assembles into pentamers and the process is driven by the transmembrane domain. We have shown how different detergents affect SH pentamerization stability. In the presence of anionic detergents (PFO), only SH-TM could maintain the pentameric structure but not the full-length protein or a longer version of SH-TM with additional seven residues at the N-terminus, SNAFWPY. SH pentamerization is most favourable in non-ionic detergents. In addition, we have found that His22 in the transmembrane domain is important for pentamerization specificity in the micelle environment.

It is still a difficult task to obtain thermodynamic data for membrane proteins oligomerization in lipid bilayers. The TOXCAT is applicable to dimerization system (Russ and Engelman 1999) and the thiol-disulfide interchange in lipid bilayers requires formation of disulfide-bond (Cristian et al. 2003). On the other hand, the detergent micelles allows for direct measurement of membrane protein association. In micelles, the detergent molecules arrange themselves around a membrane protein, providing a membrane mimetic environment to study

membrane protein interactions. As a consequence of this arrangement, the membrane protein is structurally more flexible than when inserted into lipid bilayers. The dimensionality of the lipid bilayers limits conformational freedom of membrane proteins and promotes specificity in oligomerization. Thus, the oligomerization should be even more favourable in lipid bilayers. In fact, the Influenza A M2 protein tetrameric conformation has been shown to be more favourable in DLPC (Cristian et al. 2003). Therefore, our prediction is that the pentameric structure of SH protein should be even more stable in lipid bilayers than in detergent.

We have obtained a model for the pentamerizing transmembrane α-helices of SH protein based on site specific infrared dichroism determined rotational orientation of isotopically labeled residues and GSMD simulation with evolutionary conserved data. This model should provide the basic for mutagenesis studies to elucidate important residues for SH, both functional and structural.

In addition, an intra-helical hydrogen bond was detected between the backbone carbonyl oxygen of Leu31 and the side chain of Ser35 in the SH-TM as evidenced by a shift in the isotopically labeled carbonyl group for Leu31. To our knowledge, this the first instance of such spectral effect observed. This implies a possible application for isopically labeled ATR-FTIR spectroscopy to detect hydrogen bonding network within the α-helical transmembrane domain, either intra-helical hydrogen bonding between the backbone carbonyl oxygen and the polar side chain at position i+4 or interhelical hydrogen bonding involving backbone carbonyl oxygen.

Attempts to obtain high resolution structure for SH protein by NMR was hampered by poor spectral resolution due to possible sample heterogeneity, improper conditions or detergent micelles used to solubilize SH protein, a common problems encounter with structural study of

membrane proteins (Sanders and Sonnichsen 2006). We have demonstrated in SE-AUC that SH protein is able to form pentamer in DPC, one of the detergents we used for SH protein solution NMR. A large membrane protein-micelle complex would result in line broadening and broad resonances lead to loss of spectrum resolution. Furthermore, we have shown in ATR-FTIR that SH protein reconstituted into lipid bilayers contains a large fraction of α-helical structure, which would contribute to poor dispersion in solution NMR spectroscopy. In future, more detergents and conditions should be screened to improve the quality of SH protein NMR spectra. Alternatively, different sample labeling strategy should be explored such as perdeuterated protein, selective ^{15}N and ^{13}C labeling or segmental labeling.

As an effort to obtain high resolution structure for SH protein or SH-TM, we tried to crystallize the SH protein and SH-TM in the presence of β-OG detergent. To our excitement, SH-TM can be crystallized in the presence of β-OG and the crystal diffracted to 2.3Å. We are now working to solve the structure for this crystal.

Functionally, conductance studies of SH-TM by BLM and mammalian whole-cell patch clamp for full-length SH protein have confirmed that SH is indeed a viroporin for hRSV. The channel is pH sensitive, presumably regulated by His51. Our results provide experimental evidence that SH is a pentameric ion channels.

4.2) Proposed general channel activation mechanism for SH protein. Based on the structural and functional results obtained in this study, we propose a mechanism of channel activation for SH channel (Fig. 4-1). It is possible that protonation of His51 opens up the channel and then subsequent protonation of His22 further widens the channel, allowing for higher ion flow. In the SH-TM, i.e., with His22 and without residue His51, protonation of

His22 is not sufficient to disrupt helix packing, the imidazole ring would arrange closer to each other in a way that repels cations.

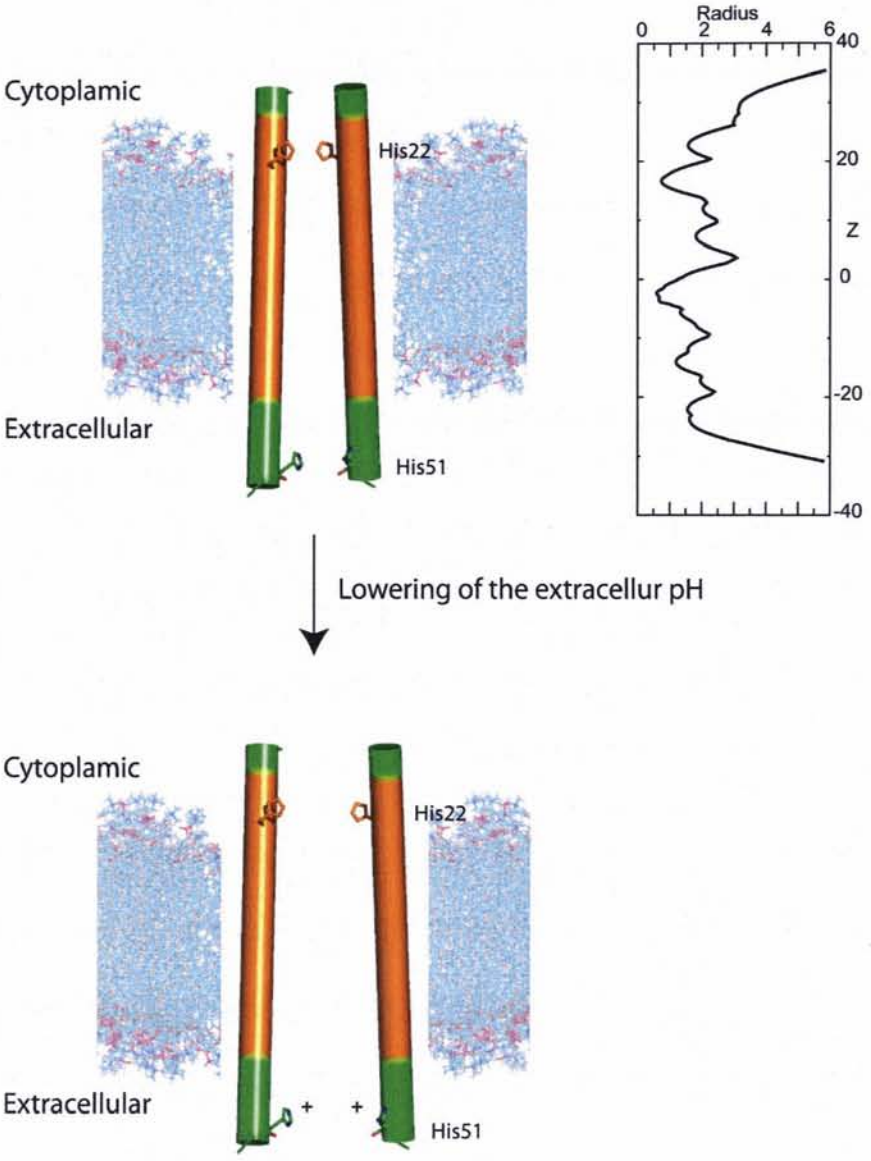


Figure 4-1. Proposed general channel activation mechanism for SH protein. The helices are shown in cylinder, with TM coloured in orange. The His22 and His51 are shown in sticks. For ease of viewing, only two out of five helixes are shown in this picture. On Top right, the channel outline (radius) is shown as a function of Z cordinate. Upon exposure to an acidic solution, His51 gets protonated and repels each other to open the channel.

From the topology of SH protein, the protein transverses the membrane once with the N-terminus facing the cytoplasm and the C-terminus being oriented extracellularly (Collins and

Mottet 1993). The His51 located at the C-terminus of the helix is directly exposed to the acidic environment, and responds to pH changes. Protonation of His51 would disrupt the packing of the helices and allow for a wider opening of the pore that permits more ions to pass through the channel as seen in the M2 proton channel. The His22 may be responding to pH changes in a different mechanism.

We have shown in SE-AUC experiment that the H22A mutant destabilized SH pentamer in C14SB, C8E5, and DPC micelles. The observation that the channel activity and responses to pH changes were not affected by this mutation suggests that it may still form pentamers in the cell membrane. The pentamers may be more stable after post-translational modification of the protein when expressed in mammalian cells. SH protein has been reported to be glycosylated and phosphorylated after post-translational modification (Collins and Mottet 1993; Rixon et al. 2005). The Influenza A M2 proton channel with His37 mutated to alanine show channel activity that is not affected by pH changes when expressed in Xenopus laevis oocytes, although this mutation destabilizes tetramerization in detergent micelles (Pinto et al. 1992; Howard et al. 2002).

The involvement of histidine as a pH sensor has been observed in other ion channels. The influenza M2 proton channel is one of the well-characterized pH sensitive ion channel with His 37 responsible for pH sensitivity and Trp41 forming a gate at the exit of the channel, the C-terminus (Schnell and Chou 2008; Stouffer et al. 2008). Other examples of histidine regulated pH sensitive ion channels include the bacterial potassium channel KscA (His25) Takeuchi et al. 2007), the potassium channel ROMK1 (four histidine residues are involved) Chanchevalap et al. 2000), the potassium channel TASK1 (His98) (Yuill et al. 2007), potassium channel TREK-1 (His126) and TREK-2 (His151) (Sandoz et al. 2009). Further mutagenesis studies involving His51 or double-mutant of His22 and His51 are necessary to confirm our results.

References

Adair, B.D., and Engelman, D.M. 1994. Glycophorin A helical transmembrane domains dimerize in phospholipid bilayers: a resonance energy transfer study. *Biochemistry* **33:** 5539-5544.

Adams, P.D., Arkin, I.T., Engelman, D.M., and Brunger, A.T. 1995. Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Nature Structural Biology* **2:** 154-162.

Agirre, A., Barco, A., Carrasco, L., and Nieva, J.L. 2002. Viroporin-mediated membrane permeabilization. Pore formation by nonstructural poliovirus 2B protein. *J Biol Chem* **277:** 40434-40441.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. 2002. *Molecular Biology of the cell*, 4th ed. Garland Science, New York.

Arkin, I.T., Adams, P.D., MacKenzie, K.R., Lemmon, M.A., Brunger, A.T., and Engelman, D.M. 1994. Structural organization of the pentameric transmembrane alpha-helices of phospholamban, a cardiac ion channel 18. *EMBO Journal* **13:** 4757-4764.

Arkin, I.T., MacKenzie, K.R., and Brunger, A.T. 1997. Site-directed dichroism as a method for obtaining rotational and orientational constraints for oriented polymers. *Journal of the American Chemical Society* **119:** 8973-8980.

Arkin, I.T., MacKenzie, K.R., Fisher, L., Aimoto, S., Engelman, D.M., and Smith, S.O. 1996. Mapping the lipid-exposed surfaces of membrane proteins. *Nature Structural Biology* **3:** 240-243.

Bachi, T., and Howe, C. 1973. Morphogenesis and ultrastructure of respiratory syncytial virus. *J Virol* **12:** 1173-1180.

Baker, E.N., and Hubbard, R.E. 1984. Hydrogen bonding in globular proteins *Progress in Biophysics and Molecular Biology* **44:** 97-179.

Balannik, V., Lamb, R.A., and Pinto, L.H. 2008. The oligomeric state of the active BM2 ion channel protein of influenza B virus. *J Biol Chem* **283:** 4895-4904.

Ballesteros, J.A., Deupi, X., Olivella, M., Haaksma, E.E.J., and Pardo, L. 2000. Serine and threonine residues bend α-helices in the χ1= g- conformation *Biophysical Journal* **79:** 2754-2760.

Barth, A., and Zscherp, C. 2002. What vibrations tell us about proteins. *Q Rev Biophys* **35:** 369-430.

Blount, R.E., Jr., Morris, J.A., and Savage, R.E. 1956. Recovery of cytopathogenic agent from chimpanzees with coryza. *Proc Soc Exp Biol Med* **92:** 544-549.

Bodelon, G., Labrada, L., Martinez-Costas, J., and Benavente, J. 2002. Modification of late membrane permeability in avian reovirus-infected cells: viroporin activity of the S1-encoded nonstructural p10 protein. *J Biol Chem* **277:** 17789-17796.

Briggs, J.A.G., Torres, J., and Arkin, I.T. 2001. A new method to model membrane protein structure based on silent amino acid substitutions. *Proteins: Structure, Function, and Genetics* **44:** 370-375.

Brown, G., Jeffree, C.E., McDonald, T., Rixon, H.W., Aitken, J.D., and Sugrue, R.J. 2004. Analysis of the interaction between respiratory syncytial virus and lipid-rafts in Hep2 cells during infection. *Virology* **327:** 175-185.

Brown, G., Rixon, H.W., and Sugrue, R.J. 2002. Respiratory syncytial virus assembly occurs in GM1-rich regions of the host-cell membrane and alters the cellular distribution of tyrosine phosphorylated caveolin-1. *J Gen Virol* **83:** 1841-1850.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54:** 905-921.

Bu, Z., and Engelman, D.M. 1999. A method for determining transmembrane helix association and orientation in detergent micelles using small angle x-ray scattering. *Biophysical Journal* **77:** 1064-1073.

Bukreyev, A., Whitehead, S.S., Murphy, B.R., and Collins, P.L. 1997. Recombinant respiratory syncytial virus from which the entire SH gene has been deleted grows efficiently in cell culture and exhibits site-specific attenuation in the respiratory tract of the mouse. *Journal of Virology* **71:** 8973-8982.

Burg, E.D., Remillard, C.V., and Yuan, J.X.J. 2006. K+ channels in apoptosis. *Journal of Membrane Biology* **209:** 3-20.

Burgess, N.K., Stanley, A.M., and Fleming, K.G. 2008. Determination of membrane protein molecular weights and association equilibrium constants using sedimentation equilibrium and sedimentation velocity. *Methods Cell Biol* **84:** 181-211.

Byler, D.M., and Susi, H. 1986. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers* **25:** 469-487.

Campanella, M., de Jong, A.S., Lanke, K.W., Melchers, W.J., Willems, P.H., Pinton, P., Rizzuto, R., and van Kuppeveld, F.J. 2004. The coxsackievirus 2B protein suppresses

apoptotic host cell responses by manipulating intracellular Ca2+ homeostasis. *J Biol Chem* **279:** 18440-18450.

Carrasco, L. 1995. Modification of membrane permeability by animal viruses. *Adv Virus Res* **45:** 61-112.

Chanchevalap, S., Yang, Z., Cui, N., Qu, Z., Zhu, G., Liu, C., Giwa, L.R., Abdulkadir, L., and Jiang, C. 2000. Involvement of histidine residues in proton sensing of ROMK1 channel. *J Biol Chem* **275:** 7811-7817.

Chew, C.F., Vijayan, R., Chang, J., Zitzmann, N., and Biggin, P.C. 2009. Determination of pore-lining residues in the hepatitis C virus p7 protein. *Biophys J* **96:** L10-12.

Chizhmakov, I.V., Geraghty, F.M., Ogden, D.C., Hayhurst, A., Antoniou, M., and Hay, A.J. 1996. Selective proton permeability and pH regulation of the influenza virus M2 channel expressed in mouse erythroleukaemia cells. *J Physiol* **494 (Pt 2):** 329-336.

Choma, C., Gratkowski, H., Lear, J.D., and DeGrado, W.F. 2000. Asparagine-mediated self-association of a model transmembrane helix. *Nature Structural Biology* **7:** 161-166.

Ciampor, F., Bayley, P.M., Nermut, M.V., Hirst, E.M., Sugrue, R.J., and Hay, A.J. 1992. Evidence that the amantadine-induced, M2-mediated conversion of influenza A virus hemagglutinin to the low pH conformation occurs in an acidic trans Golgi compartment. *Virology* **188:** 14-24.

Clarke, D., Griffin, S., Beales, L., Gelais, C.S., Burgess, S., Harris, M., and Rowlands, D. 2006. Evidence for the formation of a heptameric ion channel complex by the hepatitis C virus p7 protein in vitro. *J Biol Chem* **281:** 37057-37068.

Clayton, D., Shapovalov, G., Maurer, J.A., Dougherty, D.A., Lester, H.A., and Kochendoerfer, G.G. 2004. Total chemical synthesis and electrophysiological characterization of mechanosensitive channels from Escherichia coli and Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **101:** 4764-4769.

Cleverley, R.M., Saleem, M., Kean, J., Ford, R.C., Derrick, J.P., and Prince, S.M. 2008. Selection of membrane protein targets for crystallization using PFO-PAGE electrophoresis. *Mol Membr Biol* **25:** 625-630.

Collins, P.L., and Mottet, G. 1993. Membrane orientation and oligomerization of the small hydrophobic protein of human respiratory syncytial virus. *Journal of General Virology* **74 (Pt 7):** 1445-1450.

Cook, G.A., and Opella, S.J. 2009. NMR studies of p7 protein from hepatitis C virus. *Eur Biophys J.*

Cristian, L., Lear, J.D., and DeGrado, W.F. 2003. Use of thiol-disulfide equilibria to measure the energetics of assembly of transmembrane helices in phospholipid bilayers. *Proc Natl Acad Sci U S A* **100:** 14772-14777.

Dawson, P.E., Muir, T.W., Clark-Lewis, I., and Kent, S.B. 1994. Synthesis of proteins by native chemical ligation. *Science* **266:** 776-779.

Doyle, D.A., Morais Cabral, J., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., and MacKinnon, R. 1998. The structure of the potassium channel: molecular basis of K+ conduction and selectivity. *Science* **280:** 69-77.

Ebie, A.Z., and Fleming, K.G. 2007. Dimerization of the erythropoietin receptor transmembrane domain in micelles. *J Mol Biol* **366:** 517-524.

Ewart, G.D., Sutherland, T., Gage, P.W., and Cox, G.B. 1996. The Vpu protein of human immunodeficiency virus type 1 forms cation-selective ion channels. *J Virol* **70:** 7108-7115.

Feldman, S.A., Hendry, R.M., and Beeler, J.A. 1999. Identification of a linear heparin binding domain for human respiratory syncytial virus attachment glycoprotein G. *J Virol* **73:** 6610-6617.

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

Fleming, K.G. 2002. Standardizing the free energy change of transmembrane helix-helix interactions. *Journal of Molecular Biology* **323:** 563-571.

Fleming, K.G. 2008. Determination of membrane protein molecular weight using sedimentation equilibrium analytical ultracentrifugation. *Curr Protoc Protein Sci* **Chapter 7:** Unit 7 12 11-17 12 13.

Fleming, K.G., and Engelman, D.M. 2001. Specificity in transmembrane helix-helix interactions can define a hierarchy of stability for sequence variants. *Proceedings of the National Academy of Sciences USA* **98:** 14340-14344.

Fleming, K.G., Ren, C.C., Doura, A.K., Eisley, M.E., Kobus, F.J., and Stanley, A.M. 2004. Thermodynamics of glycophorin A transmembrane helix dimerization in C14 betaine micelles. *Biophys Chem* **108:** 43-49.

Fox, J.D., Kapust, R.B., and Waugh, D.S. 2001. Single amino acid substitutions on the surface of Escherichia coli maltose-binding protein can have a profound impact on the solubility of fusion proteins. *Protein Sci* **10:** 622-630.

Franzin, C.M., Teriete, P., and Marassi, F.M. 2007. Structural similarity of a membrane protein in micelles and membranes. *J Am Chem Soc* **129:** 8078-8079.

Fuentes, S., Tran, K.C., Luthra, P., Teng, M.N., and He, B. 2007. Function of the respiratory syncytial virus small hydrophobic protein. *Journal of Virology* **81:** 8361-8366.

Gan, S.W., Ng, L., Lin, X., Gong, X., and Torres, J. 2008. Structure and ion channel activity of the human respiratory syncytial virus (hRSV) small hydrophobic protein transmembrane domain. *Protein Sci* **17:** 813-820.

Gardiner, A.T., Southall, J., and Cogdell, R.J. 2006. Crystallization of Membrane Proteins. In *Structural Genomics on Membrane Proteins*

(ed. K.H. Lundstrom). CRC Press, FL.

Gonzalez, M.E., and Carrasco, L. 2003. Viroporins. *FEBS Letters* **552:** 28-34.

Gratkowski, H., Dai, Q.H., Wand, A.J., DeGrado, W.F., and Lear, J.D. 2002. Cooperativity and specificity of association of a designed transmembrane peptide. *Biophys J* **83:** 1613-1619.

Griffin, S.D., Beales, L.P., Clarke, D.S., Worsfold, O., Evans, S.D., Jaeger, J., Harris, M.P., and 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:** 34-38.

Groothuis, J.R., and Nishida, H. 2002. Prevention of respiratory syncytial virus infections in high-risk infants by monoclonal antibody (palivizumab). *Pediatr Int* **44:** 235-241.

Heminway, B.R., Yu, Y., Tanaka, Y., Perrine, K.G., Gustafson, E., Bernstein, J.M., and Galinski, M.S. 1994. Analysis of respiratory syncytial virus F, G, and SH proteins in cell fusion. *Virology* **200:** 801-805.

Holsinger, L.J., and Lamb, R.A. 1991. Influenza virus M2 integral membrane protein is a homotetramer stabilized by formation of disulfide bonds. *Virology* **183:** 32-43.

Howard, K.P., Lear, J.D., and DeGrado, W.F. 2002. Sequence determinants of the energetics of folding of a transmembrane four-helix-bundle protein. *Proc Natl Acad Sci U S A* **99:** 8568-8572.

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

Hu, J., Asbury, T., Achuthan, S., Li, C., Bertram, R., Quine, J.R., Fu, R., and Cross, T.A. 2007a. Backbone structure of the amantadine-blocked trans-membrane domain M2 proton channel from Influenza A virus. *Biophys J* **92:** 4335-4343.

Hu, J., Qin, H., Li, C., Sharma, M., Cross, T.A., and Gao, F.P. 2007b. Structural biology of transmembrane domains: efficient production and characterization of transmembrane peptides by NMR. *Protein Sci* **16:** 2153-2165.

Jorgensen, W.L., and Tirado-Rives, J. 1988. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society* **110** 1657-1666.

Kang, C., Tian, C., Sonnichsen, F.D., Smith, J.A., Meiler, J., George, A.L., Jr., Vanoye, C.G., Kim, H.J., and Sanders, C.R. 2008. Structure of KCNE1 and implications for how it modulates the KCNQ1 potassium channel. *Biochemistry* **47:** 7999-8006.

Kapust, R.B., and Waugh, D.S. 1999. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8:** 1668-1674.

Kochendoerfer, G.G., Jones, D.H., Lee, S., Oblatt-Montal, M., Opella, S.J., and Montal, M. 2004. Functional characterization and NMR spectroscopy on full-length Vpu from HIV-1 prepared by total chemical synthesis. *J Am Chem Soc* **126:** 2439-2446.

Kochendoerfer, G.G., Salom, D., Lear, J.D., Wilk-Orescan, R., Kent, S.B., and DeGrado, W.F. 1999. Total chemical synthesis of the integral membrane protein influenza A virus M2: role of its C-terminal domain in tetramer assembly 1. *Biochemistry* **38:** 11905-11913.

Kochva, U., Leonov, H., and Arkin, I.T. 2003. Modeling the structure of the respiratory syncytial virus small hydrophobic protein by silent-mutation analysis of global searching molecular dynamics. *Protein Science* **12:** 2668-2674.

Kostolansky, F., Russ, G., Mucha, V., and Styk, B. 1988. Changes in the influenza virus haemagglutinin at acid pH detected by monoclonal antibodies to glycopolypeptides HA1 and HA2. *Arch Virol* **101:** 13-24.

Kovacs, F.A., and Cross, T.A. 1997. Transmembrane four-helix bundle of influenza A M2 protein channel: structural implications from helix tilt and orientation. *Biophysical Journal* **73:** 2511-2517.

Kovacs, F.A., Denny, J.K., Song, Z., Quine, J.R., and Cross, T.A. 2000. Helix tilt of the M2 transmembrane peptide from influenza A virus: an intrinsic property. *Journal of Molecular Biology* **295:** 117-125.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L.L. 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology* **305:** 567-580.

Krusat, T., and Streckert, H.J. 1997. Heparin-dependent attachment of respiratory syncytial virus (RSV) to host cells. *Arch Virol* **142:** 1247-1254.

Kukol, A., Adams, P.D., Rice, L.M., Brunger, A.T., and Arkin, T.I. 1999. Experimentally based orientational refinement of membrane protein models: A structure for the Influenza A M2 H+ channel. *Journal of Molecular Biology* **286:** 951-962.

Kukol, A., and Arkin, I.T. 1999. vpu transmembrane peptide structure obtained by site-specific fourier transform infrared dichroism and global molecular dynamics searching 55. *Biophysical Journal* **77:** 1594-1601.

Kukol, A., Torres, J., and Arkin, I.T. 2002. A structure for the trimeric MHC class II-associated invariant chain transmembrane domain. *Journal of Molecular Biology* **320:** 1109-1117.

Lama, J., and Carrasco, L. 1992. Expression of poliovirus nonstructural proteins in Escherichia coli cells. Modification of membrane permeability induced by 2B and 3A. *J Biol Chem* **267:** 15932-15937.

Lamb, R.A. 1993. Paramyxovirus fusion: a hypothesis for changes. *Virology* **197:** 1-11.

Lang, F., Foller, M., Lang, K.S., Lang, P.A., Ritter, M., Gulbins, E., Vereninov, A., and Huber, S.M. 2005. Ion channels in cell proliferation and apoptotic cell death. *Journal of Membrane Biology* **205:** 147-157.

Laue, T.M. 1995. Sedimentation equilibrium as thermodynamic tool. *Methods Enzymol* **259:** 427-452.

Lemmon, M.A., Flanagan, J.M., Treutlein, H.R., Zhang, J., and Engelman, D.M. 1992. Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry* **31:** 12719-12725.

Li, C., Qin, H., Gao, F.P., and Cross, T.A. 2007. Solid-state NMR characterization of conformational plasticity within the transmembrane domain of the influenza A M2 proton channel. *Biochim Biophys Acta* **1768:** 3162-3170.

Li, R.H., Gorelik, R., Nanda, V., Law, P.B., Lear, J.D., DeGrado, W.F., and Bennett, J.S. 2004. Dimerization of the transmembrane domain of integrin aIIb subunit in cell membranes. *Journal of Biological Chemistry* **279:** 26666-26673.

Lin, X., Tan, S.M., Law, S.K.A., and Torres, J. 2006a. Two types of transmembrane homomeric interactions in the integrin receptor family are evolutionarily conserved. *Proteins* **63:** 16-23.

Lin, X., Tan, S.M., Law, S.K.A., and Torres, J. 2006b. Unambiguous prediction of human integrin transmembrane heterodimer interactions using only homologous sequences. *Proteins: Structure, Function and Genetics* **65:** 274-279.

Liu, C.F., and Tam, J.P. 1994. Peptide segment ligation strategy without use of protecting groups. *Proc Natl Acad Sci U S A* **91:** 6584-6588.

Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X., and Chang, W. 2004. Crystal structure of spinach major light-harvesting complex at 2.72 A resolution. *Nature* **428:** 287-292.

MacKenzie, K.R., and Fleming, K.G. 2008. Association energetics of membrane spanning alpha-helices. *Curr Opin Struct Biol* **18:** 412-419.

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

Madan, V., Castello, A., and Carrasco, L. 2008. Viroporins from RNA viruses induce caspase-dependent apoptosis. *Cell Microbiol* **10:** 437-451.

Madan, V., Garcia Mde, J., Sanz, M.A., and Carrasco, L. 2005. Viroporin activity of murine hepatitis virus E protein. *FEBS Lett* **579:** 3607-3612.

Marley, J., Lu, M., and Bracken, C. 2001. A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* **20:** 71-75.

Marshall, J., Molloy, R., Moss, G.W., Howe, J.R., and Hughes, T.E. 1995. The jellyfish green fluorescent protein: a new tool for studying ion channel expression and function. *Neuron* **14:** 211-215.

Martin, K., and Helenius, A. 1991. Nuclear transport of influenza virus ribonucleoproteins: the viral matrix protein (M1) promotes export and inhibits import. *Cell* **67:** 117-130.

McCurdy, L.H., and Graham, B.S. 2003. Role of plasma membrane lipid microdomains in respiratory syncytial virus filament formation. *J Virol* **77:** 1747-1756.

Melero, J.A. 2007. Molecular Biology of Human Respiratory Syncytail Virus In *Respiratory Syncytail Virus*. (ed. P. Cane). Elsevier.

Melton, J.V., Ewart, G.D., Weir, R.C., Board, P.G., Lee, E., and Gage, P.W. 2002. Alphavirus 6K proteins form ion channels. *J Biol Chem* **277:** 46923-46931.

Mould, J.A., Paterson, R.G., Takeda, M., Ohigashi, Y., Venkataraman, P., Lamb, R.A., and Pinto, L.H. 2003. Influenza B virus BM2 protein has ion channel activity that conducts protons across membranes. *Dev Cell* **5:** 175-184.

Neumoin, A., Cohen, L.S., Arshava, B., Tantry, S., Becker, J.M., Zerbe, O., and Naider, F. 2009. Structure of a double transmembrane fragment of a G-protein-coupled receptor in micelles. *Biophys J* **96:** 3187-3196.

Newstead, S., Ferrandon, S., and Iwata, S. 2008. Rationalizing alpha-helical membrane protein crystallization. *Protein Sci* **17:** 466-472.

Neznanov, N., Kondratova, A., Chumakov, K.M., Angres, B., Zhumabayeva, B., Agol, V.I., and Gudkov, A.V. 2001. Poliovirus protein 3A inhibits tumor necrosis factor (TNF)-induced apoptosis by eliminating the TNF receptor from the cell surface. *J Virol* **75:** 10409-10420.

Okada, A., Miura, T., and Takeuchi, H. 2001. Protonation of histidine and histidine-tryptophan interaction in the activation of the M2 ion channel from influenza a virus. *Biochemistry* **40:** 6053-6060.

Olmsted, R.A., and Collins, P.L. 1989. The 1A protein of respiratory syncytial virus is an integral membrane protein present as multiple, structurally distinct species. *Journal of Virology* **63:** 2019-2029.

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

Page, R.C., Lee, S., Moore, J.D., Opella, S.J., and Cross, T.A. 2009. Backbone structure of a small helical integral membrane protein: A unique structural characterization. *Protein Sci* **18:** 134-146.

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

Parthasarathy, K., Ng, L., Lin, X., Liu, D.X., Pervushin, K., Gong, X., and Torres, J. 2008. Structural flexibility of the pentameric SARS coronavirus envelope protein ion channel. *Biophys J* **95:** L39-41.

Pastey, M.K., Crowe, J.E., Jr., and Graham, B.S. 1999. RhoA interacts with the fusion glycoprotein of respiratory syncytial virus and facilitates virus-induced syncytium formation. *J Virol* **73:** 7262-7270.

Pastey, M.K., Gower, T.L., Spearman, P.W., Crowe, J.E., Jr., and Graham, B.S. 2000. A RhoA-derived peptide inhibits syncytium formation induced by respiratory syncytial virus and parainfluenza virus type 3. *Nat Med* **6:** 35-40.

Perez, M., Garcia-Barreno, B., Melero, J.A., Carrasco, L., and Guinea, R. 1997. Membrane permeability changes induced in Escherichia coli by the SH protein of human respiratory syncytial virus. *Virology* **235:** 342-351.

Pervushin, K., Tan, E., Parthasarathy, K., Lin, X., Jiang, F.L., Yu, D., Vararattanavech, A., Soong, T.W., Liu, D.X., and Torres, J. 2009. Structure and inhibition of the SARS coronavirus envelope protein ion channel. *PLoS Pathog* **5:** e1000511.

Peterson, B.Z., DeMaria, C.D., Adelman, J.P., and Yue, D.T. 1999. Calmodulin is the Ca2+ sensor for Ca2+ -dependent inactivation of L-type calcium channels. *Neuron* **22:** 549-558.

Pinto, L.H., Holsinger, L.J., and Lamb, R.A. 1992. Influenza virus M2 protein has ion channel activity. *Cell* **69:** 517-528.

Plotkowski, M.L., Kim, S., Phillips, M.L., Partridge, A.W., Deber, C.M., and Bowie, J.U. 2007. Transmembrane domain of myelin protein zero can form dimers: possible implications for myelin construction. *Biochemistry* **46:** 12164-12173.

Plugge, B., Gazzarrini, S., Nelson, M., Cerana, R., Van Etten, J.L., Derst, C., DiFrancesco, D., Moroni, A., and Thiel, G. 2000. A potassium channel protein encoded by chlorella virus PBCV-1. *Science* **287:** 1641-1644.

Ramjeesingh, M., Huan, L.J., Garami, E., and Bear, C.E. 1999. Novel method for evaluation of the oligomeric structure of membrane proteins. *Biochemical Journal* **342 (Pt 1):** 119-123.

Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G., and Deber, C.M. 2009. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc Natl Acad Sci U S A* **106:** 1760-1765.

Reynolds, J.A., and Tanford, C. 1976. Determination of molecular weight of the protein moiety in protein-detergent complexes without direct knowledge of detergent binding. *Proc Natl Acad Sci U S A* **73:** 4467-4470.

Rixon, H.W., Brown, G., Aitken, J., McDonald, T., Graham, S., and Sugrue, R.J. 2004. The small hydrophobic (SH) protein accumulates within lipid-raft structures of the Golgi complex during respiratory syncytial virus infection. *Journal of General Virology* **85:** 1153-1165.

Rixon, H.W., Brown, G., Murray, J.T., and Sugrue, R.J. 2005. The respiratory syncytial virus small hydrophobic protein is phosphorylated via a mitogen-activated protein kinase p38-dependent tyrosine kinase activity during virus infection. *Journal of General Virology* **86:** 375-384.

Rosenthal, P.B., Zhang, X., Formanowski, F., Fitz, W., Wong, C.H., Meier-Ewert, H., Skehel, J.J., and Wiley, D.C. 1998. Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. *Nature* **396:** 92-96.

Ruiz-Arguello, M.B., Martin, D., Wharton, S.A., Calder, L.J., Martin, S.R., Cano, O., Calero, M., Garcia-Barreno, B., Skehel, J.J., and Melero, J.A. 2004. Thermostability of the

human respiratory syncytial virus fusion protein before and after activation: implications for the membrane-fusion mechanism. *J Gen Virol* **85:** 3677-3687.

Russ, W.P., and Engelman, D.M. 1999. TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 863-868.

Russ, W.P., and Engelman, D.M. 2000. The GxxxG motif: A framework for transmembrane helix-helix association. *Journal of Molecular Biology* **296:** 911-919.

Sal-Man, N., Gerber, D., Bloch, I., and Shai, Y. 2007. Specificity in transmembrane helix-helix interactions mediated by aromatic residues. *J Biol Chem* **282:** 19753-19761.

Sal-Man, N., Gerber, D., and Shai, Y. 2005. The identification of a minimal dimerization motif QXXS that enables homo- and hetero-association of transmembrane helices in vivo. *J Biol Chem* **280:** 27449-27457.

Sanders, C.R., and Sonnichsen, F. 2006. Solution NMR of membrane proteins: practice and challenges. *Magn Reson Chem* **44 Spec No:** S24-40.

Sandoz, G., Douguet, D., Chatelain, F., Lazdunski, M., and Lesage, F. 2009. Extracellular acidification exerts opposite actions on TREK1 and TREK2 potassium channels via a single conserved histidine residue. *Proc Natl Acad Sci U S A* **106:** 14628-14633.

Sanz, M.A., Madan, V., Carrasco, L., and Nieva, J.L. 2003. Interfacial domains in Sindbis virus 6K protein. Detection and functional characterization. *J Biol Chem* **278:** 2051-2057.

Sanz, M.A., Perez, L., and Carrasco, L. 1994. Semliki Forest virus 6K protein modifies membrane permeability after inducible expression in Escherichia coli cells. *J Biol Chem* **269:** 12106-12110.

Sato, T., Kawakami, T., Akaji, K., Konishi, H., Mochizuki, K., Fujiwara, T., Akutsu, H., and Aimoto, S. 2002. Synthesis of a membrane protein with two transmembrane regions. *J Pept Sci* **8:** 172-180.

Schnell, J.R., and Chou, J.J. 2008. Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* **451:** 591-595.

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

Sharpe, S., Yau, W.M., and Tycko, R. 2006. Structure and dynamics of the HIV-1 Vpu transmembrane domain revealed by solid-state NMR with magic-angle spinning. *Biochemistry* **45:** 918-933.

Shim, J.W., Yang, M., and Gu, L.Q. 2007. In vitro synthesis, tetramerization and single channel characterization of virus-encoded potassium channel Kcv. *FEBS Lett* **581:** 1027-1034.

Simmerman, H.K., Kobayashi, Y.M., Autry, J.M., and Jones, L.R. 1996. A leucine zipper stabilizes the pentameric membrane domain of phospholamban and forms a coiled-coil pore structure. *Journal of Biological Chemistry* **271:** 5941-5946.

Stanley, A.M., Chuawong, P., Hendrickson, T.L., and Fleming, K.G. 2006. Energetics of outer membrane phospholipase A (OMPLA) dimerization. *J Mol Biol* **358:** 120-131.

Stanley, A.M., and Fleming, K.G. 2005. The transmembrane domains of ErbB receptors do not dimerize strongly in micelles. *J Mol Biol* **347:** 759-772.

Stouffer, A.L., Acharya, R., Salom, D., Levine, A.S., Di Costanzo, L., Soto, C.S., Tereshko, V., Nanda, V., Stayrook, S., and DeGrado, W.F. 2008. Structural basis for the function and inhibition of an influenza virus proton channel. *Nature* **451:** 596-599.

Stouffer, A.L., DeGrado W.F., Lear J.D. 2006. Analytical Ultracentrifugation Studies of the Influenza M2 Humotetramerization Equilibrium in Detergent Solutions. *Progr Colloid Polym Sci* **131:** 108-115.

Stouffer, A.L., Nanda, V., Lear, J.D., and DeGrado, W.F. 2005. Sequence determinants of a transmembrane proton channel: an inverse relationship between stability and function. *J Mol Biol* **347:** 169-179.

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

Sunstrom, N.A., Premkumar, L.S., Premkumar, A., Ewart, G., Cox, G.B., and Gage, P.W. 1996. Ion channels formed by NB, an influenza B virus protein. *J Membr Biol* **150:** 127-132.

Szabo, I., Adams, C., and Gulbins, E. 2004. Ion channels and membrane rafts in apoptosis. *Pflugers Archiv European Journal of Physiology* **448:** 304-312.

Takeuchi, K., Takahashi, H., Kawano, S., and Shimada, I. 2007. Identification and characterization of the slowly exchanging pH-dependent conformational rearrangement in KcsA. *J Biol Chem* **282:** 15179-15186.

Tamm, L.K., and Tatulian, S.A. 1997. Infrared spectroscopy of proteins and peptides in lipid bilayers. *Quarterly Reviews of Biophysics* **30:** 365-429.

Techaarpornkul, S., Barretto, N., and Peeples, M.E. 2001. Functional analysis of recombinant respiratory syncytial virus deletion mutants lacking the small hydrophobic and/or attachment glycoprotein gene. *Journal of Virology* **75:** 6825-6834.

Teriete, P., Franzin, C.M., Choi, J., and Marassi, F.M. 2007. Structure of the Na,K-ATPase regulatory protein FXYD1 in micelles. *Biochemistry* **46:** 6774-6783.

Torres, J., Adams, P.D., and Arkin, I.T. 2000a. Use of a new label, $^{13}C{=}^{18}O$, in the determination of a structural model of phospholamban in a lipid bilayer. Spatial restraints resolve the ambiguity arising from interpretations of mutagenesis data. *Journal of Molecular Biology* **300:** 677-685.

Torres, J., and Arkin, I.T. 2000. Recursive use of evolutionary conservation data in molecular modeling of membrane proteins A model of the multidrug H+ antiporter emrE. *European Journal of Biochemistry* **267:** 3422-3431.

Torres, J., and Arkin, I.T. 2002. C-deuterated alanine: a new label to study membrane protein structure using site-specific infrared dichroism. *Biophysical Journal* **82:** 1068-1075.

Torres, J., Briggs, J.A., and Arkin, I.T. 2002a. Convergence of experimental, computational and evolutionary approaches predicts the presence of a tetrameric form for CD3-zeta. *Journal of Molecular Biology* **316:** 375-384.

Torres, J., Briggs, J.A., and Arkin, I.T. 2002b. Multiple site-specific infrared dichroism of CD3-z, a transmembrane helix bundle. *Journal of Molecular Biology* **316:** 365-374.

Torres, J., Briggs, J.A.G., and Arkin, I.T. 2002c. Contribution of energy values to the analysis of global searching molecular dynamics simulations of transmembrane helical bundles. *Biophysical Journal* **82:** 3063-3071.

Torres, J., Kukol, A., and Arkin, I.T. 2000b. Use of a single glycine residue to determine the tilt and orientation of a transmembrane helix. A new structural label for infrared spectroscopy. *Biophysical Journal* **79:** 3139-3143.

Torres, J., Kukol, A., and Arkin, I.T. 2001. Mapping the energy surface of transmembrane helix-helix interactions. *Biophysical Journal* **81:** 2681-2692.

Torres, J., Parthasarathy, K., Lin, X., Saravanan, R., Kukol, A., and Liu, D.X. 2006. Model of a putative pore: the pentameric a-helical bundle of SARS coronavirus E protein in lipid bilayers. *Biophysical Journal* **91:** 938-947.

Torres, J., Wang, J., Parthasarathy, K., and Liu, D.X. 2005. The transmembrane oligomers of coronavirus protein E. *Biophysical Journal* **88:** 1283-1290.

Tsuboi, M. 1962. Infrared dichroism and molecular conformation of à-form poly-g-benzyl-L-glutamate. *Journal of Polymer Science* **59:** 139-153.

Van Horn, W.D., Kim, H.J., Ellis, C.D., Hadziselimovic, A., Sulistijo, E.S., Karra, M.D., Tian, C., Sonnichsen, F.D., and Sanders, C.R. 2009. Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. *Science* **324:** 1726-1729.

van Kuppeveld, F.J., Hoenderop, J.G., Smeets, R.L., Willems, P.H., Dijkman, H.B., Galama, J.M., and Melchers, W.J. 1997. Coxsackievirus protein 2B modifies endoplasmic reticulum membrane and plasma membrane permeability and facilitates virus release. *EMBO J* **16:** 3519-3532.

von Heijne, G. 1996. Principles of membrane protein assembly and structure. *Prog Biophys Mol Biol* **66:** 113-139.

Wang, C., Lamb, R.A., and Pinto, L.H. 1994. Direct measurement of the influenza A virus M2 protein ion channel activity in mammalian cells. *Virology* **205:** 133-140.

Wang, J., Kim, S., Kovacs, F., and Cross, T.A. 2001. Structure of the transmembrane region of the M2 protein H(+) channel. *Protein Science* **10:** 2241-2250.

Wellings, D.A., and Atherton, E. 1997. Standard Fmoc protocols. *Methods in Enzymology* **289:** 44-67.

Werling, D., Hope, J.C., Chaplin, P., Collins, R.A., Taylor, G., and Howard, C.J. 1999. Involvement of caveolae in the uptake of respiratory syncytial virus antigen by dendritic cells. *J Leukoc Biol* **66:** 50-58.

Whitehead, S.S., Bukreyev, A., Teng, M.N., Firestone, C.Y., St Claire, M., Elkins, W.R., Collins, P.L., and Murphy, B.R. 1999. Recombinant respiratory syncytial virus bearing a deletion of either the NS2 or SH gene is attenuated in chimpanzees. *Journal of Virology* **73:** 3438-3442.

Yu, L., Sun, C., Song, D., Shen, J., Xu, N., Gunasekera, A., Hajduk, P.J., and Olejniczak, E.T. 2005. Nuclear magnetic resonance structural studies of a potassium channel-charybdotoxin complex. *Biochemistry* **44:** 15834-15841.

Yuill, K.H., Stansfeld, P.J., Ashmole, I., Sutcliffe, M.J., and Stanfield, P.R. 2007. The selectivity, voltage-dependence and acid sensitivity of the tandem pore potassium channel TASK-1: contributions of the pore domains. *Pflugers Arch* **455:** 333-348.

Zhou, F.X., Merianos, H.J., Brunger, A.T., and Engelman, D.M. 2001. Polar residues drive association of polyleucine transmembrane helices. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 2250-2255.

PUBLICATIONS AND CONFERENCES

blications:

1. Structure and ion channel activity of the human respiratory syncytial virus (hrsv) small hydrophobic protein transmembrane domain. SW Gan, LF Ng, X Lin, XD Gong and J Torres. Protein Science 2008, 17: 813-820.

2. The transmembrane homotrimer of ADAM 1 in model lipid bilayers. SW Gan, X Lin, and J Torres. Protein Science 2007, 16: 285-292.

al presentation in major conferences:

le: Model of pentameric ion channel formed by transmembrane domain of SH protein of hRSV.

nference:

1) 1st International Conference on Viral Membrane Proteins, 2007 December, 6th-8th. National Yang-Ming University, Taipei, Taiwan.

2) Joint 5th Structural Biology & Functional Genomics and 1st Biological Physics International Conference, 2008 December, 9th-11th. National University of Singapore, Singapore.