Structure Determination of a Double Transmembrane Fragment of the G protein-coupled Receptor Ste2p in Membrane Mimetic Environments

By

Leah S. Cohen

A dissertation submitted to the Graduate Faculty in Biochemistry in partial fulfillment of the requirements for the degree of Doctor in Philosophy, The City University of New York.

2010

© 2010

Leah S. Cohen

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Biochemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date	Dr. Fred Naider Chair of Examining Committee
Date	Dr. Edward Kennelly Executive Officer

Dr. Ruth E. Stark

Dr. Lesley Davenport

Dr. Jeffery M. Becker

Dr. Sebastien Poget

Supervisory Committee

The City University of New York

Abstract

STRUCTURE DETERMINATION OF A DOUBLE TRANSMEMBRANE FRAGMENT OF THE G PROTEIN-COUPLED RECEPTOR STE2P IN MEMBRANE MIMETIC ENVIRONMENTS

By

Leah S. Cohen

Mentor: Professor Fred Naider

G-protein coupled receptors (GPCRs) are relevant in cellular signal transduction pathways and are targets for disease therapeutics. Since the sequencing of the human genome, there have been close to 1000 GPCRs predicted and many have been characterized by biological and biochemical analysis. Though these integral membrane proteins (IMPs) have little sequence similarity, they show strong putative structural similarities. All GPCRs contain an N-terminal extracellular domain (NT), 7 transmembrane helical regions (TM) connected by intra- and extracellular loops (IL and EL, respectively), and a C-terminal intracellular tail (CT). The extracellular domains are thought to play a role in ligand-receptor interactions and together with the TM domains form the ligand binding site for many GPCRs.

The number of structures in the protein structural database is increasing exponentially, but there are only a few high-resolution GPCR structures, those of rhodopsin, the β-adrenergic receptors and the adenosine A_{2A} receptor. GPCRs are difficult to crystallize and the protein:detergent complex size can be restrictive for solution NMR. As is true of GPCRs in general, limited structural information is available for Ste2p, a yeast GPCR which recognizes the α-factor tridecapeptide mating pheromone. Many groups, including the Naider lab, have been working with smaller fragments of GPCRs in an attempt to elucidate the structure of individual

regions by either crystallization or NMR and to combine these to achieve a better picture of GPCR structure. The use of fragments of GPCRs to study the structure of these large membrane receptors remains controversial. I have used Ste2p as a paradigm to answer the following questions: 1) Can GPCR fragments fold into a tertiary structure without the context of the full protein? and 2) Do organic:aqueous solvents such as trifluoroethanol(TFE):water result in GPCR fragment structures similar to those found in detergents and lipids?

I have cloned and expressed two 2TM regions of Ste2p, G31-T110 (TM1-TM2) and R231-S339 (TM6-TM7-CT40). The constructs were chosen based on the biological relevance of the domain and a buried surface area analysis that was performed using a rhodopsin-templated model of Ste2p and the NACCESS software package. Expression in *Escherichia coli* in minimal media and CNBr cleavage conditions for these 2TM fusion proteins were optimized and the fusion protein was labeled with [^{15}N], [^{15}N,^{13}C], [^{15}N,^{13}C,^{2}H] and with selectively [^{15}N]-labeled amino acids to aid in NMR assignments. Purification of isotopically labeled peptides was performed using RP-HPLC with a Zorbax C3 column and an acetonitrile:water gradient containing isopropanol and 0.1% TFA. Yields of Ste2p(G31-T110) were 6-20 mg per L culture. Far UV CD analysis indicated that Ste2p(G31-T110) formed a highly helical peptide in TFE:water and various micellar environments, whereas the helicity was reduced for Ste2p(R231-S339). Initial [^{15}N,^{1}H]-HSQC analysis showed that CD can be used as an NMR screening technique to determine conditions for high resolution NMR analysis.

Ste2p(G31-T110) was used to determine a high-resolution structure of a 2TM GPCR fragment in TFE:water(+0.1% TFA) (1:1; v/v) by multidimensional NMR spectroscopy. The 80-residue polypeptide was folded into a helical hairpin, but the overall convergence of the structure bundle was poor with an RMSD of ~ 6.7 Å. Comparison to an NMR study of TM1-

TM2 in LPPG micelles concluded that very different structures occurred in the two membrane mimetic media and it was observed that relative conformation of TM2 to TM1 in TFE:water was twisted in comparison to that in the LPPG micelle so that the interactions between helices were on opposite sides. Although Ste2p(G31-T110) folds into a tertiary structure in TFE:water (1:1; v/v) the structure is quite flexible compared to that in LPPG micelles.

This dissertation is dedicated to my husband Ari for his love and support and to my children, Rivkah, Aaron, Miriam and P'nina, for their understanding and smiles.

Acknowledgements

There are many people that played a role in convincing me to go back to school and help me complete my doctoral dissertation. I would like to thank Dr. Naider for his extremely important role in my scientific career. He has trusted me to perform to the best of my ability and pushed me to achieve higher goals. His suggestion to join the Doctoral Program at the Graduate Center, City University of New York began a career changing journey in which I learned new scientific techniques and the art of scientific writing. The journey culminated in the attendance of the 60th Annual Meeting of Nobel Laureates in Lindau, Germany which was the opportunity of a lifetime.

I would also like to thank my committee members: Dr. Becker, Dr. Stark, Dr. Davenport and Dr. Poget. The long time collaboration between Dr. Becker and Dr. Naider is a unique venture into the art of structure:function relationships and I am glad to be a part of it. The help of Dr. Stark with my training in NMR and our scientific discussions have been very helpful over the years. Dr. Davenport's guidance throughout my graduate career in the Graduate Center Biochemistry PhD program was invaluable. Finally, the discussions with Dr. Poget about recent NMR papers and the difficulties in my project have been beneficial in increasing my knowledge in membrane protein structure analysis.

My past has played a strong role in where I am today and I would like to acknowledge the guidance of support of my previous mentors Dr. Robert Osuna and Dr. Richard E. Davis.

A laboratory is not an environment in which one can survive alone. The members of the Naider Lab, past and present, have affected me and my science. Boris has guided me through peptide synthesis, HPLC purification, MS, CD and NMR. He has always taken the time to help me when I have needed it and I am grateful for his support. Past lab members Racha Estephan

and Jacqueline Englander brought me into their membrane peptide biosynthesis world and got me started on this challenging path. Katrina Caroccia, my fellow graduate student in the trenches, has been there with me since I started. She has supported me in all that I do both scientifically and personally and I value her friendship as one of the greatest rewards of this program.

None of this would have been possible without the support of my family. My parents and my in-laws have helped in countless ways. From babysitting for weeks at a time to preparing Shabbat meals, their support has been unconditional. My sister Rachel and my sister-in-law Naomi have not hesistated to pick up my kids and watch them when an experiment has gone long or I have had to attend a conference and I could not have done this without them. I am forever grateful to my children, Rivkah, Aaron, Miriam and P'nina, who have made this the most interesting graduate career that I could have imagined. Though their illnesses always seemed to fall in the middle of the greatest experiment, their hugs and kisses were all I needed to go to work in the morning and push my way through to completion. Finally, to my wonderful and supportive husband, Ari: we have been through graduate school once before and now that I have finally finished I can say that we have survived it all. You are my friend, my partner and my one and only and I cannot wait to move forward on this journey with you and our family. Your love was the rock that kept me gounded and the sanity when mine was lost.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ºC	degree Celsius
1D	One-dimensional
2D	two-dimensional
2TM	double TM
3D	three-dimensional
4D	four-dimensional
*Bam*HI	*Bacillus amyloliquefaciens* H restriction enzyme I
BSA	buried surface area
CD	circular dichroism
COSY	correlation spectroscopy
ct	constant time
CT	Carboxy-terminal tail
D_2O	deuterium oxide
DAGK	diacyl glycerol kinase
DDM	dodecyl maltoside
DE	directly expressed
DHPC	1,2-dihexanoyl-*sn*-glycero-3-phosphocholine
DPC	dodecylphosphocholine
*Dpn*I	*Diplococcus pneumoniae* restriction enzyme I
DMSO	dimethyl sulfoxide
DNA	deoxynucleic acid
dNTP	deoxynucleotide triphosphate

DSS	2,2-dimethyl-2-silapentane-5- sulfonate sodium salt
DYANA	DYnamics Algorithm for Nmr Applications
E. coli	*Escherichia coli*
EL	extracellular loop
ESI-MS	electrospray-ionization mass spectroscopy
f_H	fraction helicity
*Fok*I	*Flavobacterium okeanokoites* restriction enzyme I
FP	fusion protein
FRET	Förster Resonance Energy Transfer
GDP	guanosine diphosphate
GnHCl	guanidine hydrochloride
GPCRs	G protein-coupled receptors
GTP	guanosine triphospahte
H-D	hydrogen-deuterium
HDPC	hexadecylphosphocholine
*Hind*III	*Haemophilus influenzae* Rd restriction enzyme III
His_6	six histidines in a row used for immunopurification and/or identification
HMQC	heternuclear multiple quantum coherence
H-NOE	$^{15}N(^{1}H)$-NOE experiment
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum correlataion
IBs	inclusion bodies

IL	intracellular loop
IMPs	integral membrane proteins
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	kiloDalton
LB	Luria-Bertani media
$LBAmp_{200}$	Luria-Bertani media containing 200 μg/mL ampicillin
$LBAmp_{50}Cam_{34}$	Luria-Bertani media containing 50 μg/mL ampicillin and 34 μg/mL chloramphenicol
LM	lysis mix
LMPG	1-myristoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol)
LPPC	1-palmitoyl-2-hydroxy-*sn*-glycero-3-phosphocholine
LPPG	1-palmitoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol)
LSPG	1-stearoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol)
MAP	mitogen activated protein
MAPK	mitogen activated protein kinase
MAT**a**	mating type a
MATα	mating type α
MD	mild detergent
min	minute
MM	M9 minimal media
MPex	Membrane Protein Explorer
msec	milliseconds
MS	mass spectroscopy

*Msl*I	*Moraxella osloensis* restriction enzyme I
MTSL	1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) Methanethio-sulfonate
MW	molecular weight
*Nde*I	*Neisseria denitrificans* restriction enzyme I
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser effect spectroscopy
NOEs	Nuclear Overhauser effects
NT	N-terminal tail
OD	optical density
PCR	polymerase chain reaction
Pfx pol	Platinum *Pyrococcus sp.* polymerase
PHD	Protein network from HeiDelberg
ppm	parts per million
RDC	residual dipolar coupling
RMSD	root mean square deviation
RP-HPLC	reverse phase- high performance liquid chromatography
RT	room temperature or
S. cerevisiae	*Sacchromyces cerevisiae*
SA	Surface area
SDS	sodium dodecylsulfate
SDS-PAGE	Sodium dodecyl sulfate -polyacrylamide gel electrophoresis
SOC	Super Optimal broth with Catabolite repression

TALOS	Torsion Angle Likelihood Obtained from Shift and sequence similarity
TBST	tris-buffered saline + 0.5% Tween-20
TFA	trifluoroacetic acid
TFE	trifluoroethanol
T_m	melting temperature
TM	transmembrane
TM1-TM2	multidomain fragment containing part of the N-terminal tail, the first transmembrane domain, the first intracellular loop, the second transmembrane domain and part of the first intracellular loop of Ste2p (G31-T110)
TM6-TM7-CT40	multidomain fragment containing the sixth transmembrane domain, the third extracellular loop, the seventh transmembrane domain and 40 residues of the cytosolic tail of Ste2p (R231-S339)
TOCSY	total correlation spectroscopy
TrpΔLE	portion of the histidine-tagged TrpΔLE1413
TROSY	transverse relaxation optimized spectroscopy
UV	ultraviolet
WT	wild-type

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Background and Significance

G Protein-Coupled Receptors

G protein-coupled receptors (GPCRs) are integral membrane proteins (IMPs) found in the plasma membrane of cells. They play a role in cell signaling and signal transduction initiated by protein-ligand interactions. In general, a ligand interacts with the extracellular and/or transmembrane region of the GPCR and causes a conformational change in the protein that is propagated to a heterotrimeric G protein that is interacting with the intracellular face of the GPCR. A change in the conformation of the G protein allows for GDP to be exchanged for GTP and the trimeric G protein complex is disrupted. In yeast, the Gα monomer containing GTP releases the Gβγ dimer which carries the downstream message and a MAP kinase cascade ensues (*1-3*). Following the release of the activated G proteins, the C-terminal tail (CT) may play a role in desensitzation of the signal by endocytosis.

GPCR ligands are highly diverse and include peptides, nucleotides, lipids, glycoproteins and light. They play a role in taste, smell, and light stimulation as well as many other cell signaling processes (*4*). Since the human genome has been sequenced, many potential GPCR genes have been predicted. These have been classified into five families known as the GRAFS system: Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2 and Secretin (*5*). GPCRs can also be classified by their predicted structural characteristics into five subfamilies: Family A, Rhodopsin and β-adrenergic related receptors; Family B, glucagon related receptors; Family C, metabotropic neurotransmitter related receptors; Families D and E, STE2 and STE3 yeast pheromone receptors, respectively; and Family F, four receptors in *Dictyostelium discoideum*

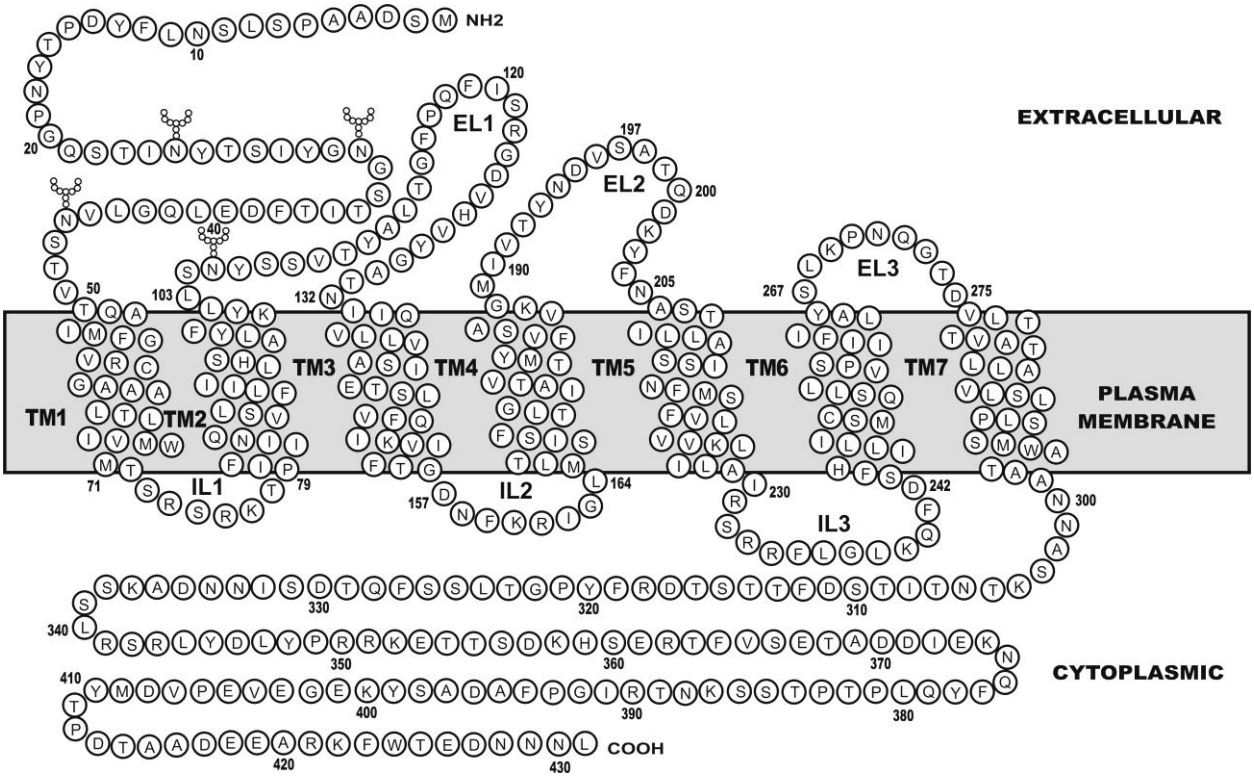


Figure 1-1. Cartoon of the yeast α-mating factor G protein-coupled receptor, Ste2p. The N-terminal domain contains four Asn glycosylation points. The domains are indicated as: TM, transmembrane; EL: extracellular loop; IL: intracellular loop. This figure was originally published in (*6*).

that are related to the production of cyclic adenosine monophosphate (*4*). These families are comprised of proteins with low sequence homology, but are based on a few conserved residues and structural similarities.

The GPCR families contain consensus sequences that may lead to insights into how the IMPs function. For example, the class A GPCRs, including rhodopsin and the β-adrenergic receptors, contain (E/D)RY sequences on TM3 and TM6, NPxxY on TM7, and disulfide bonds (*4, 7, 8*). These consensus sequences are not necessarily found between families, but structural and functional similarities have been observed. Each GPCR structure is predicted by hydropathy analysis to consist of 7 transmembrane (TM) helices that are connected by intra- and extra-

cellular loops (IL and EL) (Figure 1-1). Comparisons of GPCRs in two different classes, β-adrenergic receptor and Ste2p, revealed functional similarities in that they both interact with their ligand in a similar manner (*9-11*) and showed that similar regions are involved in activation (*12-14*) and desensitization (*15, 16*). Therefore, although Ste2p is in a GPCR class distinct from those of many medically relevant GPCRs, it is still a relevant model system for the generation of structural information about GPCRs.

Ste2p as a Model System

Cell Signaling. Yeast genetics is a very powerful tool because the mechanisms that are derived in yeast tend to be applicable to higher eukaryotes (*3*). It is relatively easy to manipulate the genome of a yeast cell to perform molecular biology experiments in yeast and cell signaling in yeast is very well understood. The yeast *Saccharomyces cerevisiae* can exist as a haploid or a diploid cell. When the two haploid yeast cells are mixed in a culture, diffusible peptide pheromones **a**-factor and α-factor, respectively, are released. The MAT**a** and MATα haploids express a GPCR, Ste2p or Ste3p, that specifically recognize α-factor and **a**-factor, respectively. Interaction of the pheromones with their cognate receptors produces conformational changes that are believed to be propagated through to the heterotrimeric G-protein complex via the third intracellular loop and portions of the C-terminus of the receptor. This change in conformation of the G proteins begins the MAP kinase cascade (*1-3*)(Figure 1-2). Control of this pathway helps to regulate the cell cycle, transcription of genes and cell polarity (*3*).

Mutations in the extracellular ends of the transmembrane domains appeared to be important in the early initiation steps of signal transduction, potentially affecting ligand binding (*17*). Two mutant receptors, F204S and Y266C, showed a decreased affinity for α-factor

binding demonstrating a role for these regions, EL2-TM5 and TM6-EL3 interfaces respectively, in specific Ste2p- α-factor contacts. The importance of Y266 was also observed by site-directed mutagenesis studies of the Ste2p region 262-270, where Y266A had no response to α-factor and a decreased binding affinity (*18*). The use of photoactivatable or chemically activatable moieties on the α-factor peptide or Ste2p, such as 4-benzyl-L-phenylalanine or 3,4-dihydroxylphenylalanine, have more specifically pinpointed the regions of interaction. The amino-terminus of the peptide has been shown to crosslink the protein to the extracellular regions of TM5-EL2 and TM6-EL3-TM7 (*9, 19, 20*)(personal communication by George Umanah and Jeff Becker). The carboxy-terminus of the peptide, specifically Tyr13, has been shown to crosslink to a small region within TM1, F55-C59 (*19-21*). The importance of these regions in ligand binding and cell-signaling may be further elucidated by high-resolution structural information. Structural studies of this interaction, in conjunction with the biochemical studies, would be beneficial in determining on an atomic level how a GPCR interacts with its peptide ligand and may be applicable to other GPCR-peptide ligand interactions.

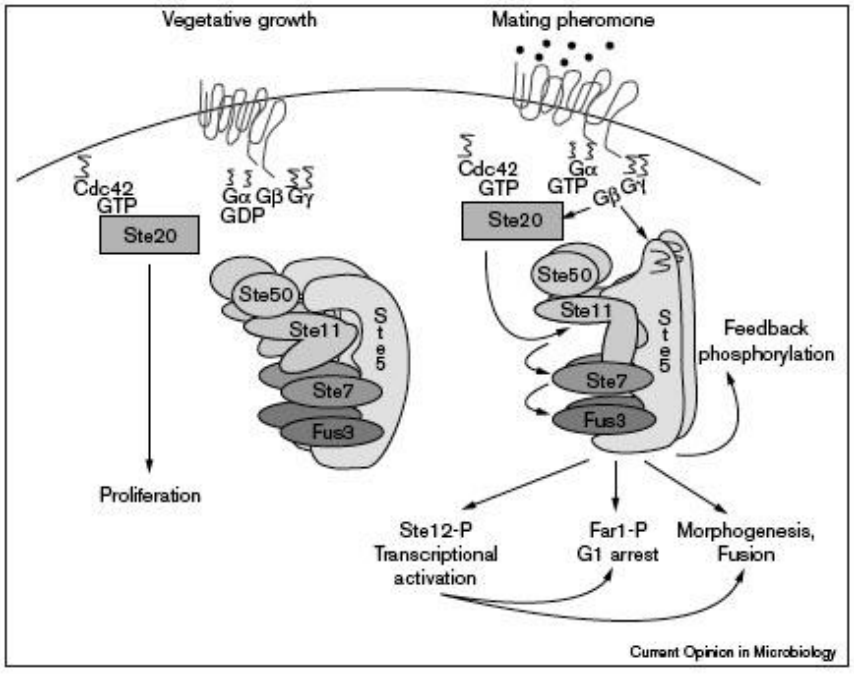


Figure 1-2. Signal transduction pathway propagated upon pheromone binding to Ste2p (*2*). A conformational change in the receptor leads to GDP to GTP exchange in Gα and release of Gβγ which activates Ste20. The signal is propagated leading to transcriptional activation, growth arrest and fusion. Figure taken from (*2*).

Yeast Functional Assays. Genetic manipulation in yeast has been used in research for a long time and multiple assays have been utilized in the analysis of yeast signal transduction. Phenotypic responses to the α-factor pheromone peptide can be observed microscopically by looking for pear-like "shmoo" formation (*22*, *23*), cell cycle arrest using the halo plate assay (*24*, *25*), and transcriptional regulation by the mating cycle pheromones using *LacZ* fusions to genes that are regulated by the MAPK cascades (*26*). The substituted cysteine accessibilty method has been utilized to determine which residues in the EL1 are exposed and/or undergo a conformational change upon pheromone binding (*27*). Direct receptor binding assays have been used with radiolabeled or fluorescent peptides (*28*-*33*). Thus the yeast system can be used to evaluate the biological effects of any changes in the amino acid sequence of Ste2p that may be necessary to better determine structural constraints.

Structural Studies of Integral Membrane Proteins and the Need for Fragment Analysis

X-ray crystallography and nuclear magnetic resonance (NMR) have been used to study the high-resolution structures of macromolecules. The worldwide Protein Data Bank (wwPDB; www.wwpdb.org) has determined that the total number of proteins deposited in the PDB is more than 60,000. These protein structures are mostly determined by Xray crystallography and are mainly soluble proteins. The number of unique IMP structures deposited for the year 2005 was less than 20, or less than 1% of all the structures in the Data Bank (*34*). The underrepresentation of IMP structures in the database belies the importance of these proteins. Many membrane proteins are used in cell-to-cell signaling, signal transduction, nutrient uptake and are related to disease. In fact, 30-60% of the current drugs on the market are targeted to membrane proteins

(*35-37*). The need for structural information on the target proteins is important for the development of more specific drugs that could increase the efficacy and decrease the side effects.

The structural analysis of membrane proteins has been limited by low protein expression levels, difficulty in crystal formation for X-ray crystallography, and size of the IMP:lipid complex for solution NMR. Expression of eukaryotic membrane proteins in bacterial systems can be toxic to the cells leading to low protein yields (reviewed in *38*). Many strains of *Escherichia coli* (*E. coli*) and plasmids with fusion tags have been developed to deal with the high toxicity levels (*39, 40*). Other methods have been introduced utilizing expression in yeast (*41-43*), insect cells (*44, 45*), mammalian cell culture (*46, 47*) and cell-free expression systems (*48, 49*) which appear to increase protein yields. Isotopic labeling of proteins, for use in NMR, with ^{15}N or ^{13}C requires growth in minimal media which generally leads to a decrease in protein expression levels but through expression optimization, the yields can be increased dramatically.

The inherent flexibility of GPCRs, one of the largest classes of IMPS, has led to difficulties in generating crystals for use in X-ray crystallography (*50*). Crystal formation is also hampered by the lipids needed to place the protein in an environment that mimics the functional conformation (*50*). Until recently, the only crystal structure of a GPCR was that of Rhodopsin which was available in mg quantities and contains a covalently bound retinal ligand thereby rigidifying the structure (*51-54*). In the last three years, high-resolution crystal structures of the β_2-adrenergic (*55*) and β_1-adrenergic receptors (*56*), ligand free opsin with (*57*) and without (*58*) a portion of the Gα monomer, and the adenosine A_{2A} receptor (*59*) were published using antibodies to stabilize structure (*60*), lyzozyme fusion to increase solubility and aid in the lattice contact formation (*55, 59, 61*), and thermostable mutations of a receptor in the antagonist state (*56*). Although these great accomplishments indicate that much structural information will be

forthcoming from analysis of GPCR crystals, most of the crystallized GPCRs are decidedly non-native, containing major insertions or significant mutation. Furthermore, X-ray crystallography cannot give insights into protein dynamics without in depth comparative analysis of all of the GPCR structures that have been determined to date (*62-65*).

A tool that is useful for structural as well as dynamic analysis is solution NMR. Interactions between proteins and ligands or proteins and other proteins can also be observed (*66-68*). Unfortunately, solution NMR has limitations of its own. The most prominent limitation is size of the protein under analysis. In order for chemically shifted resonances to be observed, fast isotropic tumbling must take place to obtain sharp lines and a good signal to noise ratio. As the molecular weight of the protein increases, the rate of tumbling decreases. Spectral resolution is further hampered by the presence of lipids or detergents which form a complex with IMPs and increase the total molecular weight of the macromolecular complex. In addition, most IMPs contain α-helical transmembrane domains consisting of non-polar amino acids which lead to spectral overlap due to the similar environments of each of the amino acids and, therefore, difficulties arise in assigning peaks. The use of detergents and/or lipids also limits the use of residual dipolar coupling techniques and the elucidation of long range NOEs (*69*). Advances in the field of protein labeling have aided residue assignment, the determination of long-range interactions and of relaxation parameters. These assignment strategies include uniform labeling (*70, 71*), selective amino acid (*72-74*) or methyl group labeling (*75*) and segmental labeling (*76-79*). Solid state NMR is also a highly relevant approach because there is no size limitation to the macromolecular complex and recently several protein structures have been solved to high-resolution (*80-83*). However, at this point, larger proteins have very poor signal to noise ratios leading to poor resolution of chemical shifts. Based on the current state of the art solution NMR

remains the most useful structural tool to study the dynamic conformational changes in IMPs and other proteins.

Structural analysis of integral membrane proteins has been an important goal of NMR spectroscopists and significant progress is being made (reviewed in *65*). However, few high-resolution structures have been reported. Many researchers are trying to focus on smaller full-length proteins, but the high-resolution structural determination has been difficult. For example, although the 40kDa homotrimeric diacyl glycerol kinase (DAGK) in DPC micelles has been studied intensively for many years and the assignment of the majority of the nuclei were reported in 2004 (*84*) and the full structure was reported five years later (*85*). Other proteins that have been studied are the *Staphylococcal* multidrug-resistance pump in bicelles (*86*), the *E. coli* multidrug transporter EmrE in chloroform:methanol:water (*87*) and the *Streptomyces lividans* KcsA potassium channel in micelles (*88*). The helical hairpin structure of subunit c of the F1Fo ATPase was solved in chloroform:methanol:water in its monomeric form (*89*). A few years ago, a group attempted to study the whole vasopressin receptor, a GPCR, by solution NMR, but only about 80 peaks were observed by TROSY NMR (*90*, *91*). Recently, a solution NMR analysis has been performed on a serpentine receptor sensory rhodopsin (*92*, *93*). The initial assignment of the backbone nuclei were important in the analysis of protein dynamics (*92*). Further assignment of the aliphatic methyl group moieties and the aromatic spin systems resulted in a high-resolution structure of a seven transmembrane domain protein by solution NMR (*93*). An important question that should be answered is whether the use of fragments of IMPs is a valid and useful strategy for the determination of protein structure.

Structural analysis by solution state NMR

Chemical shift assignments of peptide backbone atoms (i.e. N, NH, Cα, CH, CO) can be made by solution NMR experiments that determine through-bond connectivities such as [^{15}N,^{1}H]-HSQC, HNCA/HNCOCA and HNCO/HNCACO (*94*). These sets of experiments can be used to determine chemical shifts for the backbone nuclei of all protein residues. Once the chemical shifts have been determined, chemical shift indexing can be used as an indication of the secondary structure for each of these atoms (*95-97*). A comparison of ϕ and ψ angles in the structure database can also be used to approximate the secondary structure in a region of a peptide. A program called TALOS can be useful in predicting these angles by performing comparative analysis with a subset of proteins from the database (*98*). The program analyzes each residue in the context of a 3 amino acid window and looks for proteins in the database that have similar 3 amino acid chemical shifts. The current database is comprised of mostly soluble proteins which may lead to the misidentification of a given dihedral angle, as more IMP structures are determined by NMR, a new database may be compiled to increase the utility of this program.

NMR has become a very useful in the determination of the tertiary structure of a protein. Once specific amino acid residues have been assigned as described above, the side chain atoms can be assigned beginning with HN(CA)CB and CBCA(CO)NH experiments (*94, 99*). Additional experiments that can be used to determine the side chain assignments are HCCH-TOCSY (*100, 101*), HCCH-COSY (*102, 103*), and the combination of H(CCO)NH- and (H)C(CO)NH-TOCSY (*104, 105*). Due to the large numbers of aliphatic side chains in the TM regions of a GPCR, selective methyl group labeling can aid in assignments by using a HMCMCBCANH experiment which transfers resonances from the methyl groups of aliphatic

residues to the amide of the backbone (*106*) along with ct-[^{13}C,^{1}H]-HSQC and ^{13}C-resolved NOESY experiments. Aromatic residue assignment is also difficult because there is low chemical shift dispersion (*107*). A series of experiments that can be performed to aid in the assignment of the aromatic residues correlate the $^{13}C^{\gamma}$ with the remaining atoms on the residue. These experiments include CG(CB)H, CG(CD)H, CG(CDCE)H, CG(CDCECZ)H and CG(C^{aro})H-TOCSY (*108*).

Once all nuclei are assigned NOESY experiments that determine through space proton-proton interactions can be used to determine short-, medium- and long-range connectivities between protons. These experiments provide distance constraints that are used to calculate a backbone structure using structural calculation programs such as DYANA (*109*). Due to the spectral overlap of α-helical transmembrane domains and significant residue repetition, not all residues may get assigned. Most critical are connectivities identified between the two helical elements of the construct, as these will constrain the relative orientation of the helices. When sufficient connectivities are obtained, structural calculations are performed and the high-resolution structure is determined. This result can be compared to the structure of the same region in the rhodopsin-templated model of Ste2p and/or other high-resolution structures of GPCRs. The structure obtained in organic:aqueous membrane mimetic media can also be compared to that obtained in micellar membrane mimetic media. Chemical shift perturbation analysis can be used to determine if there are significant changes in the fragment structure in the two different environments.

The Integral Membrane Protein Folding Model

The biological function of any protein is dependent on its assuming the proper tertiary conformation, its native state. Intramolecular interactions are important for this folding to occur in an efficient manner. Folding of helical transmembrane proteins has been proposed to occur in a two-step model by Popot and Engelman (*110*). In this model, first the helices are formed and enter the membrane bilayer. Next, the helices will find each other and generate interhelical contacts, which help the protein to fold into the tertiary structure necessary for function. Synthetic single TM domains of GPCRs have been shown to form highly helical structures (*111-118*). In the case of bacteriorhodopsin, the structures of these peptides can be overlayed on the crystal structure with low RMSD values (*119*). Double TM constructs have also been generated that are highly helical as observed by circular dichroism (*120-122*). Receptor reconstitution assays *in vivo* that were performed using coexpressed fragments of Ste2p, have shown that when fragmented in the connecting loop regions between the TM α-helices Ste2p can reconstitute into a functional receptor (*123*). The data above indicate that sections of a protein can fold properly outside the context of the full receptor in support of the two-step folding model and can be used as validation for the use of fragment analysis for the determination of atomic resolution structures of Ste2p.

Förster Resonance Energy Transfer (FRET) has been used to elucidate interactions between helices in folded and partially unfolded proteins as well as to discern interactions between individual TMs. The helical interactions in bacterioopsin were studied using energy transfer from a dansyl group incorporated at Lys41 to the surrounding Trp residues (*124*). A loss of FRET was observed due to a partial unfolding which leads to a separation of helices in regions containing 2TMs, 3TMs and/or 5TMs. Coexpression of Ste2p fragments containing either

yellow or cyan fluorescent proteins at the C-termini resulted in FRET transfer, indicating that the 2 segments are close in space (*125*). More recently, biophysical studies utilizing FRET and circular dichroism (CD) analyses of individual TMs of the adenosine A_{2A} receptor indicate that, in SDS micelles, two individual TMs can find each other and form heterodimers (*118*). In some cases, the helicity of the individual helices increased indicating that heterodimer formation influenced the structure of the individual domains of the multimer. Based on the above experiments, one could argue that fragment analysis is a valid method of structural determination. Some of the questions that are still unanswered are: 1) what is the best membrane mimetic for these analyses and 2) how many TMs does a fragment need to contain to adopt a native-like conformation.

Stabilization of Structural Conformation in a Peptide Fragment

When deciding upon which fragment to study when performing fragment analysis of an IMP one must consider whether the two (or more) TM domains would have a tendency to form a tertiary structure. One way to think about this would be to use buried surface area analysis as an indicator of potential bond formation between the helices. If a pair of helices had a low buried surface area, then the assumption would be that the water and/or solvent molecules would have room to cage the helices and therefore prevent interhelical contacts. Conversely, if there were a high buried surface area, the bonding between helices would be better able to exclude water and/or solvent and a tertiary fold (or helical hairpin) would be more likely to form. A structural model of Ste2p based on rhodopsin is available (*8*) and would be very useful for this type of analysis.

However, the use of fragments for structural analysis raises the question about how these domains fold without the rest of the receptor. If the other helices are important for conformational stabilization and this stabilization is lost, the fragment might not assume a unique tertiary fold. To stabilize such a fold, covalent bonds between helices can be inserted. A possible strategy is to place two Cys residues into each TM in positions where the side chains are in close proximity. The length of a disulfide bond is approximately 2.05 Å so the two residues should be predicted to be close enough so as to not greatly perturb the structure. The resulting disulfide crosslink would generate a conformation that might mimic that of this region in the context of the full receptor. The introduction of non-native disulfide bonds has been used to stabilize many soluble proteins as indicated by a decrease in the effect of heat in unfolding or heat inactivation (*126-129*). Most recently it has been used to stabilize the structure of an immunoglobulin fold domain (*130*). Structural studies indicated that the residues that were mutated to Cys were within the hydrophobic core of the protein and were within the distance of a normal disulfide bond. Another method used to constrain a domain region is chemical ligation to generate a circular peptide. A linear fragment of the cCrk adapter protein containing an SH3 domain was cyclized by *in vivo* intramolecular chemical ligation, resulting in a 10-fold increase in binding affinity and the native SH3 folded structure as observed by NMR (*131*). Use of the rhodopsin-templated Ste2p model can predict residues that are in close proximity and which, therefore, are candidates for mutation to Cys residues. Expression of this construct and subsequent structural analysis by solution NMR can provide experimental data on the importance of covalent constraints in determining the tertiary structure of discrete regions of GPCRs. By using the biological assays described above to determine the effect of this mutation on function, the relevance of the structural data to the native state of the receptor can be evaluated critically.

Such information would be very valuable in comparing the suitability of various membrane mimetic media for structural analysis.

Unanswered Questions from the Literature

The use of fragments to study protein structure remains a controversial but widespread experimental design. The question of whether fragments can fold into a tertiary structure without the rest of the protein and whether this fold is the same in the context of the full IMP is still not completely answered. Furthermore, since the addition of detergents and/or lipids increases the MW of the macromolecular complex and decreases its isotropic tumbling in solution NMR, many researchers utilize organic aqueous solutions like trifluoroethanol (TFE) and dimethyl sulfoxide (DMSO) to solubilize membrane fragments they wish to study. The question arises as to whether such membrane mimetic media are comparable to the acyl chains of detergents and lipids. The goal of my research is to systematically investigate whether 1) a double transmembrane peptide fragment can fold into a tertiary conformation in an organic aqueous environment, 2) a double TM (2TM) peptide fragment folds into the same conformation in an organic aqueous environment and a lipid environment and 3) disulfide bond formation can stabilize a unique tertiary fold in both membrane mimetic environments. Success in addressing these questions will provide valuable information to guide biophysicists who wish to study the structure of other IMPs.

Specific Aims

To determine the three dimensional structure of a double TM fragment in membrane mimetic environments.

i. Choose a 2TM fragment that will potentially form a stable tertiary structure that has biological relevance to the ligand binding and/or signal transduction mechanisms of Ste2p.
ii. Determine a high-resolution structure of this double transmembrane fragment in organic aqueous solvents such as TFE:water and DMSO.
iii. Perform comparative analysis of the structure(s) determined above with the structure found in a micellar environment to determine whether organic:aqueous solutions and micellar solutions allow similar stable tertiary structures.
iv. Attempt to "lock" a tertiary fold using disulfide bond formation.

Chapter 2

Cloning, Expression and Purification of 2TM constructs of Ste2p

2a. Introduction

Prior to cloning fragments of GPCRs, or any other IMP, for structural analysis it is important to evaluate which fragments would have the highest potential for forming a tertiary fold. This can be done using a buried surface area analysis. The buried surface area is the region of a multiple TM fragment that is excluded from solvent interaction and can be determined by comparing the surface area of single TMs and that of their corresponding multi-TM regions. Theoretically, the higher the buried surface area, the more potential contact points between each of the TM domains. Discerning a construct with a large number of potential interhelical contacts provides a basis for choosing multiple TM domains that should have a high probability for folding. The surface area of single and multiple TM fragments of a GPCR can be calculated based on the structure or model of a protein using a software package such as NACCESS (*132*). For example, I can use a rhodopsin-templated model of Ste2p (*8*) and a structure of rhodopsin (*52*) to analyze various multi-TM fragments and decide which ones have the higher buried surface areas. Although this exercise may be helpful in determining which region of Ste2p to choose for further study, there are biochemical and molecular biological data, as described in Chapter 1, concerning regions of Ste2p that are important for ligand binding and/or G protein activation. Therefore, the biological relevance of the domain is also taken into account when considering a specific construct for further study.

Heterologous IMPs expressed in bacteria can be toxic because they may integrate into the membrane and could disrupt its integrity. One solution to this problem is to force the expressed

protein into insoluble inclusion bodies (IBs) shielding it from the membrane. The TrpΔLE fusion protein was first used by Staley and Kim to express bovine pancreatic trypsin inhibitor at high levels (*133*) and was generated from two *E. coli* trp leader peptide sequences, L and E with a deletion of region LE 1413 (*134, 135*). The promoter that drives the expression of the fusion protein is a T7 promoter sequence that is recognized by the T7 RNA polymerase which has a much faster processivity than that of the *E. coli* RNA polymerase (*136*). Fusion proteins containing the N-terminal TrpΔLE tag have been found in IBs when expressed in bacteria. These include fusions of BPTI (*133*), the HIV-1 gene *vpu* (*137, 138*), fragments of the cannabinoid receptor (*122, 139*), an isotopically labeled peptide fragment of a charged multi-vesicular body protein (*140*), caveolin (*141*), and the p7 protein from hepatitis C (*142*). Our lab has used this system to express and purify a single transmembrane fragment (*6*) and a double domain fragment, Ste2p(S267-S339), that included a single TM domain and a large portion of the CT (*115*). Therefore, cloning downstream of the TrpΔLE gene can be a useful strategy to ensure that the extremely hydrophobic membrane protein fragments described below are expressed into the insoluble fraction of the cells. Other fusion proteins that have been utilized with IMPs are glutathione-S-transferase (*39*) and thioredoxin (*120*). Expression of Ste2p(S267-S339) was also tested using the thioredoxin tag, but expression was reduced (*115*). Therefore, the 2TM protein fragments that were chosen using the buried surface area (BSA) analysis and biological information were cloned into the pMMHa based vector (*133*), which encodes the N-terminal TrpΔLE fusion protein, and expression was optimized as to expression strain, temperature and media such that yields of at least 5 mg of isotopically labeled purified peptide per L of culture were obtained.

2b. Materials and Methods

Buried Surface Area Analysis. Constructs containing one TM domain or two consecutive TM domains, but excluding loop residues, were generated from a rhodopsin-templated model of Ste2p (STE220614) (*8*) and a crystal structure of rhodopsin (1GZM) (*52*). Surface area (SA) analysis was performed on these model constructs using the computer program NACCESS (*132*). The amount of surface area that is buried between two or three helices can be determined by calculating the amount of SA of the packed TMs and subtracting this from the total SA of the individual TMs. The buried surface area (BSA) was determined using the equation:

$$BSA_{TM1TM2}=(SA_{TM1}+SA_{TM2})-SA_{TM1TM2}$$

where SA_{TM1} is the surface area of the first TM, SA_{TM2} is the surface area of the second TM and SA_{TM1TM2} is the surface area when the two helices are in the same 2TM construct. The difference is the amount of surface area buried from the solvent in the TM pair.

Table 2-1. Primers used in cloning and mutagenesis of Ste2p(G31-T110, TM1-TM2).

Primer Name[a]	Sequence of Primer[b,c]
M1M2 pSW02 for	5'-AGTACGCGCT**AAGCTT**TGGATGGGGAATGGATCTACCATCAC-3'
M1M2 pSW02 rev	5'-CTCGGTACCCGG**GGATCC**TCAAGTCACTGAAGAGTAATTAG-3'
M54L for	5'-CAGGCCATTttgTTTGGTGTCAG-3'
M54L rev	5'-CACCAAAcaaAATGGCCTGAGTAAC-3'
M69V/M71I for	5'-GATTGTCgttTGGatcACATCGAGAAGCAGAAAAACG-3'
M69V/M71I rev	5'-CGATGTgatCCAaacGACAATCAAAGTCAAAGCAGC-3'

[a]for=forward primer, rev= reverse primer, pSW02=name of original template.
[b]Underlined and bold sequences are restriction enzyme sites used for cloning.
[c]Lower case nucleotides are the nucleotides that are mutated.

Cloning of 2TM Protein Fragments. Cloning of pREJ04, which contains Ste2p (R231-S339, TM6-TM7-CT40), was performed previously by Racha Estephan and Jackie Englander (*143*). The region of Ste2p encoding Ste2p(G31-T110, TM1-TM2) was amplified by PCR using a Cys-less Ste2p gene in the pBEC2 plasmid (all primers are listed in Table 2-1). The vector pSW02 (*6*), which contained the TrpΔLE followed by the TM6 region of Ste2p, and the PCR product

amplified above, were digested with *Haemophilus influenzae* Rd restriction enzyme III (*Hind*III) and *Bacillus amyloliquefaciens* H restriction enzyme I (*Bam*HI) to remove the TM6 insert and replace it with the G31-T110 region by ligation using T4 DNA Ligase. Restriction digestion of the resulting plasmid was performed to verify the size of the inserted DNA and the clones of the correct size were sent to University of Tennessee for sequencing.

In order to remove the TrpΔLE from the fusion protein and liberate the target peptide a Met residue was incorporated between the fusion tag and the peptide of interest to allow for CNBr cleavage. Mutation of any native methionine residue found in the Ste2p(G31-T110) and/or Ste2p(R231-S339) regions was, therefore, necessary to ensure no internal cleavage of the GPCR fragments. The methionine residues in the Ste2p(R231-S339) were mutated previously (*143*). Three methionines in the Ste2p(G31-T110) region were mutated using a primer based double stranded DNA mutagenesis protocol. The specific mutations were chosen based on amino acid hydrophobicity (*144*) and retention of Ste2p functionality using results from a mutagenesis study of the intact receptor performed by the Dumont group (*145*). Furthermore, the codons chosen for use in the mutagenic primers were high usage *E. coli* codons (*146*). Briefly, two (forward & reverse) primers (Table 2-1) that have a T_m value higher than 58ºC were designed to yield the desired mutated nucleotide sequence. One primer annealed to the Watson strand of the DNA and replicated the entire plasmid, while the other primer replicated the entire plasmid by annealing to the Crick strand. For correct incorporation of the mutation into the plasmid DNA proper annealing of the primers is necessary. It is necessary to add at least 15 nucleotides to the 3' side of the desired mutation so that primer annealing is specific to the template. Primer melting temperatures and other characteristics were determined at http://www.basic.northwestern.edu/biotools/oligocalc.html. Invitrogen Platinum *Pyrococcus sp.*

polymerase (*Pfx* pol) was used under the following reaction conditions: ~400 ng template plasmid DNA, 50 pmoles of each primer, 2 mM dNTP mix, 1X Amplification Buffer, 2.5 mM $MgSO_4$ and 3.75 units Platinum *Pfx* polymerase. The negative control reaction included everything above except for the primers. The reaction cycle ran as follows: 94ºC for 3 min, 94ºC for 30 sec, 55ºC for 30 sec, 68ºC for 1.5 min/1 kb of plasmid (~5 min) (steps 2-4 cycled 35 times), 68ºC for 20 min and then the reaction was kept at 4ºC until it was removed from the thermalcycler. After the elongation protocol, the positive and negative reactions were mixed with water and 10X *Diplococcus pneumoniae* restriction enzyme I (*Dpn*I) (New England Biolabs) digestion buffer to a final volume of 400 μl. Eighty units of *Dpn*I were added to each mixture and digestion was allowed to proceed for several hours at 37°C. The *Dpn*I digestions were cleaned up using the Wizard DNA Clean Up Kit (Promega) according to the manufacturer's instructions. A small volume of the eluted DNA (1 μl) was transformed into 35 μl of MAX efficiency DH5α Competent Cells (Invitrogen) according to the manufacturer's instructions. After transformation, the cells were diluted into 250 μl of SOC media and then incubated with shaking at 37ºC for 1 hour. The cells were then plated on LBAmp plates (Luria-Bertani agar plates containing 200 μg/mL ampicillin) split 10% on one plate and 90% on a second plate and incubated overnight at 37ºC. Transformant colonies were picked and inoculated into 5mL of LBAmp (Luria-Bertani broth containing 200 μg/mL ampicillin) media, grown overnight, and miniprepped using the Wizard SV Miniprep Kit (Promega). When the M69V/M71I mutations were made the DNA was mutated such that there was a change in the restriction sites of *Moraxella osloensis* restriction enzyme I (*Msl*I) and *Flavobacterium okeanokoites* restriction enzyme I (*Fok*I). These enzymes were used to analyze the resulting plasmids to determine which ones contained the mutations. Plasmids exhibiting the correct digestion pattern were sent for

sequencing. The other mutation reactions were screened only by sequencing. Clone pLC01 was found to have the correct sequence to express TrpΔLE-Ste2p(G31-T110:M54L,C59S,M69V,M71I). pREJ04 had been previously determined to contain the correct sequence for TrpΔLE-Ste2p(R231-S339:M250A,C252S,M294A)(*143*).

Expression of TM1-TM2-FP in BL21(DE3)pLysS. The plasmid pSWM1M2(C59S), Ste2p(G31-T110, C59S) prior to Met mutagenesis, was transformed in BL21(DE3)pLysS. An overnight culture was grown at 37°C and then diluted 1:100 in 50 mL $LBAmp_{50}Cam_{34}$ (50 μg/mL ampicillin and 34 μg/mL chloramphenicol, respectively) at 37°C with shaking at 250 rpm to OD_{600} ~0.6-0.7 and induced with 1 mM IPTG. One mL samples were collected for analysis by Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) (see below). Western blot analysis was used to verify that the main band was the fusion protein. Expression optimization was performed by adjusting overnight dilution factors, growth and/or induction temperatures, and inducer concentration and/or induction OD_{600}.

Expression Optimization of 2TM Fusion Proteins. Plasmids pLC01 and pREJ04 were transformed into BL21(DE3)pLysS, BL21Star(DE3)pLysS and BL21-AI (Invitrogen) according to the manufacturer's instructions. Overnight cultures were each grown in $LBAmp_{200}Cam_{34}$ (200 μg/mL ampicillin and 34 μg/mL chloramphenicol, respectively) for the DE3 strains or $LBAmp_{200}$ for the AI strain. The overnight culture was diluted 1:20 into 50 mL of the respective LB media and the cells were grown to OD_{600} ~0.6 for DE3 strains and ~0.4 for the AI strain. The BL21(DE3)pLysS and BL21Star(DE3)pLysS cells were induced with 1 mM IPTG and the BL21-AI cells were induced with 0.2% L-arabinose. Cells were collected every hour to follow

the expression of the fusion protein and IB preps were generated for SDS-PAGE analysis. Expression for both 2TM fragments was optimized by testing different temperatures (22°C, 30°C, 37°C) and inducer concentrations (0.2%, 0.5%, 1% L-arabinose or 0.25 mM, 0.5 mM, 1 mM IPTG).

Overnight cultures of BL21-AI containing pLC01 or pREJ04 were diluted 1:20 into 2-500 mL flasks containing 100 mL $LBAmp_{200}$. These cultures were grown at 37°C at 250 rpm until they reached an OD_{600} of 0.1 and 0.4, respectively. The cells were then gently pelleted and split into 4-50 mL cultures containing M9 minimal media (MM:1g/L NH_4Cl or $^{15}NH_4Cl$, 20 mM KH_2PO_4, 48 mM Na_2HPO_4, 8.6 mM NaCl, 0.4% glucose or ^{13}C-glucose, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, and 200 μg/mL ampicillin). The culture that had grown to an absorbance of 0.1 was split into 2-50 mL cultures and grown at 37°C to an OD_{600} of ~0.4 and then induced with 1% L-arabinose and 1 mM IPTG. The cells that had grown to an absorbance of 0.4 were split into 2-50 mL cultures and induced right away as above. One of each set (0.1 and 0.4) was induced at 37°C and one of each set was induced at room temperature (~22°C). All four of the cultures were allowed to grow overnight and 1 mL samples were taken over time to follow the expression by SDS-PAGE.

Expression of Labeled Ste2p(G31-T110). Incorporation of ^{15}N and/or ^{13}C was accomplished using $^{15}NH_4Cl$ and/or ^{13}C-glucose in the MM using the optimized expression protocol. BL21-AI cells containing pLC01 [Ste2p(G31-T110)] are optimally induced at 0.4 OD_{600} using 0.5% L-arabinose and 1 mM IPTG with incubation at room temperature overnight. BL21-AI cells with pREJ04 [Ste2p(R231-S339)] are induced at the same optical density with 0.5% L-arabinose but

no IPTG and incubation at 37ºC overnight. Incorporation of the label was determined by ESI-MS.

Selective amino acid labeling was performed on BL21-AI containing pLC01. An overnight culture was grown in LB to 0.4 OD_{600} and then [^{15}N]-Ala was added just prior to induction with 0.5% L-arabinose (*74*). Incorporation of the ^{15}N moiety appeared low so other conditions in MM supplemented with unlabeled and labeled amino acids were evaluated (see below).

Deuteration of TM1-TM2-FP. Induction of expression from pLC01 in BL21-AI without adaptation was performed by inoculation of unlabeled minimal media with an overnight culture of pLC01 in BL21-AI (1:20) followed by growth at 37ºC until OD_{600} ~0.32. The cells were then pelleted gently and resuspended in minimal media containing various amounts of D_2O (39%, 59%, 79%) and $^{15}NH_4Cl$. Expression was induced with 0.5% L-arabinose and 1mM IPTG at room temperature overnight. Further analysis was performed with growth to OD_{600} 0.3-0.4 in LB or MM and then induction in ~99% D_2O minimal media (M9 salts were prepared in D_2O; glucose, $MgSO_4$, $CaCl_2$, Amp were prepared in H_2O). Further optimization was performed by growing the cells in 200 mL LBAmp to OD_{600} 0.3-0.4 then gently pelleting the cells and inducing in 50 mL of ~99% D_2O minimal media. Fusion protein was purified by HPLC and ESI-MS was used to determine incorporation of the ^{2}H and ^{15}N moieties. The expression levels in deuterated media was analyzed by SDS-PAGE and was quantitated using ^{15}N-TM1-TM2-FP expression as a control with the Odyssey Infrared Imaging System (LI-COR Biosciences, Lincoln, NB) (*147*).

Induction conditions were also examined with prior adaptation in deuterated minimal media. A 5 mL overnight culture was grown at 37ºC in LBAmp and then used to inoculate a 5 mL culture of 40% D_2O media, followed by transfer into 60% or 80% D_2O media. The culture in 60% D_2O MM was inoculated into 100 mL of 99% D_2O MM and a 5 mL 80% D_2O MM overnight culture. The 100 mL culture was induced as described above and pelleted after 21 hours. The 80% D_2O MM overnight culture was inoculated into MM containing 99% D_2O or 90% D_2O and they were induced as described. Final induction conditions that were used were to inoculate a 50% D_2O MM 5 mL overnight culture and use that to inoculate an 80% D_2O MM 50 mL overnight culture. The 50 mL overnight culture was split and 33 mL was used to inoculate 1L MM of $^{2}H(99\%)/^{15}N$ MM and 16 mL was used to inoculate $^{2}H(99\%)/^{15}N/^{13}C$ MM. The cultures were induced with 0.5% L-arabinose and 1 mM IPTG at OD_{600} ~0.35 and then grown at room temperature for 28 hours.

Selective protonation of methyl groups in Ile, Leu and Val[1]. Selective labeling of the aliphatic methyl groups (i.e. Val, Leu, Ile) can be very important in the determination of long-range connectivities in integral membrane peptide fragments. The following labeling protocol was adapted from Tugarinov and Kay (*75*). Cells were streaked onto a LB agar plate containing ampicillin and incubated overnight at 37°C. A 6 mL LBAmp culture was inoculated from a single colony and grown at 37°C until OD_{600} ~0.75 and then the cells were pelleted gently by centrifugation at 1400xg and resuspended in 20 mL M9 medium (as described above) to OD_{600} ~0.05-0.1. These cells were incubated with shaking at 37°C until OD_{600} ~0.6 and then they were gently pelleted in sterile tubes and resuspended in 100 mL M9 medium in 100% D_2O with 4 g/L

[1] This section was taken from Cohen *et al*. (*140*).

$^{13}C/^{2}H$-glucose and 1 g/L $^{15}NH_4Cl$. Incubation at 37°C with shaking occurred until OD_{600} ~0.4-0.5. The cells were then diluted to 200 mL in the same medium and grown again to OD_{600} ~0.4-0.5 at which point the cells were diluted to 1L in labeling medium (100% D_2O, 4 g/L $^{13}C/^{2}H$-glucose, 1 g/L $^{15}NH_4Cl$, 70 mg/L α-ketobutyric acid ($^{13}C_4$, 98%; 3,3-D_2, 98%) and 120 mg/L α-ketoisovaleric acid (1,2,3,4-$^{13}C_4$, 99%; 3,4',4',4'-D_4, 98%)). The precursor α-ketobutyric acid is converted in the cells to isoleucine that is protonated only at the δ-methyl group, whereas the α-ketobutyric acid is converted to both leucine and valine in which one of the branch point methyl groups is protonated. These cells were grown to OD_{600} ~0.3-0.4 and induced with 0.5% L-arabinose, grown at 37°C for 6-8 hours and then pelleted by centrifugation at 4000xg for IB preparation.

Isolation of Inclusion Bodies. The small (1 mL) inclusion body (IB) preps were generated by sonication-assisted lysis in 100-200 μl lysis mix which is lysis buffer (50 mM Tris-HCl pH 8.7, 1 mM ethylenediaminetetracetic acid) with 300 μg/mL lysozyme and 1 mM phenylmethylsulfonyl fluoride. Sonication was performed until the mixture was resuspended uniformly in the lysis mix followed by centrifugation at 21100xg for 20 minutes at 4°C. The supernatant was decanted and the pellet was resuspended by sonication in mild detergent (1% igepal Ca-630 and 1% deoxycholic acid in lysis buffer) and pelleted as above. The supernatant was discarded again and 100 μl SDS-PAGE loading buffer (4% SDS, 12% glycerol, 256 mM 2-mercaptoethanol, 50 mM Tris-HCl, pH 6.8 and a small amount of Coomassie Brilliant Blue G) was added and the pellet was resuspended by sonication.

Large-scale (50 mL to 1 L) IB preparations were generated as described (*6*, *115*, *148*). Briefly, frozen cells were thawed slightly, resuspended in lysis mix and sonicated in a room

temperature water bath until the cells appeared as a creamy uniform mixture. The 50 mL culture pellets were decanted into a 2 mL microfuge tube and centrifuged in a refrigerated tabletop microcentrifuge (Sorvall Legend Micro 21R). The larger pellets were transferred into 50 mL centrifuge tubes and centrifuged in a Sorvall (Sorvall RC5C Plus) centrifuge. The mixture was then centrifuged at 21100xg (for 2 mL microfuge tubes) and 39000xg (for 50 mL centrifuge tubes) for 20 minutes at 4°C. The resulting supernatant was decanted and the pellet was resuspended by sonication in lysis buffer and pelleted by centrifugation as above, and then the supernatant was decanted. Next, the pellet was resuspended by sonication in mild detergent and pelleted by centrifugation. Finally, the pellet was resuspended by sonication in water and pelleted to remove any remaining detergent. This washing procedure in water was performed twice. After the second sonication, the suspension was divided into eight-one mL aliquots per liter of the original growth in coated 1.5 mL microcentrifuge tubes. The samples were pelleted at 4ºC using 21100xg for 1 hour. For optimal protein yield, the supernatant at this last pelleting step should be clear. If it is a little cloudy then there will be fusion protein losses and longer centrifugation times are necessary.

SDS-PAGE Analysis. SDS-PAGE was used to analyze the expression of the fusion proteins. In general 16% gels were poured following a protocol supplied by Sigma based on Schagger et al 1987 (*149*). The following solutions were made in the preparation of the SDS-PAGE gels: 48% acrylamide solution (32:1 acrylamide:N,N'-methylene-bis-acrylamide), gel buffer (3M Tris, 0.1M SDS, pH 8.45). Gels were prepared with three layers; a 16% acrylamide separating gel, a 9.7% acrylamide spacer gel, and a 3.8% acrylamide stacking gel. Polymerization was induced by the addition of ammonium persulfate and N,N,N',N'-tetramethylethylenediamine.

The gels were run in a discontinuous buffer system. The two buffers used when running the gels were: anode buffer (0.2M Tris, pH8.9) and cathode buffer (0.1M Tris,0.1M Tricine, 0.1% SDS, pH ~8.2). The samples for SDS-PAGE were boiled for 2-3 minutes and 25 µl of each was loaded in the wells along with molecular weight markers. The gels were run at constant Amp (~20 mA for 2 gels) until the blue dye ran off the bottom of the gel. The gels were then stained in staining solution (0.1% Coomassie brilliant blue R, 40% methanol, 10% acetic acid) overnight and then destained in destaining solution (40% methanol, 10% acetic acid).

Western Blot Analysis. Western blots were performed using the SuperSignal West HisProbe Kit (Pierce) according to the manufacturer's instructions. The SDS-PAGE gels were run as described above and the gel was then incubated in cold Towbin buffer (25 mM Tris, pH ~8.3, 192 mM glycine, 20% methanol) for 15 minutes along with 2 pieces of thick filter paper and 1 piece of nitrocellulose paper per gel. One piece of filter paper was placed onto the Semi-Dry Transblot transfer apparatus and bubbles were removed by rolling with a pipette. The nitrocellulose was placed on top of the filter paper followed by the gel and then the second filter paper, removing bubbles at each step for complete transfer. The protein was transferred from the gel to the nitrocellulose at 12V with a 0.7 Amp limit for 30 minutes.

After the transfer, the membrane was incubated in 10 mL blocking buffer (1% BSA in TBST (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5% Tween-20)) with shaking at 4°C overnight. The blot was washed twice in 15 mL TBST for 10 minutes each and then incubated with the His probe-Horseradish perxidase diluted 1:5000 in 10 mL TBST for 1 hour at 4°C. The blot was then washed four times with 15 mL TBST for 10 minutes each at 4°C. Equal volumes (5 mL each) of the exposure solutions, SuperSignal West Pico Luminol/Enhancer Solution and

SuperSignal West Pico Stable Peroxide Solution, were mixed and the membrane was incubated in the solution for 5 minutes and then exposed to X-ray film for times ranging from 10 seconds to 1 minute depending on the intensity of the signal. The film was then developed.

Purification of TrpΔLE-Ste2p(G31-T110) by Ni^{2+}-NTA. IBs from a 125 mL culture of TrpΔLE-Ste2p(G31-T110) were solubilized in 2 mL of guanidine hydrochloride (GnHCl) solution (6M GnHCl, 10 mM Na_2HPO_4, 138 mM NaCl, 2.7 mM KCl; pH 7.4) with 5 mM imidazole by sonication. The solution was incubated with 0.5 mL Ni^{2+}-NTA Agarose (Qiagen) in a 15 mL conical tube overnight at room temperature with gentle shaking. Denaturing purification was performed at room temperature as described by Qiagen with some minor modifications. The beads were split into two columns and the flow through was collected. Both columns were washed with 2 mL urea buffer (8M urea, 20 mM sodium phosphate, 500 mM NaCl, 50 mM imidazole, pH 8). One column was then eluted with an imidazole gradient (250 mM imidazole, 500 mM imidazole and 1M imidazole) and one with a pH gradient (pH 6.8, pH 5.3, pH 4.0). Purification was analyzed by SDS-PAGE.

Purification of 2TM Fusion Proteins by RP-HPLC. Purification of the TrpΔLE-Ste2p(G31-T110) was performed on a Zorbax 300SB-C3 PrepHT column (21.2x150 mm) with a 30-90% acetonitrile, 0.1% TFA gradient at 50°C. Contamination noticed during further purifications (see below) led to adjustments made in the purification procedures. The analytical column Zorbax 300SB-C3 analytical column was cleaned by running a 40-80% isopropanol:butanol gradient in 20 minutes and 80-100% in 5 minutes at 50°C one time. After cleaning the acetonitrile gradient was run again at 60°C. The Zorbax 300SB-C3 PrepHT column was washed with 1 mL DMSO

in the acetonitrile gradient at 60°C until the levels of protein were insignificant. Fusion protein was then purified with a 40-90% acetonitrile, 0.1% TFA gradient at 60°C. After these first rounds of column cleaning, the columns were generally cleaned by injecting 70% TFA and running the acetonitrile:water gradient at the end of a set of purifications. MWs of the fusion proteins were analyzed by ESI-MS. Optimization of FP purification for both 2TM proteins are described below.

Cleavage of Purified 2TM-FP by CNBr. Purified TrpΔLE-Ste2p(G31-T110:M54L,C59S,M69V,M71I; TM1-TM2) (0.5-2 mg) was cleaved with a 1500-fold molar excess of CNBr (0.7-1 M) in ~50-100 μl 60% TFA. Reactions were incubated at room temperature in the dark. Cleavages were followed over time by analytical RP-HPLC as described above or with a similar gradient containing 10% isopropanol. Peaks were collected and analyzed by ESI-MS. Upon complete cleavage, the remaining reaction mixture was purified on a Zorbax 300SB-C3 Prep-HT column using a 36%-90% acetonitrile gradient with 10% isopropanol and 0.1% TFA.

Cleavage of 2TM-FP Directly in Inclusion Bodies and Purification by RP-HPLC. An IB pellet from a 500 mL culture was solubilized in 4 mL GnHCl buffer by sonication. A portion was removed (100 μl) and mixed with a large excess of CNBr. The reaction was followed over time by analytical RP-HPLC and purified on a Zorbax 300SB-C3 Prep-HT using a 45-80% acetonitrile gradient with 10% isopropanol, 0.1% TFA at 60°C. Peaks were collected and analyzed by ESI-MS. Another IB pellet from a 500 mL culture was solubilized in 60-70% TFA and treated with CNBr as above. Remaining IBs were sonicated in water and aliquoted into

silicanized microfuge tubes, pelleted at maximum speed and 4°C. These were stored at -20°C for further use.

2c. Results

Buried Surface Area Analysis and Choice of 2TM Constructs. Using the rhodopsin-templated model of Ste2p (*8*) the buried surface area was determined for each of the consecutive 2TM pairs (Table 2-2). The surface area of each TM domain and each consecutive 2TM pairs was determined using the NACCESS software. These TMs were taken directly from the templated model of Ste2p so they are in the predicted conformation that they would be in the 7TM construct. The 2TM construct with the highest BSA was TM1TM2. The high BSA indicates that the helices are close together in space (Figure 2-1, panel A) which would increase the potential for interhelical interactions and would therefore be predicted to have the highest potential to form a stable tertiary fold. These interhelical interactions would be from the conformation in the context of the full receptor and may not reflect the interactions in the individual 2TM. As the BSA decreases, the distance between the two helices increases which is clearly illustrated in Figure 2-1 which shows three space-filled models of TM1TM2 (high BSA, panel A), TM6TM7 (intermediate BSA, panel C) and TM4TM5 (low BSA, panel B). As the

Table 2-2. Buried surface area analysis of an X-ray crystal structure of Rhodopsin (1GZM) and a computer model of Ste2p (STE220614).

Ste2p TM regions	Buried Surface Area ($Å^2$)	Rhodopsin TM regions[a]	Buried Surface Area ($Å^2$)
TM1TM2	1348.5	TM6TM7	1343.3
TM2TM3	1253.5	TM1TM2	1232.7
TM3TM4	868.7	TM2TM3	1073.5
TM6TM7	844.6	TM3TM4	973.5
TM5TM6	793.8	TM5TM6	664.9
TM4TM5	472.2	TM4TM5	346.6

[a]Fragments analyzed contained only TM regions and no connecting loops.

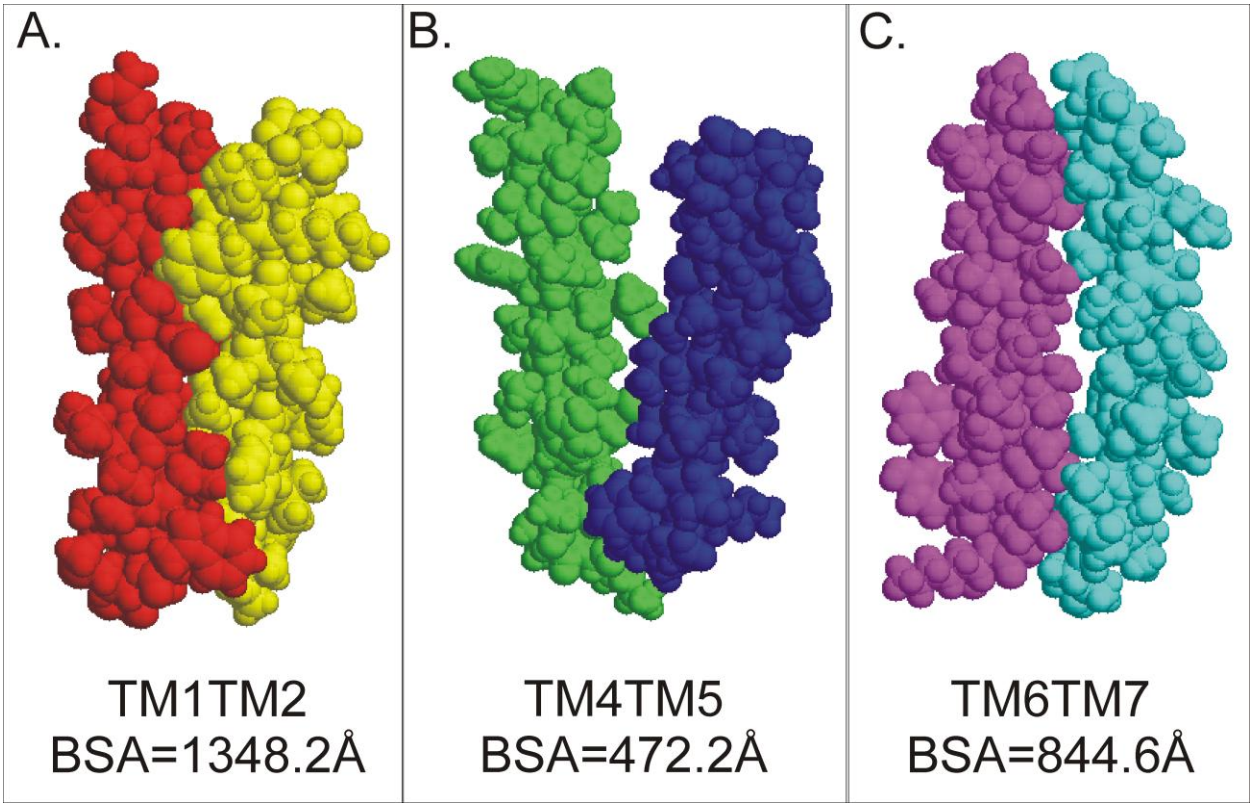


Figure 2-1. Three examples of BSA models. Space-filled models of three 2TM constructs were generated in Rasmol using the rhodopsin-templated Ste2p model developed by the Konopka group (*8*). A) TM1TM2: TM1 (red) and TM2 (yellow) are in close proximity to each other and correlate with the calculated buried surface area (BSA). B) TM4TM5: TM4 (green) and TM5 (blue) are not predicted to interact with each other as observed by the distance between the two helices and indicated by the low calculated BSA. C) TM6-TM7: TM6 (lilac) and TM7 (cyan) are predicted to have a few contacts as observed by the model and the intermediate BSA.

distance between the two helices increases, the potential for interhelical contacts decreases. For example, in these models the TM1TM2 helices are closely packed whereas the TM4TM5 helices splay apart and have few interhelical contacts. This led me to conclude that TM1TM2 would have a greater potential to form a stable tertiary fold.

The model of Ste2p was used to predict some of the interactions that play an important role in structural stability of this GPCR. Hydrogen bonds, electrostatic interactions and hydrophobic interactions may be helpful in the determination of which domains to express and purify for high-resolution structural determination. The Konopka group (*8*) have looked very

closely at these interactions in both rhodopsin and the Ste2p model. The 2TM construct with the highest BSA, TM1TM2, is predicted to have multiple interactions between the two helices. There are hydrogen bonds predicted between R58/H94/Q51 and hydrophobic packing interactions between A62/L88/I92 which are similar to a leucine zipper interaction. Other 2TM fragments, for example TM2TM3, are predicted to have interactions such as hydrogen bonding between S87 and S145 and between H94, Y98 and Q135. The two TM pairs that have intermediate to low BSA (TM3TM4, TM4TM5 and TM5TM6) are predicted to interact at only one residue for each pair. The low BSA results along with the few predicted interactions indicate that these 2TM fragments would not be desireable fragments for the structure analysis of 2TM fragments. Another 2TM pair with intermediate BSA is TM6TM7 which is predicted to have interactions between Q253 with S288 and S292. However, most predicted interactions for this pair are with other TMs. The helices that are implicated in signal transduction through ligand binding and conformational changes include TM1-TM2, TM5-TM6 and TM6-TM7 (*21, 150-153*).

Based on the above BSA, the predicted interactions between TM1 and TM2 and the functional importance of this region as described in the Introduction, we decided to move forward with the structural characterization of the 2TM fragment TM1-TM2 [Ste2p(G31-T110)]. Furthermore, although TM6TM7 had only an intermediate BSA, this region is important in signal transduction and since we had previously determined the structure of TM7 in organic aqueous and micellar environments (*115, 117*), we were also interested in pursuing the structural determination of TM6-TM7-CT40 [Ste2p(R231-S339)].

As a control, the BSA of a structure of rhodopsin was calculated as described above (Table 2-2). For most of the 2TM constructs, the order of BSA was similar. One major

difference between the two proteins analyzed was the BSA between the TM6TM7 helices. Based on buried surface area data alone, it would appear that the strength of the interaction of these helices is much stronger in rhodopsin than in Ste2p. I believe that the covalently bound ligand, retinal, forces the conformations of TMs 6 and 7 close to each other so that the interaction for these two domains is stronger in rhodopsin than in Ste2p. In fact, in the crystal structure of rhodopsin there are three observed contact points for the ligand retinal, within TM6 which also contacts TM7 (*8*). The differences in the way that the retinal ligand interacts with rhodopsin and the α-factor interacts with Ste2p are indicative of the need for high-resolution structural analysis of peptide binding GPCRs.

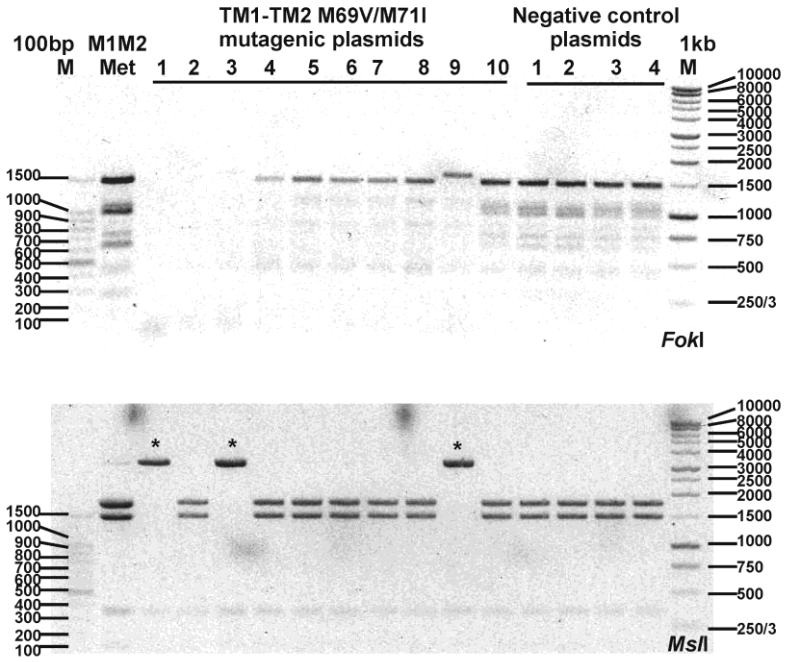
Further evaluation of each 2TM construct was performed to analyze the hydrophobicity which may play a role in purification and future analysis. Each Ste2p fragment was analyzed based on the Liu and Deber scale (*144*). Parts of the N terminus and the loops were included in the constructs to decrease the hydrophobicity of the molecules in order to aid in purification. The constructs that I considered for mutation took into account Met residues which would have to be mutated for future CNBr cleavage from fusion proteins (Table 2-3). I took into account mutations that needed to be made and the overall hydrophobicity of the resulting peptide. As the hydrophobicity number decreases, the potential solubility increases. TM6-TM7-CT40 was

Table 2-3. Hydrophobicity analysis of 2TM Ste2p domains based on amino acid composition.

Construct	**Residues**	**Mutations**	**Hydropbicity**
M1M2	G31-T110	M54L, C59S, M69V, M71I	41.08
M2M3	S73-R161	No mutations	40.02
M3M4	S121-K202	M165T, M180T, M189A	20.63
M4M5	T155-K239	M165T, M180T, M189A, M218S	32.02
M5M6 revisited	N194-D275	M218S, M250L, C252S	35.60

already constructed so it was not included in this evaluation. The optimal mutations that are presented in Table 2-3 had low hydrophobicity while considering the activity of mutated proteins *in vivo* that were isolated from a screen of the TM domains of Ste2p (*145*). All of the chosen mutations retain at least 80% activity *in vivo* and some had more than 110% activity. The mutated proteins found in the screen often contained multiple mutations. In my analysis I chose specific mutations that were found in several independent isolates. The final construct chosen for TM1-TM2 incorporated L, S, V and I residues (Table 2-3).

Cloning of 2TM Fragments of Ste2p. The regions of Ste2p that were chosen above were cloned by PCR and ligated downstream of the TrpΔLE sequence creating plasmids that coded for expression of TrpΔLE-Ste2p(G31-T110) and of TrpΔLE-Ste2p(R231-S339). All expressed fusion proteins were made to allow removal of the N-terminal TrpΔLE fusion protein by CNBr cleavage without cleaving the desired fragment. Analysis of the DNA for of the M69V,M71I mutagenesis indicated that three of the clones that were generated potentially contained the correct sequence (Figure 2-2). The Table in the Figure indicates the expected size of the DNA fragments after digestion with *Fok*I and *Msl*I. The results of the FokI digest were difficult to interpret based on size separation, but the MslI digest was definitive that clones 1, 3 and 9 had a change in sequence that was expected based on restriction analysis. The clones were sent for sequencing and clone 9 contained the correct sequence to code for the desired final product. The final constructs, pLC01 and pREJ04, contain the sequences for the expression of TrpΔLE-Ste2p(G31-T110:M54L,C59S,M69V,M71I), and TrpΔLE-Ste2p(R231-S339:M250A,C252S, M294A), respectively (Table 2-4). A cartoon map of pLC01 is shown in Figure 2-3. The cloning of pREJ04 had been done previously (*143*).



Restriction Enzymes	Restriction Fragment Lengths	
	M1M2(M54L,C59S)	M1M2(M54L,C59S,M69V,M71I)
*Fok*I	1478	1601
	643	643
	612	612
	463	463
	287	287
	181	181
	123	
*Msl*I	1778	3269
	1491	
	359	359
	159	159

Figure 2-2. Restriction enzyme analysis of M69V/M71I mutagenesis. Double-stranded DNA mutagenesis was performed to generate codon changes to create Ste2p(G31-T110; M69V,M71I). These codon changes changed the restriction patterns for two restriction enzymes, *Fok*I (top gel) and *Msl*I (bottom gel), which are described in the Table next to the gels. Restriction digests were performed on miniprepped plasmids that had no mutagenesis performed (M1M2 Met), that were from the mutagenesis reaction (TM1-TM2 M69V/M71I mutagenic plasmids) and plasmids that were from the negative control reaction (Negative control plasmids). The positive plasmids for mutagenesis are indicated by an asterisk.

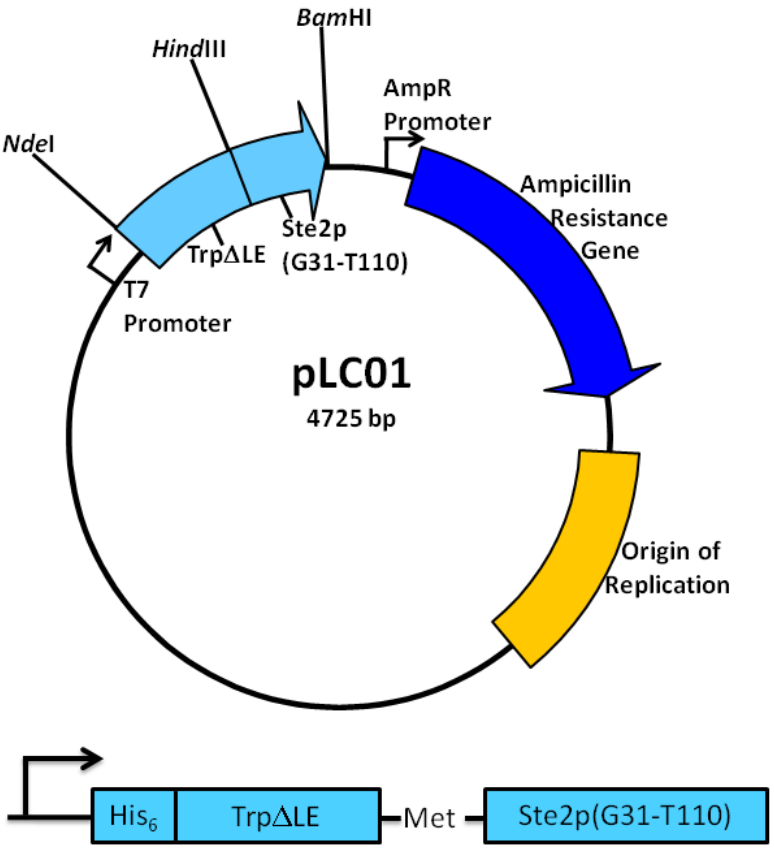


Figure 2-3. Plasmid that expresses the TrpΔLE-Ste2p(G31-T110) fusion protein. The plasmid pLC01 was based on pMMHa (*133*). A) The Origin of Replication allows for multiple copies to be made in an *E. coli* strain. The Ampicillin Resistance Gene allows *E. coli* cells containing the plasmid to be selected in growth media. Promoter regions direct the expression of the genes using either *E. coli* RNA polymerase (AmpR Promoter) or T7 RNA polymerase (T7 Promoter). The T7 RNA polymerase has a higher rate of processivity so more mRNA and therefore more protein can be made. The cloning sites that were used to create this plasmid are indicated above. B) The fusion protein being expressed is blown up in a cartoon format indicating the Met residue where CNBr cleavage takes place. This cleavage releases the target receptor fragment.

Initial expression optimization of 2TM constructs. A plasmid that expressed TrpΔLE-Ste2p(G31-T110, C59S), prior to Met mutagenesis, was transformed into the *E. coli* expression strain BL21(DE3)pLysS and induction with 1 mM IPTG was followed over time (Figure 2-4A). As judged by the band with a MW of ~22kDa, no expression of the protein was observed before induction with IPTG. The expression of this protein was observed to increase over time with maximum expression between 3 and 4 hours (Figure 2-4A, arrow). Although induction of this fusion protein appeared to be good, when compared to the expression of TM7-CT40-FP (*115*) about half as much protein was obtained. In order to acquire the amount of final peptide required to study the biophysical properties of this double domain fragment optimization of expression was needed. To confirm that the indicated protein band was the TM1-TM2-FP (calculated MW=22398.19 Da), the protein was purified from the inclusion bodies and ESI-MS was performed (Figure 2-5). The observed MW was 22384.83 Da which was within 3 Da of the correct MW so we moved forward with optimizing expression.

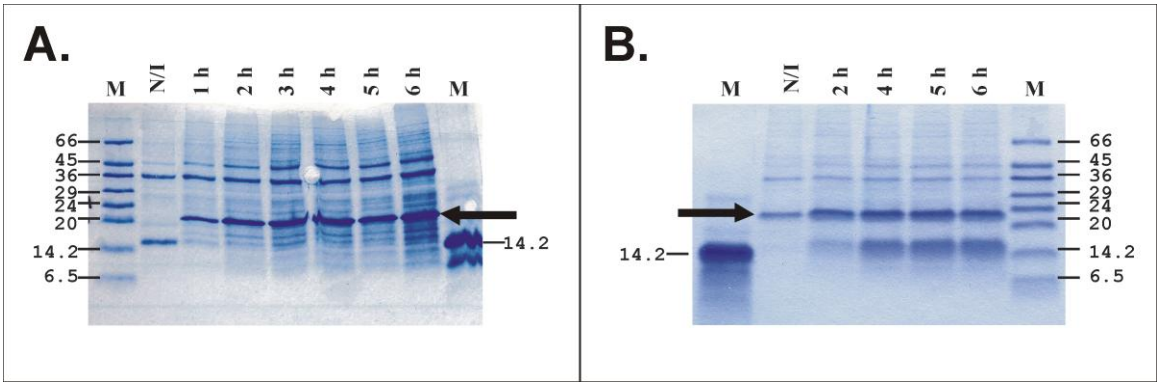


Figure 2-4. Initial expression of 2TM Fusion Proteins in BL21(DE3)pLysS. A) A plasmid expressing Ste2p(G31-T110) was transformed into BL21(DE3)pLysS and expressed in LB at 37°C for 6 hours. Cells (1ml) were collected every hour and IBs were generated to follow expression by SDS-PAGE. The arrow indicates a protein of ~22kDa that was induced upon addition of 1mM IPTG. B) The plasmid pREJ04 was transformed into BL21(DE3)pLysS and expression was followed as in A. The arrow indicates a protein of ~25kDa that was induced upon addition of 1mM IPTG. Analysis of protein expression from this construct indicates some expression prior to IPTG induction.

Table 2-4. Sequence and molecular weights of 2TM fusion proteins and peptide fragments with various isotopic label incorporations.

				Molecular Weight (Da)			
Protein Name	Plasmid	Amino acid sequence	Label Incorporated	Calculated	Observed	% Incorporation	Yield (mg/L)
TM1-TM2-FP	pLC01	MHHHHHHHHH-KAIFVLKGSLDRDLDSRIELELRTDHKELSEHLLLVDLARNDLARIATPGSRYVADLTKVDRYSYVLHLVSRVVGELRHDLDALHAYRAALNLGTLSGAPKVRAKLW(TrpΔLE)-M-GNGSTITFDELQGLVNSTVTQAILFGVRSGAAALTLIVVWITSRSRKTPIFIINQVSLFLIILHSALYFKYLLSNYSSVT	No label	22315	22318	N/A[a]	8
			^{15}N	22604	22587.7	95%	
			^{15}N,^{13}C	23614	23563	96%	
			^{15}N-Ala	22331	22328.6	85%	
			^{15}N-Val	22331	22330.3	96%	
			^{15}N-Ile	22327	22338.7	198%	
			^{15}N-Leu	22347	N/D[b]	N/D	
			^{15}N-Phe	22321	N/D	N/D	
TM1-TM2		GNGSTITFDELQGLVNSTVTQAILFGVRSGAAALTLIVVWITSRSRKTPIFIINQVSLFLIILHSALYFKYLLSNYSSVT	No label	8751.3	8749.2	N/A	10
			^{15}N	8852.3	8847.9	96%	10
			^{15}N,^{13}C	9258.3	9240.5	97%	8
			^{15}N-Ala	8755.0[c]	8752.6	52%	11
			^{15}N-Val	8756.5	8757.4	113%	11
			^{15}N-Ile	8759.2	8760.7	117%	20
			^{15}N-Leu	8760.9	8761.3	104%	16
			^{15}N-Phe	8754.9	8754.1	83%	21
TM6-TM7-CT40-FP	pREJ04	MHHHHHHHHH-TrpΔLE-MRSRRFLGLKQFDSFHILLIASSQSLLVPSIIFILAYSLKPNQGTDVLTTVATLLAVLSLPLSSAWATAANNASKTNTITSDFTTSTDRFYPGTLSSFQTDSINNDAKSS	No label	25325		N/A	
			^{15}N	25651	25639.9	97%	
TM6-TM7-CT40		RSRRFLGLKQFDSFHILLIASSQSLLVPSIIFILAYSLKPNQGTDVLTTVATLLAVLSLPLSSAWATAANNASKTNTITSDFTTSTDRFYPGTLSSFQTDSINNDAKSS	No label	11760		N/A	
			^{15}N	11898	11893.1	96%	3

[a]N/A is not applicable because no isotopes were incorporated.
[b]N/D was not determined for the samples indicated.
[c]Percent Incorporation of ^{15}N-labeled amino acids was determined using adjusted calculated MW based on unlabeled observed MW performed on the same day as the labeled MW.

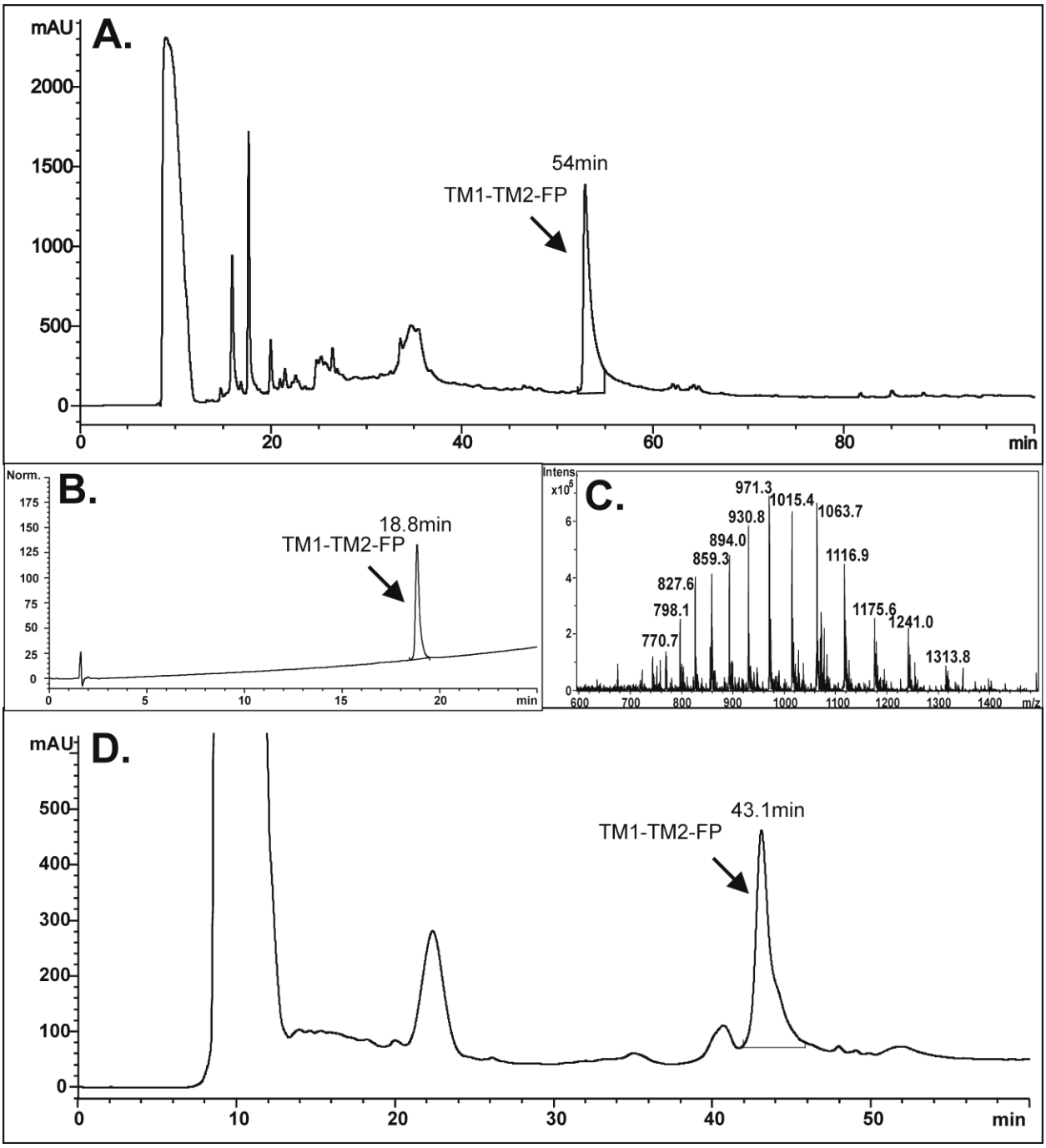


Figure 2-5. Purification of TM1-TM2-FP from Inclusion Bodies. A) The proteins solubilized from inclusion bodies were applied to a Zorbax 300-SB C3 Prep-HT column and separated by a 30-90% acetonitrile:water:0.1% TFA gradient at 50ºC, 5 ml/min for 90 minutes. The peak with a retention time of about 54 minutes (as indicated by the arrow) was collected, was observed to be homogeneous by analytic HPLC (B) and the MW was determined by ESI-MS (C). D) The same fusion protein was purified using a 40-80% acetonitrile:water gradient with 0.1 % TFA at 60ºC, 5 ml/min for 60 minutes. The TM1-TM2-FP eluted at approximately 43 minutes as verified by ESI-MS (data not shown).

Expression was first tested at two different temperatures, 30°C and 37°C. Cells were grown at either 30°C and 37°C to OD_{600} ~0.6 and induced with 1mM IPTG. Then both cultures were incubated at 37°C for 6 hours and IBs from 1-mL pellets were isolated for SDS-PAGE analysis (Figure 2-6A). The expression of TM1-TM2-FP increased over time and appeared to level off again between 3 and 5 hours post-induction. Similar expression was observed at 30°C and 37°C (Figure 2-6A). As plasmids containing the methionine mutations were generated [i.e. TrpΔLE-Ste2p(G31-T110, M54L, C59S)], the induction was tested and the expression was similar to that observed in Figure 2-6 (data not shown). Western blot analysis was performed to verify that the main band observed was the fusion protein. The anti-His antibody bound to the protein at 22kDa (data not shown) indicating that this protein was the fusion protein of interest.

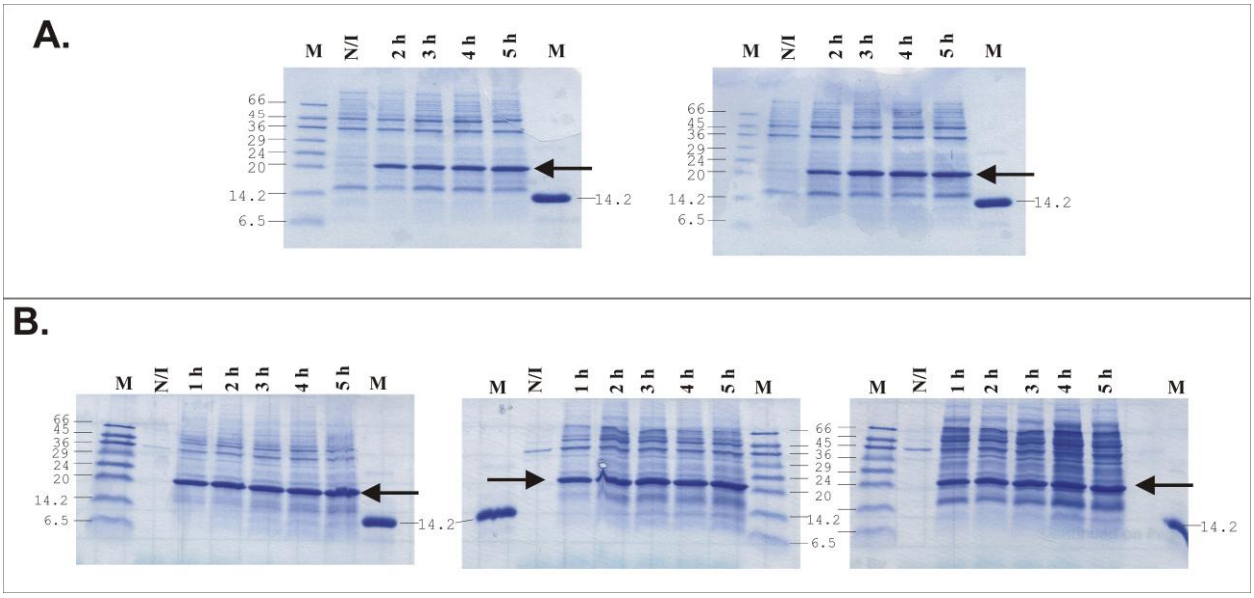


Figure 2-6. Optimization of the Expression of TM1-TM2-FP in BL21(DE3)pLysS. A) Optimization by initial growth temperature. BL21(DE3)pLysS cells that contained a plasmid expressing TrpΔLE-Ste2p(G31-T110) were grown at either 37°C (left gel) or 30°C (right gel) to OD_{600} ~0.6 and then both were induced with 1mM IPTG at 37°C. B) Optimization of expression for cells induced at various OD_{600}. BL21(DE3)pLysS cells that contained a plasmid expressing TrpΔLE-Ste2p(G31-T110) were grown at 37°C and induced with 1mM IPTG at OD600 0.6 (mid-log, left gel), 1.0 (late-log, middle gel), 1.5 (stationary, right gel). For all gels, cells were collected at the times indicated at the top of the gels, IBs were prepared and analyzed by SDS-PAGE. M=marker, N/I=non-induced.

Expression was also checked at different stages of the growth curve. Expression was induced at OD_{600} of 0.6, 1.0 and 1.5 (mid-log, late log, stationary phase). There appeared to be little difference in the expression of our target protein, but the amount of background proteins appeared to increase as the OD at which I induced was increased (Figure 2-6B). This was verified by the small-scale purification of the fusion proteins generated in this growth (data not shown). As the induction OD increased, the weights of the IB pellets increased. However, the peak areas of the fusion protein as determined by analytical RP-HPLC were very similar. I concluded that neither change in growth temperature nor OD of induction in BL21(DE3)pLysS significantly increased the level of expression of the TM1-TM2-FP.

Expression of the other 2TM construct, TM6-TM7-CT40, TrpΔLE-Ste2p(R231-S339; C252S, M250A, M294A), was also attempted in BL21(DE3)pLysS. A 1L culture was inoculated 1:20 with an overnight culture and induced at OD_{600} ~0.6 with 1mM IPTG at 37°C. SDS-PAGE analysis indicated that a low level of expression of a protein at approximately 24 kDa was observed prior to induction with IPTG and that expression of this protein appeared to increase over time and level off between 4 and 5 hours (Figure 2-4B). Western blots using an anti-His antibody indicated the 24 kDa protein was TrpΔLE-TM6-TM7-CT40 (data not shown). This 2TM fusion protein expressed at lower levels than both Ste2p(G31-T110)-FP (see above) and Ste2p(S267-S339)-FP (*115*) under these growth conditions. This data indicates that the induction level would not be sufficient to generate enough peptide for biophysical analysis so further expression optimization was required.

Expression of 2TM constructs in different expression strains. As described earlier, membrane proteins can be toxic to bacteria and strains BL21Star(DE3)pLysS and BL21-AI have been

developed specifically for the expression of toxic proteins. BL21Star(DE3)pLysS contains a mutation of the *rne* gene that encodes a truncated RNase E protein which increases mRNA stability because it lacks the ability to degrade RNA(*154-156*). BL21-AI has the T7 RNA polymerase gene as a chromosomal insertion, as opposed to a λ lysogen, which is under the regulation of the araBAD promoter and is, therefore, arabinose inducible (*157, 158*). It is also

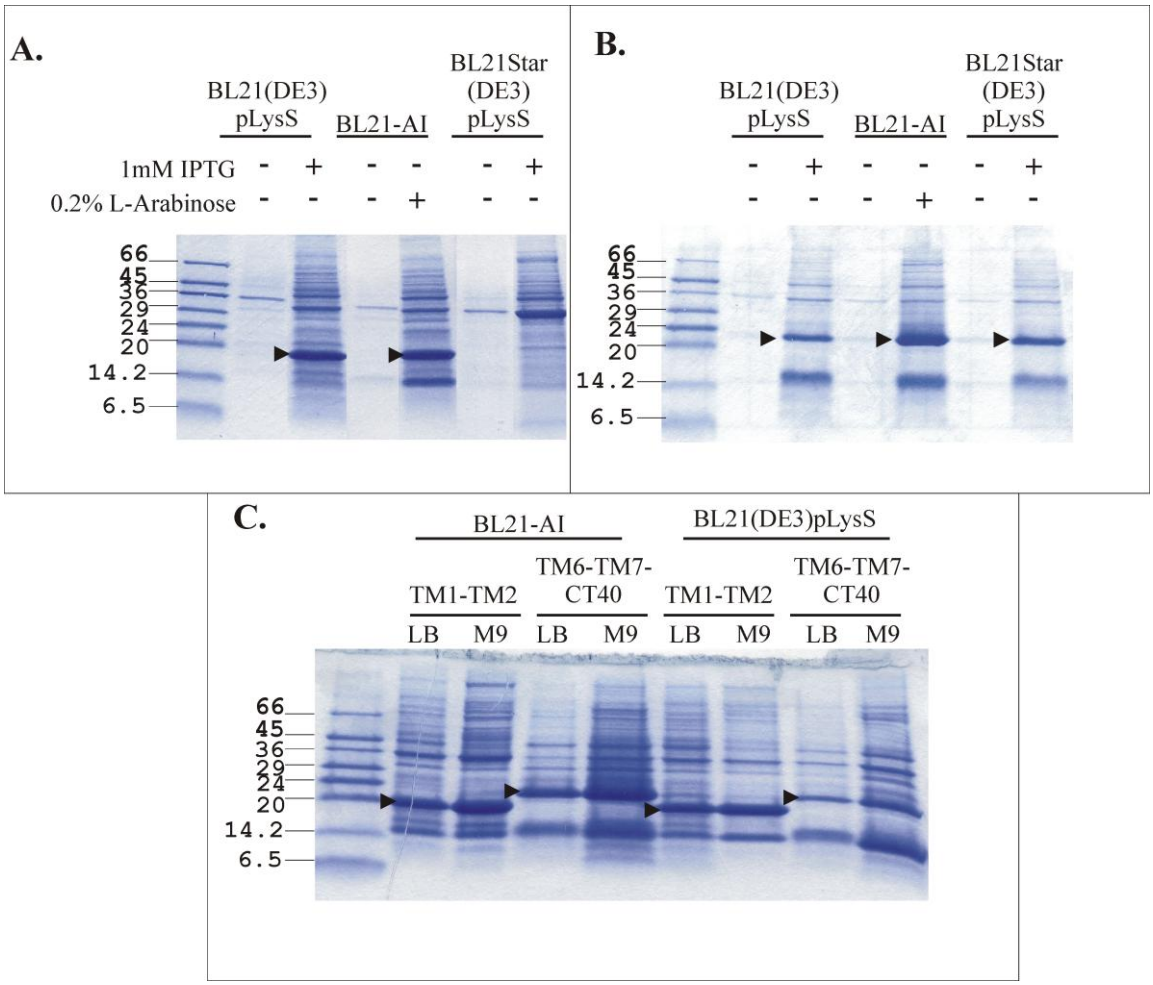


Figure 2-7. SDS-PAGE comparison of TM1-TM2-FP and TM6-TM7-CT40-FP expression levels in three different *E. coli* expression strains. A) TM1-TM2-FP and B) TM6-TM7-CT40-FP. Expression was induced with either 1 mM IPTG or 0.2% L-arabinose (as indicated at the top of the gels) and grown in LB at 37°C for 5 hours (TM1-TM2-FP) or 6 hours (TM6-TM7-CT40-FP) at 250 rpm. Inclusion bodies were prepared for SDS-PAGE analysis. The predicted proteins are marked by an arrow and the expected MWs are A) 22.3 kDa for TM1-TM2-FP and B) 25.3 kDa for TM6-TM7-CT40-FP. C) Comparison of expression in LB and minimal media for both double domain fusion proteins in BL21-AI and BL21(DE3)pLysS.

deficient in the outer membrane protease, OmpT, which reduces the degradation of heterologous proteins (*159*). pLC01 and pREJ04 were transformed separately into both strains, protein expression was induced according to the manufacturer's instructions (BL21-AI: 0.2% L-arabinose; BL21Star(DE3)pLysS: 1mM IPTG), and IBs were isolated and analyzed by SDS-PAGE. The expression of TM1-TM2-FP appeared to be similar in both BL21(DE3)pLysS and BL21-AI, but the correct size protein, indicated by the arrows, was absent in the BL21Star(DE3)pLysS cells (Figure 2-7A). Based on the SDS-PAGE analysis, TM6-TM7-CT40-FP appeared to be expressed in all three strains, with the highest expression observed in BL21-AI (Figure 2-7B).

Optimization of expression as to L-arabinose concentration and the presence of IPTG was performed in strain BL21-AI. TM1-TM2-FP was expressed in BL21-AI in LB under four different induction conditions at 37°C: 0.2% L-arabinose, 0.2% L-arabinose + 1 mM IPTG, 1% L-arabinose, and 1% L-arabinose + 1 mM IPTG (Figure 2-8A). The IPTG may help to repress the effects of the *lacI* gene which reduces expression if present. Even though the *lacI* gene was not present on the pLC01 plasmid, the expression at 5 hours appears to be slightly increased in the presence of 1mM IPTG when induced with 1% L-arabinose. When BL21-AI cells containing pREJ04 were induced, the expression appeared to be increased (0.2% L-arabinose, Figure 2-8B lanes 2 vs. 3), unchanged (0.5% L-arabinose Figure 2-8B lanes 4 vs 5) or decreased (1% L-arabinose Figure 2-8B lanes 6 vs 7) by the presence of IPTG. The optimal expression in LB medium for TM6-TM7-CT40-FP appears to be with 1% L-arabinose.

The expression of the fusion proteins was also optimized in M9 minimal media (MM). For the expression of the TM1-TM2-FP in MM, expression was tested under four different

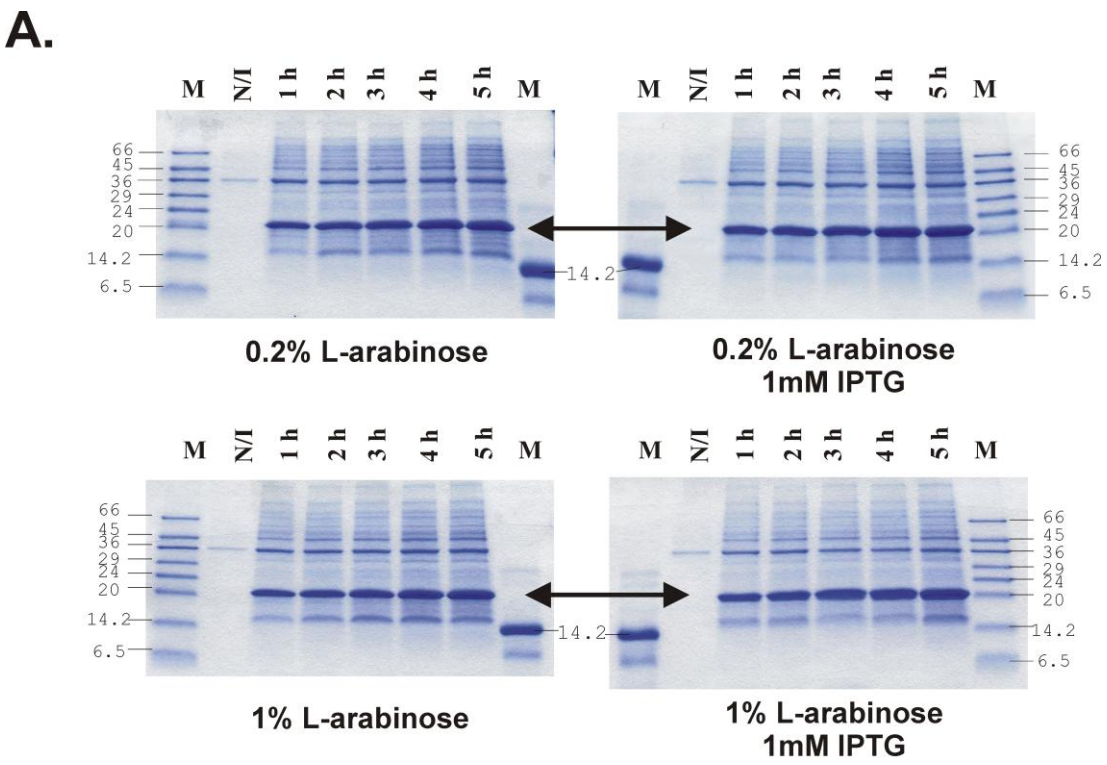


Figure 2-8. Expression optimization of 2TM fusion proteins in BL21-AI. A) BL21-AI cells containing pLC01 were grown at 37°C and induced with L-arabinose (0.2% or 1%) with and without IPTG and expression was followed every hour for five hours. SDS-PAGE analysis was performed on 1 ml IB preparations to check induction conditions. B) BL21-AI cells containing pREJ04 were grown at 37°C and induced as described above the SDS-PAGE gel. The first lane is a non-induced sample. The remaining inductions were performed at various L-arabinose amounts (0.2%, 0.5%, and 1%) with or without IPTG. The samples shown have been induced for 6 hours.

conditions (Figure 2-9). The cultures were grown at 37°C until they reached an OD_{600} of 0.1 and 0.4. The culture that had grown to an absorbance of 0.1 was pelleted and split into 2-50 mL cultures in minimal media containing $^{15}NH_4Cl$ and grown at 37°C to an OD_{600} of ~0.4 and then induced with 1% L-arabinose and 1 mM IPTG at either 37°C or room temperature (~22°C). The cells that had grown to an absorbance of 0.4 were split as above and induced at the two different temperatures immediately. All four of the cultures were allowed to grow overnight and expression was followed by SDS-PAGE. The samples that were transferred into MM at OD_{600} 0.1 did not induce well as indicated by the lack of a 22 kDa protein in Figure 2-9A. When transferred to minimal media at OD_{600} 0.4, the 22 kDa band appeared at 5 hours post-induction at room temperature (Figure 2-9B, left gel) and 2 hours post-induction at 37°C (Figure 2-9B, right gel). The protein expression at 37°C appeared to level off at 2 hours whereas the expression at room temperature peaked after an overnight growth. The overnight growth had the highest level of expression and the background was relatively low. A comparison of expression using 0.5% and 1% L-arabinose was performed due to the cost of L-arabinose and there was very little difference observed in the expression levels (see Figure 2-7C lanes 2 and 3 for expression with 0.5% L-arabinose).

Expression of the TM6-TM7-CT40-FP in minimal medium was tested with induction at room temperature and 37°C with 0.5% L-arabinose and 1mM IPTG (Figure 2-9C). Induction at room temperature resulted in very little of the fusion protein being expressed (Figure 2-9C, left gel). On the other hand, when induced at 37°C, the fusion protein begins to appear at 2 hours with the highest expression resulting from growth overnight (Figure 2-9C, right gel). Final optimization resulted in the protein being induced without the addition of IPTG (see below).

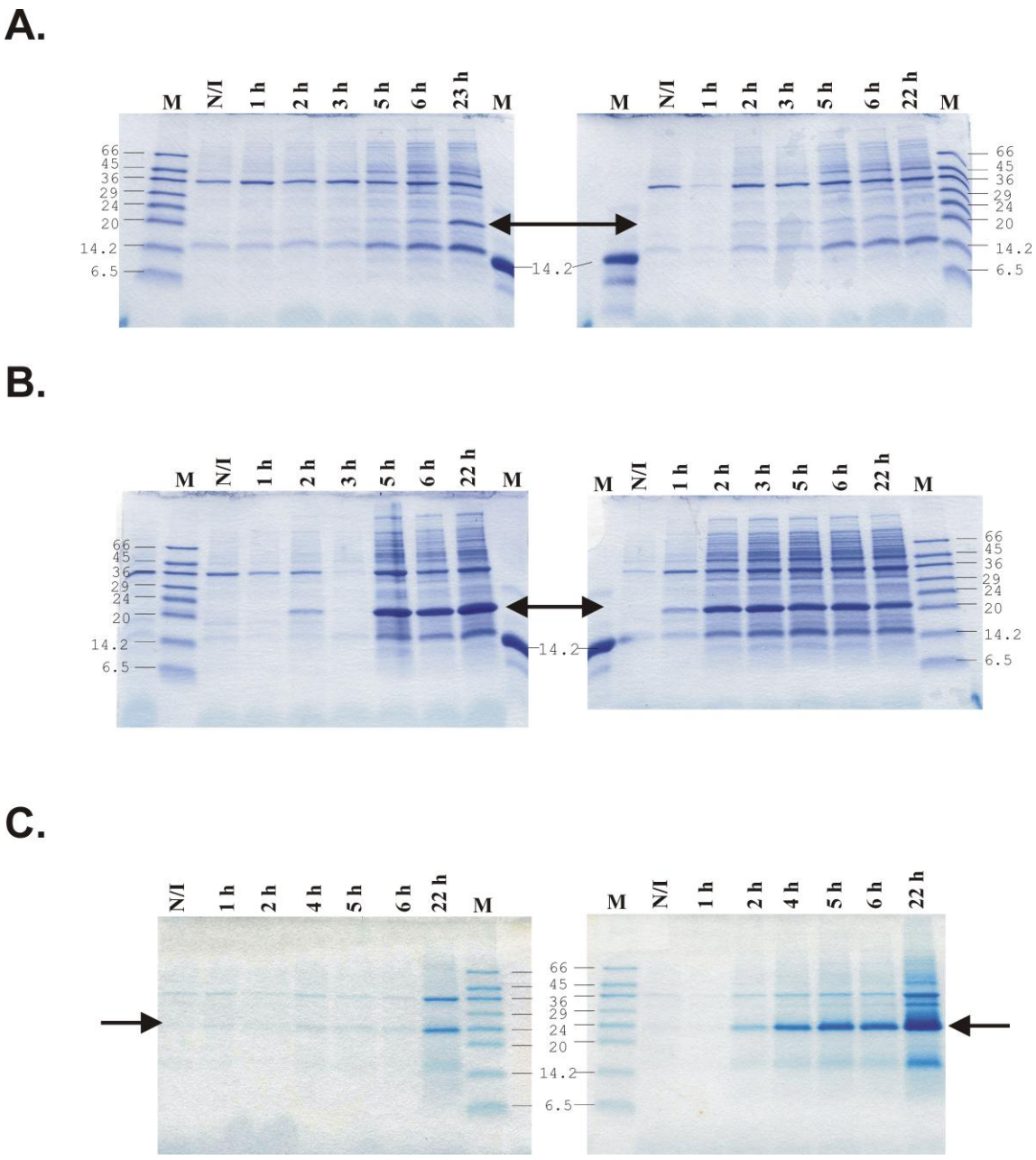


Figure 2-9. Expression optimization of 2TM fusion proteins in BL21-AI in minimal media. A) BL21-AI cells containing pLC01 were grown in LB until OD_{600} ~0.1. These cells were gently pelleted, resuspended in minimal media and induced at 37°C (right gel) and room temperature (RT; left gel) with 0.5% L-arabinose and 1mM IPTG when OD_{600} ~0.4. B) BL21-AI cells containing pLC01 were grown in LB until OD_{600} ~0.4. These cells were gently pelleted, resuspended in minimal media and induced at 37°C (right gel) and RT(left gel) with 0.5% L-arabinose and 1mM IPTG. C) BL21-AI cells containing pREJ04 were grown in LB until OD_{600} ~0.4. These cells were gently pelleted, resuspended in minimal media and induced at 37°C (right gel) and RT (left gel) with 0.5% L-arabinose and 1mM IPTG when OD_{600} ~0.4.

Somewhat surprisingly we observed that, in strain BL21-AI, the expression levels for both proteins were higher than in rich medium when the cells were induced in minimal media under optimized conditions (Figure 2-7C). An increase of expression in BL21(DE3)pLysS in MM was also observed for TM1-TM2-FP but not for TM6-TM7-CT40. The observation of high expression in MM was important for the preparation of isotopically labeled protein in sufficient amounts for NMR structural analysis (see below). The two fusion proteins were expressed in slightly different ways. TM1-TM2-FP was expressed optimally with induction at 0.4 OD_{600} using 0.5% L-arabinose and 1 mM IPTG with incubation at 22°C overnight. TM6-TM7-CT40-FP was induced at the same optical density with 0.5% L-arabinose but no IPTG and incubation at 37ºC overnight. These conditions were used to generate isotopically labeled fusion proteins containing ^{15}N and/or ^{13}C moieties (see below).

Large-scale expression of isotopically labeled TM1-TM2-FP. In order to produce mg quantities of our target peptides a 2 L culture of ^{15}N-labeled TM1-TM2-FP in BL21-AI was grown under our optimized conditions. The cells were pelleted and inclusion bodies were generated, cleaved with CNBr in 70% TFA, and the peptide was purified by RP-HPLC using a modified acetonitrile:water gradient (36-72% acetonitrile:water gradient with 10% isopropanol and 0.1% TFA at 60°C). The quality of the peptide as judged by high performance liquid chromatography (HPLC) is described in sections that follow. After multiple cleavages I recovered 10 mg/liter of culture of [^{15}N]-labeled peptide ($MW_{calculated\ for\ 100\%\ 15N\ incorporation}$=8852.3, $MW_{observed}$=8847.6).

In order to generate double labeled peptide for NMR analysis, expression of TM1-TM2-FP was induced in BL21-AI in minimal medium containing ^{13}C-glucose and $^{15}NH_4Cl$ as

described above. Analysis by SDS-PAGE indicated a high level of expression of the fusion protein at 22 hours after induction (Figure 2-10). Purification of the fusion protein for MS analysis showed excellent incorporation of both the ^{13}C and the ^{15}N moieties (96%, $MW_{calculated}$=22614, $MW_{observed}$=23562.95). The final yield of 97% labeled [^{15}N,^{13}C]-TM1-TM2 was 8.3 mg per liter of culture.

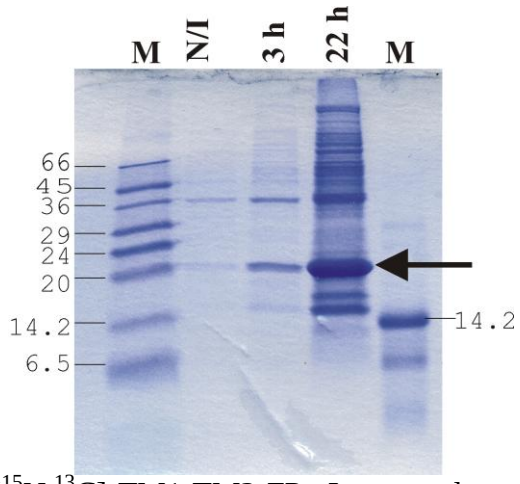


Figure 2-10. Expression of [^{15}N,^{13}C]-TM1-TM2-FP. Large-scale expression of TM1-TM2-FP in minimal media containing $^{15}NH_4Cl$ and ^{13}C-glucose was performed and analyzed by SDS-PAGE.

Expression of deuterated TM1-TM2-FP was performed in order to prepare NMR samples for use in 2D and 3D experiments (see Chapter 4). Originally the growth was performed without prior adaptation in deuterated media. Four different concentrations of D_2O were used (39%, 59%, 79%, and 99%). The growths were performed as described in the methods and the resulting cell pellets were analyzed by HPLC and ESI-MS (Table 2-5). SDS-PAGE analysis indicated that as the level of D_2O in the minimal media increased the expression level decreased (Figure 2-11A). The level of expression for the samples grown without adaptation in MM with up to 79% D_2O appeared to be sufficient to generate mg quantities of peptide for NMR analysis. However,

expression in 99% D_2O was too low. The cultures that were started in LB and then transferred to MM had lower levels of induction and would probably not be suitable as final expression conditions (data not shown). The results of the Odyssey analysis of the fusion protein expression correlate well with the HPLC data (Table 2-5).

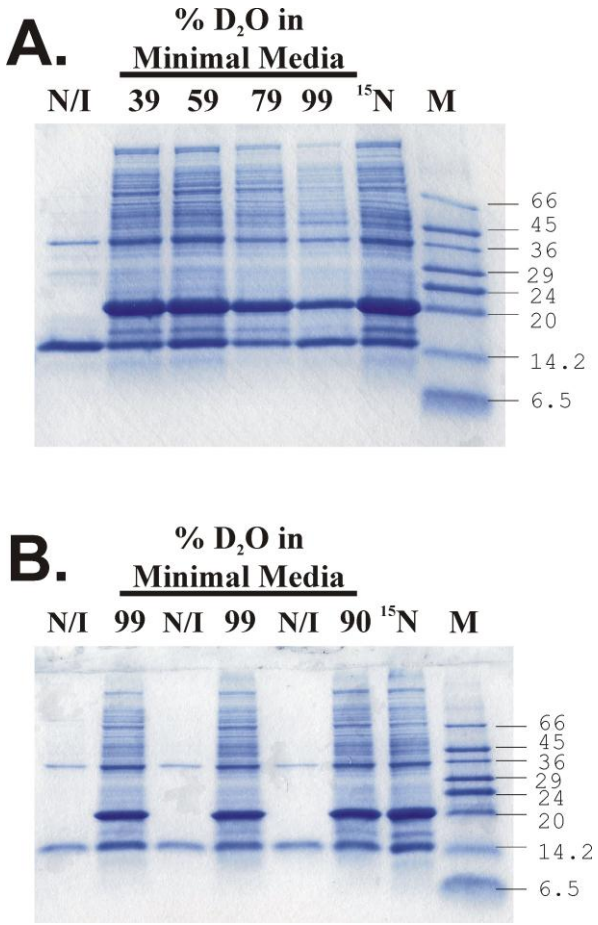


Figure 2-11 Expression of Deuterated TM1-TM2-FP. A) Expression of [^{2}H]-TM1-TM2-FP without prior adaptation. The cells were grown to OD_{600} ~0.4 in MM, gently pelleted and then induced in MM containing D_2O according to the percentages indicated at the top of the gel. [^{15}N]-TM1-TM2-FP was run on the same gel as a control of expression levels. B) Expression of [^{2}H]-TM1-TM2-FP with prior adaptation in MM containing 60% and/or 80% D_2O. An overnight culture was grown in 40% D_2O/MM and then transferred to 60% D_2O/MM or 80% and grown overnight. The cells that were grown in 60% D_2O/MM were used to inoculate 99% D_2O/MM (Lane 2) and to inoculate an 80% D_2O/MM overnight culture. This was used to inoculate 99% and 90% D_2O/MM cultures for induction (Lanes 4 and 6, respectively). [^{15}N]-TM1-TM2-FP was run on the same gel as a control of expression levels.

Expression of TM1-TM2-FP in BL21-AI cells that were adapted in MM containing D_2O was attempted in order to increase the expression in minimal media with higher amounts of D_2O. Cells were adapted at 40% D_2O/MM and then split into 60% and 80% D_2O/MM. The growth in 80% D2O/MM was poor, but the growth in 60% was good. The 60% D_2O/MM was inoculated into 100 mL 99% D_2O/MM containing $^{15}NH_4Cl$ for expression and 5 mL 80% D_2O MM for further adaptation. This 80% was split into 90% D_2O/MM containing $^{15}NH_4Cl$ and 99% D_2O/MM containing $^{15}NH_4Cl$ and the cells were grown and induced as described. Once again the samples were analyzed by HPLC, ESI-MS and SDS-PAGE (Figure 2-11B).In comparison to the expression of [^{15}N]-TM1-TM2 (Figure 2-11B, lane 7), it appears that the preadaptation in deuterated media increased the levels of expression of the fusion protein. This increase appears to be enough to get mg amounts of peptide after cleavage and purification. The levels of deuteration were dependent on the amount of D_2O in the minimal media (Table 2-5) and this information can be used as a guide for determining what labeling scheme to use for precise ^{2}H incorporation. This is important because NMR studies require different labeling schemes for optimal chemical shift assignment and NOE determination (see Chapter 4).

To prepare fusion protein that is approximately 80% deuterated, the cells were adapted in 50% D_2O/MM, followed by 80% D_2O/MM and finally inoculated in 100% D_2O/MM containing $^{15}NH_4Cl$ and unlabeled glucose or 99% D_2O/MM containing $^{15}NH_4Cl$ and ^{13}C-glucose. Expression was induced as described and the cells were pelleted, IBs generated and peptide cleaved from the fusion protein. Purified peptide was analyzed by ESI-MS and quantitated by UV to determine the amount of peptide. TM1-TM2 labeled with $^{15}N/^{2}H$ resulted in ~4 mg per L of culture with ~80% ^{2}H incorporation. TM1-TM2 labeled with $^{15}N/^{13}C/^{2}H$ resulted in ~2 mg peptide per L culture with ~60% ^{2}H incorporation (Table 2.5). Dr. Zerbe requested TM1-TM2

Table 2-5. Analysis of Protein Expression Levels and ^{2}H Incorporation into TM1-TM2-FP.

	Peak Area (per ml culture)	Relative Levels of Expression[b]	MW FP determined	FP % Deuteration	MW peptide determined	Peptide % Deuteration	Odyssey Analysis (Relative to ^{15}N)[c]
^{15}N, no deuteration							1.0
No Adaptation							
39% D_2O	42017	1.00	22931	42.5	9133	42.5	0.91
59% D_2O	30675	0.73	23706	57.5	9036	51.5	0.92
79% D_2O	31266	0.74	23333	67.5	8972	67.5	0.60
99% D_2O (LB)	1046	0.02	23459	70.0	n/d	n/d	0.19
99% D_2O (MM)	3316	0.08	23588	80.5	n/d	n/d	0.34
99% D_2O (4XLB)	921	0.02	25589	244.3	n/d	n/d	0.35
Adaptation							
40-60-99 D_2O and ^{15}N	10381	0.25	23571	79.1	9228	79.5	0.87
40-60-80-90 D_2O and ^{15}N	n/d	n/d	n/d[a]	n/d	9171	70.6	0.89
40-60-80-99 D_2O and ^{15}N	5223	0.12	n/d	n/d	9232	80	0.86
50-80-99(^{2}H,^{15}N,^{13}C)	5330	0.13	n/d	n/d	9633	58	0.75
50-80-99(^{2}H,^{15}N)	6017	0.14	n/d	n/d	9235.4	80.6	0.87

[a]n/d The MW of the peptide was not determined if the expression was low and cleavages were not performed. The fusion protein MW was not determined if cleavages were performed immediately.

[b]The Relative levels of Expression was determined by comparing the expression HPLC peak area of each fusion protein to the peak area of lowest amount of D_2O (39% D_2O).

[c]A sample of ^{15}N-TM1-TM2 was analyzed on each SDS-PAGE gel to normalize the Odyssey analysis.

that was ~60% ^{2}H incorporated with ^{15}N and ^{13}C labeling and based on the previous analysis I hypothesized that a growth in 70% D_2O MM would be appropriate. An overnight sample was prepared in 5 mL 40% D_2O MM, transferred to 50 mL 60% D_2O MM and then this culture was split evenly into 500 mL ^{2}H(99%) D_2O MM with ^{15}NH4Cl or ^{2}H(99%) D_2O with $^{15}NH_4Cl$ and ^{13}C-glucose. The double-labeled expression resulted in 78% ^{2}H incorporated peptide ($MW_{experimental}$=9215.6) and the triple-labeled expression resulted in 58% ^{2}H incorporated peptide ($MW_{experimental}$=9475.0).

Selective protonation of Ile, Leu and Val methyl groups in an otherwise deuterated background has been developed by the lab of Lewis Kay (*75, 106*). It is a particularly useful tool in the structure determination of membrane proteins due to their highly helical nature and the abundance of these hydrophobic residues. Furthermore, van der Waals contacts play an important role in the long-range interactions between helices. The precursor for Ile is α-ketobutyric acid and was purchased to incorporate a proton at the δ methyl position (Figure 2-12B, right). Leu and Val share a precursor, α-ketoisovaleric acid, which was used to incorporate protons at one of the methyl groups and deuterons at the other (Figure 2-12B, left). One of the drawbacks to this precursor is that it is generally sold as a racemic mixture and therefore is incorporated in a non-stereospecific manner. Adaptation of the Kay labeling procedure for our expression system was performed using the protocol described above, but with unlabeled glucose and α-ketobutyric acid and without α-ketoisovaleric acid. When the cells were ready for induction they were split into 4 different flasks and grown at 37ºC or room temperature for 6 hours or overnight. SDS-PAGE analysis was performed to determine the best growth conditions (Figure 2-12A) and the rest of the cells were pelleted, IBs were generated, the TrpΔLE was cleaved and the peptide was purified for analysis of isotope incorporation. HPLC analysis as

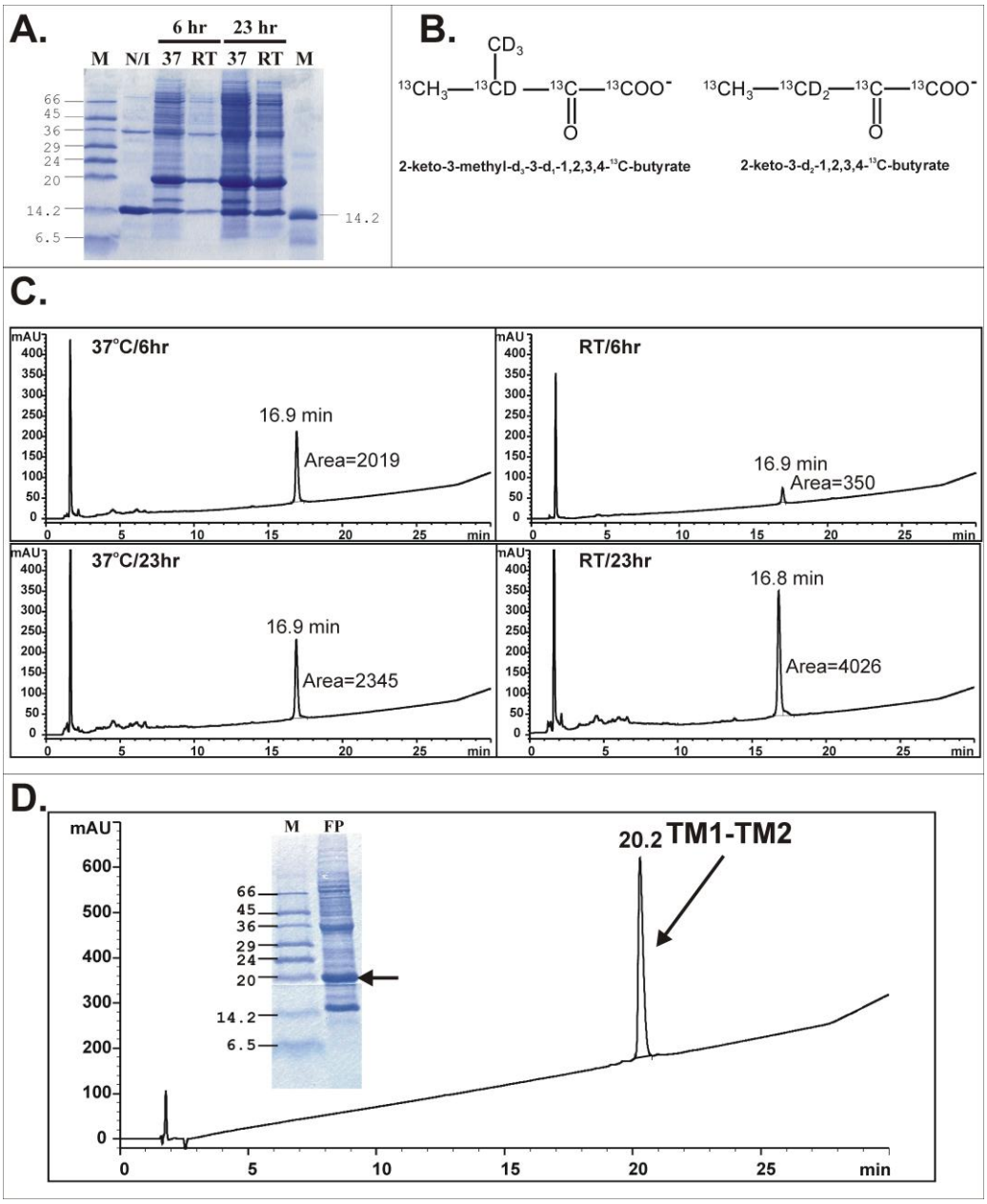


Figure 2-12. Optimization of expression of TM1-TM2-FP in selective protonation media. A) A test growth was performed using the conditions described except the glucose was not deuterated and the only precursor used, α-ketobutyric acid was not labeled. Induction of TM1-TM2-FP expression was performed as described and expression was followed at room temperature (RT) and 37ºC for 6 hours and 23 hours. IBs were generated and analyzed by SDS-PAGE. B) The α-ketoisovaleric acid precursor (left) and the α-ketobutyric acid (right) shown in the form used to incorporate selectively protonated methyl groups. C) Analytical HPLC analysis of the test growths. The samples are from the same growths as the gel in A. D) Inclusion bodies were generated from a large-scale growth in complete selectively protonated methyl group medium. The inset gel indicates the level of expression in this highly deuterated media. The chromatogram indicates the final cleaved and purified peptide as analyzed by analytical HPLC after purification.

well as the SDS-PAGE analysis indicated that induction at RT for 23 hours resulted in the highest expression levels (Figure 2-12C). The peptides that were generated from the 6h and 23h growths at 37ºC were analyzed by MS ($MW_{experimental\text{-}6h}$=9243.10, $MW_{experimental\text{-}23h}$=9243.33) and NMR. One-dimensional proton analysis of both peptides indicated that most of the peaks in the α-carbon region between 1.2 and 4 ppm were lost due to deuteration, but that the aliphatic region contained some protons (Figure 2-13, compare boxed regions in A and B). The aromatic region around 7 ppm in Figure 2-13B also indicates a loss of signal due to deuteration. Furthermore, each of the 2-dimensional [^{15}N,^{1}H]-HSQC had strong, sharp peaks which indicated that this sample would be useful in 3D experiments (see Figure 4-7B). The final optimized expression was performed using induction at RT for 23 hours which resulted in 6 mg purified peptide (Figure 2-12D) per liter of culture ($MW_{calculated}$=9702.29, $MW_{experimental}$=9660.42). The MW determined for the final peptide included incorporation of ^{15}N, ^{13}C, ^{2}H as well as ^{1}H at exchangeable protons and the indicated methyl groups. The overall incorporation was consistent for the expected labeling pattern (>99%).

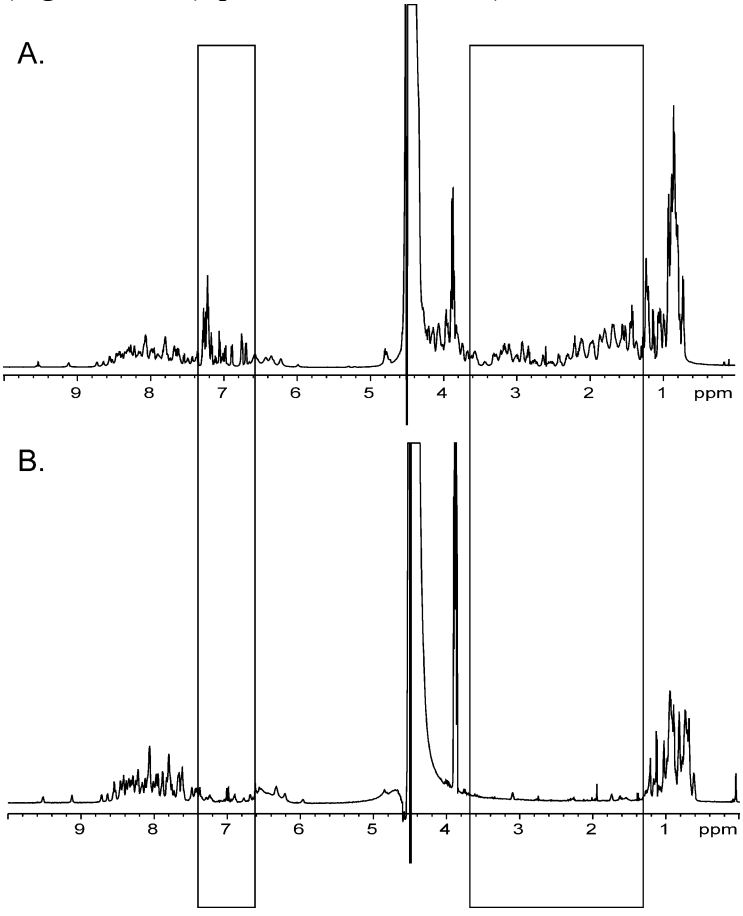


Figure 2-13. 1D ^{1}H NMR spectra of fully protonated and selectively methyl group protonated TM1-TM2. A) The ^{15}N-TM1-TM2 was prepared in TFE:water+0.1%TFA (v:v, 1:1) and a 1D ^{1}H spectra acquired at 45°C. B) The [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-TM1-TM2 was prepared in TFE:water+0.1%TFA (v:v, 1:1) and a 1D ^{1}H spectra acquired at 45°C. The boxes indicate changes in spectra due to the deuteration of the protein.

Large-scale expression of isotopically labeled TM6-TM7-CT40-FP. A 2L growth of ^{15}N-TM6-TM7-CT40-FP was performed. An overnight culture was diluted 1:20 into 2L of LB and grown to OD_{600} 0.375. The cells were pelleted in sterile centrifuge bottles and resuspended in minimal media containing $^{15}NH_4Cl$ and induced with 0.5% L-arabinose. The cells were then incubated with shaking at 37°C overnight as determined above and expression was analyzed by SDS-PAGE (Figure 2-14). Analytical purification of the FP and subsequent MS analysis indicated 97% ^{15}N incorporation based on the MW determined. The final peptide amount was low (~3 mg per L of culture) due to complications during the purification process (see below).

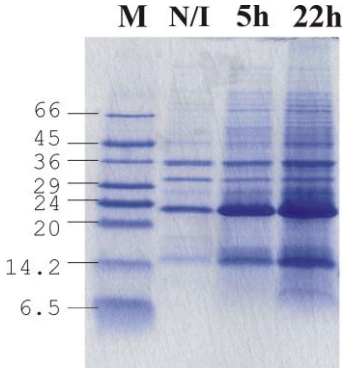


Figure 2-14. Expression of ^{15}N-TM6-TM7-CT40-FP. The plasmid pREJ04 was transformed into BL21-AI and expression of the fusion protein was induced as described. Cells were collected before induction (N/I), and 5 hours and 22 hours after induction (5h and 22h, respectively. IBs were generated and analyzed by SDS-PAGE.

Optimization of selective amino acid labeling. Selective labeling of the fusion protein with specific amino acids can be helpful in the assignment of chemical shifts in crowded [^{15}N,^{1}H]-HSQC spectra. We expected that the NMR spectra of the 2TM fragments would not be well dispersed because they are predicted to be highly helical and have high amino acid redundancy. Therefore, I optimized the incorporation of [^{15}N]-amino acids into TM1-TM2-FP. Originally, BL21-AI containing pLC01 was grown in LB to 0.4 OD_{600}, sterile [^{15}N]-Ala in water was added, and expression was induced with 0.5% L-arabinose and 1mM IPTG. The fusion protein was

purified on an analytical Zorbax 300SB-C3 column (see below) and ESI-MS was taken to check the incorporation of the [^{15}N]-Ala. Initial analysis determined that the incorporation was lower than expected, 66% in the fusion protein and 11% in the peptide, so the expression was then tested in M9 media supplemented with 2 different concentrations of 19 unlabeled amino acids, with and without NH_4Cl, and [^{15}N]-Ala. The optimal incorporation appeared to be 59% (Table 2-6). The incorporation was still not as high as I would have liked, but the yield of final peptide was greatly increased. The ESI-MS of the selectively labeled peptides was performed using unlabeled and selectively labeled on the same day due to a 2 Da difference in the calculated MW of TM1-TM2 and the observed MW of the unlabeled peptide. This normalization was very important in determining the proper percent incorporation especially when trying to determine the incorporation of five ^{15}N moieties into TM1-TM2 as was the case with Ala. The optimal condition for expression of [^{15}N-Ala] selective labeled TM1-TM2-FP was growth in minimal media supplemented with 0.5g/L of all of the unlabeled amino acids except Ala. The cells were grown at 37°C to OD_{600} 0.4 and then supplemented with 1 g/L [^{15}N]-Ala and induced with 0.5% L-arabinose and 1mM IPTG at 22°C for 22 hours. This resulted in about 52% incorporation of the [^{15}N]-Ala with a yield of 11 mg/L of purified peptide (Table 2-6). Even though our lab had previously optimized selective amino acid labeling in LB (*74*), the increase in peptide yield led me to continue using the protocol described here for the incorporation of isotopically labeled Val, Ile, Leu, Phe. The resulting incorporation was >80% and the yield was between 11-21 mg per L of culture of the final peptide (Table 2-4).

Table 2-6. Optimization of expression and incorporation of ^{15}N-amino acids into TM1-TM2-FP.

Growth Conditions[a]	$MW_{observed}$ Fusion Protein $MW_{calculated}$=22331	% Incorporation Fusion Protein	$MW_{observed}$ Peptide	% Incorporation Peptide[b]
LB[b] with ^{15}N Ala added at inoculation, 37°C, 5 hour induction	22325.55	66%	8751.65	31%
LB with ^{15}N Ala added at induction, 37°C, 5 hours	22325.11	63%	8751.85	35%
LB to 0.4 at 37°C, then MM with ^{15}N Ala at RT overnight	22347.46	203%	8760.18	202%
LB to 0.4 at 37°C, then MM+ 0.1g/L AA+^{15}N Ala at RT overnight induction	22320.35	33%	8751.14	21%
LB to 0.4 at 37°C, then MM+ 0.5g/L AA+^{15}N Ala at RT overnight induction	22323.52	53%	8751.70	32%
MM without NH_4Cl with 0.5g/L unlabeled AA at 37°C to 0.4 then add 1g/L ^{15}N Ala and induce at RT overnight	22328.57	85%	8752.62	59%
MM with NH_4Cl with 0.5g/L unlabeled AA at 37°C to 0.4 then add 1g/L ^{15}N Ala and induce at RT overnight	22324.06	57%	8753.00	52%
MM with NH_4Cl with variable g/L(*72*) unlabeled AA at 37°C to 0.4 then add 1g/L ^{15}N Ala and induce at RT overnight	22324.86	62%	n/d	n/d

[a]LB is Luria-Bertani broth, MM is M9 minimal media, AA is amino acids

[b]Analysis based on an average MW of unlabeled TM1-TM2 ($MW_{observed}$ unlabeled=8750.09) or unlabeled TM1-TM2 analyzed the same day.

Purification of TrpΔLE-Ste2p(G31-T110) by Ni^{2+}-NTA. The presence of a His_6 tag on the N-terminus of the fusion protein raised the possibility of purification of these constructs by Ni^{2+}-NTA chromatography. In order to test this, IB pellets were generated from a cell pellet of a 125 mL culture. The IBs were solubilized by sonication in 2 mL GnHCl solution as described above and incubated with the Ni^{2+}-NTA beads. Elution with imidazole or pH gradients was followed by SDS-PAGE analysis. In the gel of the imidazole elution (Figure 2-15, top gel) most of the

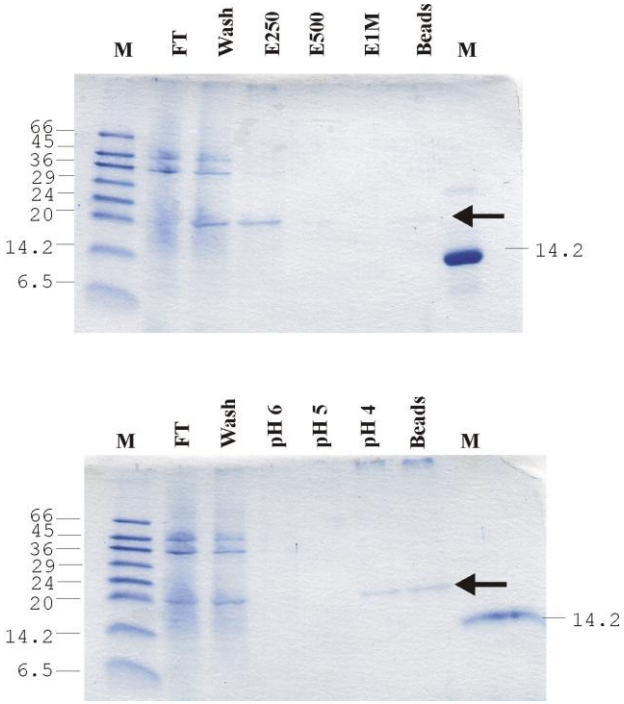


Figure 2-15. Purification of TM1-TM2-FP by Nickel affinity column chromatography. Top gel: Purification of TM1-TM2-FP under denaturing conditions by an imidazole gradient. The arrow indicates where the protein of interest would run. FT=flow through, W= 8M urea buffer with 50 mM imidazole,, E250=8M urea buffer with 250 mM imidazole, E500=8M urea buffer with 500 mM imidazole, E1M=8M urea buffer with 1M imidazole. Beads= Ni^{2+}-NTA resin after the E1M elution resuspended in water and an equivalent volume pelleted and resuspended in SDS loading buffer. Bottom gel: Purification of TM1-TM2-FP under denaturing conditions by a pH gradient. The arrow indicates where the protein of interest would run. FT=flow through, W= 8M urea buffer with 50 mM imidazole at pH 8,, pH6=8M urea buffer at pH6, pH5=8M urea buffer at pH5, pH4=8M urea buffer at pH4. Beads= Ni^{2+}-NTA resin after the pH4 elution resuspended in water and an equivalent volume pelleted and resuspended in SDS loading buffer.

fusion protein appeared to elute in the wash fraction, which contained 50 mM imidazole, and with the first imidazole elution, which contained 250 mM imidazole. The protein in the wash was not purified from the other proteins, but the protein in the 250 mM sample did seem to be cleaner. The amount of protein was too low to determine if all of the contaminants were removed. When elution was performed with a pH gradient (Figure 2-15, bottom gel), the protein did not bind to the column as determined by a band in the FT and the wash fractions. Furthermore, the protein that did bind was partially eluted at pH 4, but a lot of protein also remained on the Ni^{2+}-NTA beads. Overall, it appeared that the Ni^{2+}-NTA purification of TM1-TM2-FP would not be a good way to purify because either the protein was washed off of the beads as in the imidazole gradient or stuck on the beads as in the pH gradient. Similar results with the TM6-TM7-CT40-FP had been found by Racha Estephan and Jacqueline Englander from our laboratory (personal communication): the protein appeared to either stay on the column or co-elute with a second protein band. For these reasons, affinity purification was discontinued and purification using a C3 column on RP-HPLC was attempted.

Purification of TM1-TM2-FP by RP-HPLC. Inclusion bodies were purified on a Zorbax 300SB-C3 Prep-HT column (21.2x150mm) with a 30-90% acetonitrile:water gradient in the presence of 0.1% TFA at 50°C (Figure 2-5A). The separation of the fusion protein (retention time≈54 minutes) from the rest of the proteins in the IBs was very good and the MW (22318.03, $MW_{expected}$=22315, Table 2-3) was verified by ESI-MS (Figure 2-5). Fractions were collected and analyzed by analytical RP-HPLC. The final purity of the recovered protein was >95% purity (Figure 2-5B). Using this method of purification 6 mg per L of culture of TrpΔLE-Ste2p(G31-T110) was recovered. Unfortunately, this would only yield a maximum of 2 mg per L of culture

purified Ste2p(G31-T110) after CNBr cleavage. Therefore, I attempted to optimize the purification conditions.

Further analysis indicated that the fusion protein was sticking to both the analytical and preparative HPLC columns under the conditions used. The analytical column was cleaned by running a 40-80% isopropanol:butanol gradient in 20 min and 80-100% in 5 min. Under these conditions the fusion protein was removed, but when more was injected it would not bind. The temperature of purification was then raised to 60°C, IBs were injected and the gradient changed to 40-80% acetonitrile:water with 0.1% TFA over 60 minutes at 5 mL/min (Figure 2-5D). The fusion protein was purified and then the column was regenerated by washing with DMSO until levels of protein coming off were insignificant. Even under these conditions the fusion protein yield was low.

Table 2-7. Determination of optimal inclusion body volume to inject onto Zorbax 300SB-C3 Prep-HT column

Volume injected onto Prep-HT	Approximate culture Volume (ml)	Peak Area	Increase in Peak Area vs. 0.5 ml injection	Increase in Peak Area vs. 1 ml injection
0.5 ml	31.25	22314.9	1	
1 ml	62.5	63230.5	2.8	1
2 ml	125	106659	4.8	1.7

To ensure that the column was not being overloaded, the amount of fusion protein loaded onto the column was varied by injecting 0.5 mL, 1 mL and 2 mL of solubilized IBs onto the column. The peak areas as determined by purification on the PrepHT column were used to determine the optimum loading conditions (Table 2-7). As I increased my loading volume from 0.5 mL to 1 mL, the peak area of the fusion protein also increased by more than 2-fold. The increase wasn't quite as high when increasing the loading volume from 1 ml to 2 ml. In this

case, the peak area increased by less than double indicating that I was probably overloading the column. Based on these results, the rest of the TM1-TM2-FP from the 1L culture was purified by injecting 1 mL solubilized IBs (equivalent to 62.5 mL culture) onto the Zorbax 300SB-C3 PrepHT column as described. Two fractions were collected, the main peak and the right shoulder. Analytical HPLC analysis indicated that both fractions were identical. After the samples were lyophilized separately for each purification, the vials were combined by washing with TFE and then 70% acetonitrile +0.1% TFA. These samples were then lyophilized as peak and as shoulder and weighed (7.62 mg TM1-TM2-FP and 1.13 mg TM1-TM2-FP shoulder) resulting in about 9 mg per L of culture of fusion protein. Though the yield increased it appeared that purification was limited to about 0.5 mg per injection which would make purification very tedious. I performed CNBr cleavage of the purified fusion protein, but in order to increase the amount of peptide purified, we also decided to try to cleave the fusion protein in the IBs with CNBr.

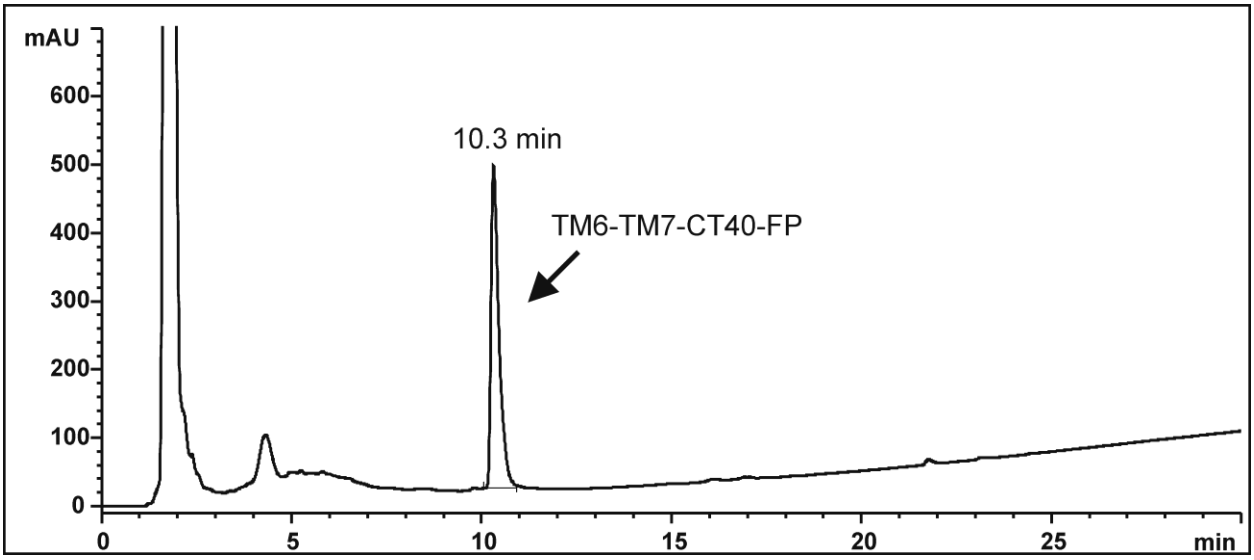


Figure 2-16. Analytical RP-HPLC analysis of TM6-TM7-CT40-FP. The plasmid pREJ04 was transformed into BL21(DE3)pLysS and induced with 1 mM IPTG for 6 hours at 37ºC. IBs were generated and analyzed by RP-HPLC on a Zorbax-C3 column using a 40-90% acetonitrile:water gradient with 0.1% TFA at 60ºC.

Purification of TM6-TM7-CT40-FP by RP-HPLC. Original purification of TM6-TM7-CT40-FP was attempted by Racha Estephan and Jacqueline Englander (personal communication). They reported that in a 30-60% acetonitrile:water (0.1%TFA) gradient at 50°C the protein appeared to aggregate on the column or that it did not stick at all. After I generated IB pellets from an unlabeled growth that was perfomed in BL21(DE3)pLysS, I was able to purify a small amount of fusion protein using an analytical Zorbax 300SB-C3 Prep-HT column with a 40-90% acetonitrile:water gradient with 0.1% TFA (Figure 2-16). The desired polypeptide eluted from the column at 10.3 min and the MW determined by ESI-MS ($MW_{calculated}$=25325.11, $MW_{observed}$=25336.6) indicated that this was the correct fusion protein. It was still very difficult to remove the fusion protein from the column, so direct cleavage was also attempted with this construct.

Cleavage of Purified TrpΔLE-Ste2p(G31-T110:M54L,C59S,M69V,M71I) with CNBr. Biophysical analysis requires that the 2TM peptide be separated from the fusion protein that was used to aid in its expression. In the case of the TrpΔLE construct a methionine residue links the fusion protein and the peptide. The peptide chain is released using CNBr in acidic conditions (*115*, *133*). Excess CNBr (0.67M, 1500 molar excess) was used to cleave 0.5 mg of purified TrpΔLE-Ste2p(G31-T110) in ~50 μl 60% TFA. The efficiency of cleavage was observed by analytical RP-HPLC and samples of the various peaks observed were collected for ESI-MS analysis but none of the peaks resulted in the expected MW.

Another cleavage was performed, this time with 1.6 mg TM1-TM2-FP in 1 M CNBr in 35 μL 60% TFA. The cleavage was followed by HPLC using a 30-90% acetonitrile:water gradient with 10% isopropanol and 0.1% TFA. Peaks were collected from the analytical HPLC

and ESI-MS was performed on the samples. Based on the data from the mass spectroscopy, the TrpΔLE eluted at about 5 min, the uncleaved TM1-TM2-FP elutes at about 18.5 min, and the TM1-TM2 peptide elutes at about 21.5 minutes (data not shown). The remaining cleavage reaction was loaded onto the Prep-HT column and the peptide eluted from the preparative column at 21.5 minutes. As I had identified conditions to purify the 2TM peptide corresponding to Ste2p(G31-T110) the remaining purified TM1-TM2-FP was cleaved and purified using a 45-80% acetonitrile gradient in water, 10% isopropanol, 0.1% TFA at 60°C.

Cleavage of TrpΔLE-Ste2p(G31-T110) directly in inclusion bodies. To increase the yield of peptide fragment recovered I hypothesized that cleavage within the IBs followed by purification of the desired peptide fragment might result in higher yields of the final product. The IB pellet was solubilized in 6M GnHCl solution and cleaved in excess CNBr and the reaction was followed over time by analytical HPLC (Figure 2-17A,B). After 3 hours the remaining reaction was purified on the Zorbax 300SB-C3 Prep-HT column using the gradient described above. The peak that eluted at about 34 minutes had a $MW_{observed}$ of 8749.15 ($MW_{calculated}$=8751.29). There was a significant proportion of TM1-TM2-FP that was uncleaved, so steps were taken to optimize the cleavage conditions. The same amount of IB was solubilized in 60% TFA and CNBr was added. The reaction was followed for 4 hours (Figure 2-17C,D) and the peptide was purified as described. The ESI-MS indicated the correct MW of the purified peptide and the amount of residual fusion protein was greatly reduced. The remaining IB pellets from this growth were sonicated in 4 mL of water and aliquoted into silicanized microfuge tubes (250ul/tube), pelleted and frozen for future cleavages.

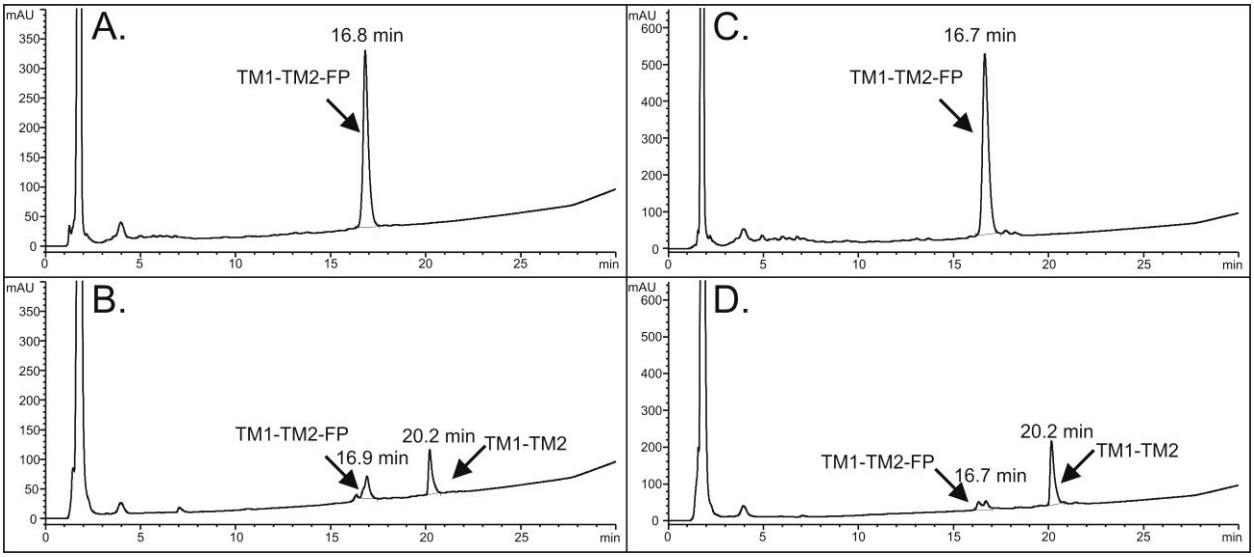


Figure 2-17: Cleavage of purified TM1-TM2-FP. A) TM1-TM2-FP was solubilized in 6M guanidinium HCl (GnHCl) and 60%TFA and analyzed by RP-HPLC on a Zorbax 300 SB-C3 column using a 36-72% acetonitrile:water gradient with 10% isopropanol and 0.1 % TFA at 60°C. CNBr was then added to the mixture and the reaction was followed over time. B) The cleavage was observed after 2 hours incubation at room temperature. The chromatogram at 4 hours was the same. C) TM1-TM2-FP was solubilized in 60% TFA and analyzed as above. CNBr was added to the mixture and the reaction was followed over time. D) The cleavage was observed after 4 hours incubation at room temperature.

Cleavage of IBs from a ^{15}N culture was performed and the peak at about 34 minutes was collected. ESI-MS analysis was performed and the ^{15}N-labeled peptide appeared to have a molecular weight of 8847.6. The expected MW if there were 100% incorporation of the ^{15}N moiety would have been 8852.3. The percent incorporation was determined by using the following formula:

$$[(MW_{labeled}-MW_{unlabeled})/(MW_{labeled\ expected}-MW_{unlabeled})]x100$$

The mass analysis of the peak collected indicated about 95% incorporation of ^{15}N into the peptide fragment. It also indicated that there was contamination of the fraction by unlabeled TM1-TM2 which was left on the column. Further optimization of the RP-HPLC conditions was needed to completely remove the peptide fragment from the column.

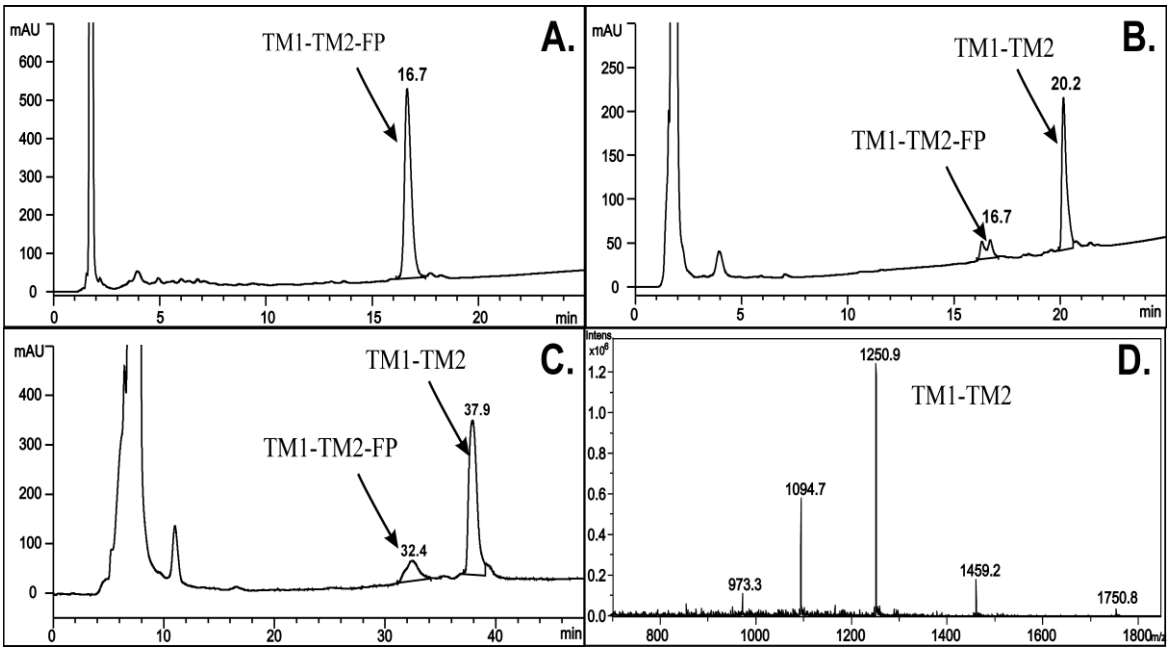


Figure 2-18. Direct cleavage of TM1-TM2-FP in inclusion bodies by CNBr. A) IB pellets containing TM1-TM2-FP were solubilized in 70% TFA and analyzed by HPLC B) 1 M CNBr was added to the reaction and cleavage was followed over time (HPLC spectra at 1.5 to 4 hours were very similar) by analytical HPLC on a Zorbax 300-SB C3 column using a 36-72% acetonitrile:water gradient with 10% isopropanol and 0.1% TFA at 60°C with a flow rate of 1 ml/min. C) The peptide was purified on a Zorbax 300-SB C3 Prep-HT column and separated by a 50-90% solvent B gradient at 60°C with a flow rate of 5 ml/min. D) ESI-MS was used to verify that the peptide was correct ($MW_{calculated}$=8751.29, $MW_{observed}$=8749.2).

IBs were solubilized in 70% trifluoroacetic acid (TFA), 1M CNBr was added, and the reaction was allowed to continue at room temperature in the dark until cleavage was complete as judged by analytical RP-HPLC. After 4 hours, the sample was injected onto a Zorbax 300SB-C3 analytical column and the TM1-TM2-FP eluted at about 16.7 minutes and the peptide eluted at about 20.2 minutes (Figure 2-18A and B). The sample was injected onto the Prep-HT column and the proteins were purified using a 36-72% acetonitrile:water gradient with 10% isopropanol and 0.1%TFA at 60°C (Figure 2-18C). ESI-MS analysis indicated that the peptide purified had a MW of 8749.28 ($MW_{calculated}$=8751.3) (Figure 2-18D). The isopropanol was used to increase the hydrophobicity of the solvent to help in the removal of the peptide. Furthermore, the increase in

temperature decreased the viscosity of the isopropanol solvent so as to decrease the column back pressure and it also helped to remove the peptide from the column. Any remaining fusion protein and the resulting peptide were well separated. The procedure resulted in highly purified peptides with yields ranging from 8 to 21 mg per L culture of TM1-TM2 depending on growth conditions (Table 2-4).

Cleavage of TrpΔLE-Ste2p(R231-S339) directly in inclusion bodies. Two mL of IB pellets from the TM6-TM7-FP induction in LB were removed from the freezer, were sonicated in 240 μl TFA and combined. Water was added slowly to a final volume of 400 μl. Some of the proteins appeared to precipitate, and 80 μl of TFA was added and the sample was sonicated to try to redissolve the proteins. CNBr was added to 1M and the cleavage was incubated in the dark at room temperature for 1 hour followed by analytical RP-HPLC purification (Figure 2-19). The cleavage was not complete as indicated by the large FP peak still present. When purified on a prepHT column, the peptide that was collected was analyzed by ESI-MS to determine the MW ($MW_{calculated}$=11760.4, $MW_{observed}$=11760.54). UV quantitation was used to determine the amount of peptide generated, which was ~0.85 mg. Further purifications were performed, though the cleavage was not optimized and a lot of FP was left uncleaved and not collected. In the end, 3 mg of TM6-TM7-CT40 was purified from 1 L of LB and 1.6 mg of ^{15}N-TM6-TM7-CT40 was purified from 2 L of minimal media. Though the purification yields were low there was sufficient TM6-TM7-CT40 to perform initial biophysical analyses (i.e. CD and [^{15}N,^{1}H]-HSQC, *vide infra*).

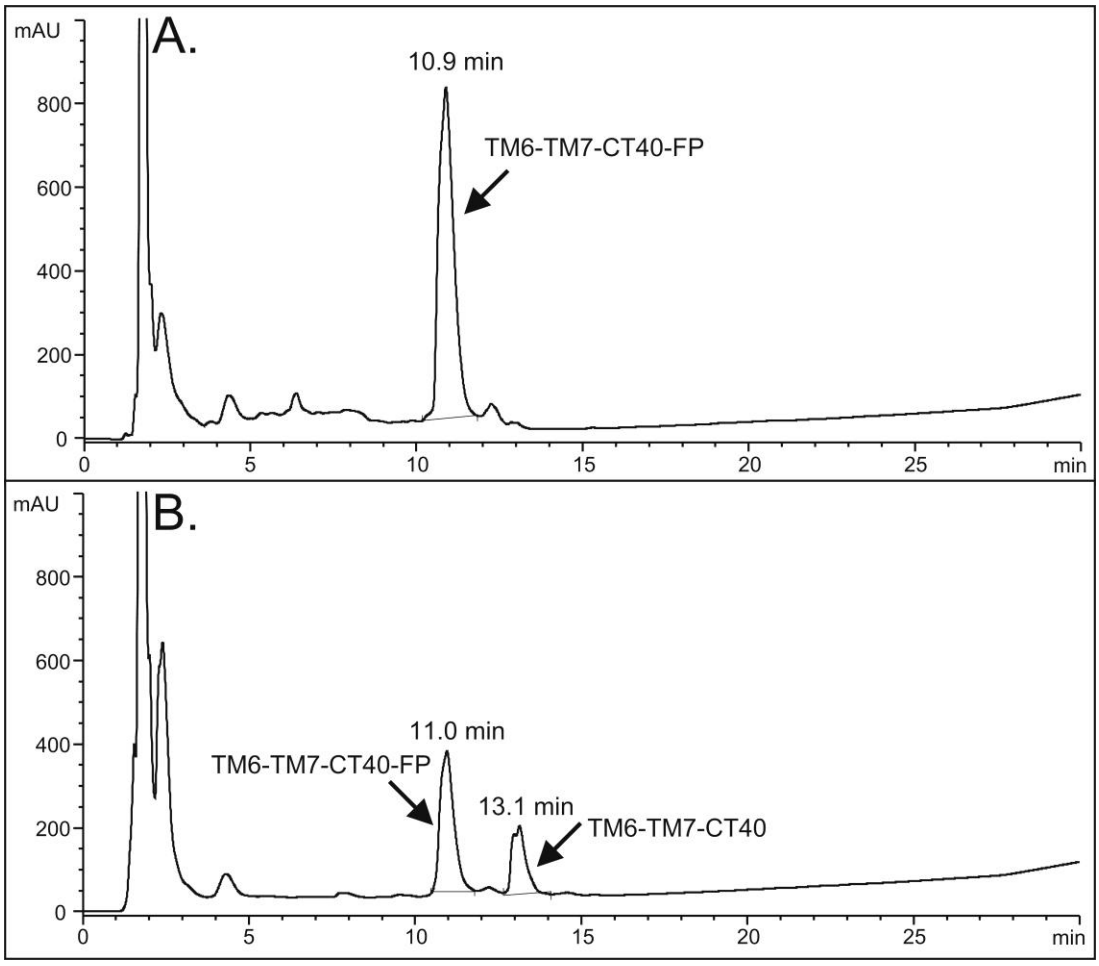


Figure 2-19. Cleavage of TM6-TM7-CT40-FP in inclusion bodies. A) IBs from a 250 ml culture of BL21-AI that expressed TM6-TM7-CT40-FP were solubilized in 70% TFA and were analyzed by RP-HPLC on a Zorbax 300SB-C3 column using a 36-72% acetonitrile:water gradient with 10% isopropanol and 0.1% TFA at 60°C. B) CNBr was added to the mixture and the reaction was checked after 1 hour. Incubations for longer times did not improve with longer incubation times. The reaction was analyzed by RP-HPLC as above.

Optimization of the cleavage reaction by increasing the amount of TFA in the reaction to 80% and the concentration of CNBr to 2M was attempted (Table 2-8). The cleavage in 80% TFA with 1M CNBr generated 0.32 mg of peptide after 2.5 h of cleavage, but there was also about 0.5 mg FP that was uncleaved. The reaction that contained 2M CNBr was purified after 2 hours and generated 0.38 mg peptide and left 0.33 mg FP uncleaved. Comparison of peak areas (Table 2-8) indicated that TM6-TM7-CT40 peptide may be aggregating. Purification conditions should be optimized to increase peptide amounts. For example, increasing the isopropanol in the solvent system to 20% may be helpful. However, initial biophysical analysis (see below) indicated that there were stability issues in NMR sample preparations (see Chapter 4) so we decided not to continue with the TM6-TM7-CT40 Ste2p fragment but rather to focus our biophysical analyses on TM1-TM2

.

Table 2-8. Comparison of CNBr cleavage conditions for TM6-TM7-CT40-FP

	1M CNBr				2M CNBr		
Time after CNBr added (h)	0	1	2	2.5	0	1	1.8
TM6-TM7-CT40-FP peak area	7681.5	2771.76	1641.56	1384.36	8566	1890.39	1203.19
TM6-TM7-CT40 peak area	0	1394.02	1467.97	1518.63		1418.01	1542.11
% FP remaining		36	21	18		22	14
% TM6-TM7-CT40 generated compared to expected peptide*		40	42	43		36	39

*This was determined by multiplying the starting FP peak area by 46% because TM6-TM7-CT40 was 46% of the FP by weight.

2d. Summary

In this chapter I have described the selection of the 2TM fragments for study, the cloning of these 2TM fragments, and optimization of their expression, cleavage and purification. The choice of fragment was very important in regards to the success of this project. The buried surface area analysis was expected to translate into interhelical connectivities that would be found in the structure analysis by NMR. Expression in multiple *E. coli* expression strains revealed that the highest yields of fusion protein occurred in BL21-AI and CNBr cleavage in the IBs increased the final yield of purified peptide. In fact, the yields of 6-10 mg per liter of culture that I obtained are significantly higher than those previously published for isotopically labeled polytopic membrane protein fragments. The following chapters will focus on biophysical analysis of these protein fragments including CD for secondary structure and NMR as a tool for structure determination.

Chapter 3

Initial Biophysical Analysis of 2TM constructs of Ste2p

3a. Introduction

After expression and purification of the 2TM peptide fragments, initial biophysical analyses such as circular dichroism (CD) and heteronuclear single quantum correlation (HSQC) NMR was used to assess the conformation and conformational stability of the peptides under various membrane mimetic conditions. Hydropathy analysis can be used to predict the amount of helical content in these transmembrane peptides. The information obtained from these computational analyses can be used in conjuction with experimental data to help ascertain more detailed structure information, specifically the assessment of the three dimensional structure of the peptides and the conformation assumed at specific residues.

Circular dichroism is an extremely useful tool in the analysis of secondary structure of a protein (*160, 161*). It is efficient in that very little peptide is needed for each experiment and that multiple conditions can be tested in a relatively short time. For helical peptides, it is a quick qualitative way to tell if the peptide is in the proper conformation. It is also useful in quantitative data analysis in determining the fraction of helical content (*162-164*), f_H, as well as differentiating between helical conformations and coiled-coils (*165*). The observations by CD are useful in determining overall secondary structure of the 2TM peptides, but the structure of individual residues or short stretches of amino acids could not be determined.

Many groups have started to use [^{15}N,^{1}H]-HSQC analysis as a screening tool for determining useful conditions to study TM peptides in 3D and 4D NMR experiments (*166-168*). This 2 dimensional (2D) NMR experiment is used to correlate the amide protons and the

nitrogen nuclei in the backbone and in some side chain residues. In comparison to the data obtained from CD, the information determined in these types of experiments relates to the conformational preferences of single amino acids and/or the structure of localized regions. The spectra observed in an HSQC experiment provide insights about conformation or lack thereof, based on peak dispersion, and on sample stability, based on the broadness and resolution of each peak. The primary sequence of the protein is used to predict the number of peaks one would expect to observe. For example, the first residue would not be observed because the exchange of the amide proton in solution would occur too quickly. Furthermore, any Pro residues would not be observed because they do not contain amide protons. In the case of TM1-TM2, I would expect to see 78 out of 80 residues and in the case of TM6-TM7-CT40 I would expect to see 106 out of 109 residues.

The disadvantage of NMR spectroscopy as an initial screening technique is the need for high concentrations of isotopically labeled samples. The experiments can also take much longer than the analysis by CD. On the other hand, CD requires much less peptide for initial analysis and the peptide does not have to be labeled. Thus we can perform initial analysis of these peptides by CD and the results we observe can be correlated to the results of the HSQC NMR analysis thereby saving time and money in the screening process.

3b. Materials and Methods[2]

Secondary Structure Prediction Analysis. The amino acid sequences for secondary structure analysis were analyzed using an online neural network from Heidelberg called PHD (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_phd.html) (*169*). The translocon analysis was performed on a downloaded program called Membrane Protein Explorer (MPex, http://blanco.biomol.uci.edu/mpex/) (*170*).

Micellar Domain Mixing Experiment and SDS-PAGE Analysis. TM1-TM2 and TM6-TM7-CT40 were solubilized in 25% TFE to a final concentration of 1 μg/μl. The peptides were portioned so that the final reactions contained 25 μM peptide. For example, TM1-TM2 and TM6-TM7-CT40 were each prepared to 25 μM to be analyzed alone, but when the peptides were mixed, 12.5 μM of each was prepared. The micelles were prepared in 20 mM phosphate buffer, pH 6.5, with either 10 or 100 mM SDS and sonicated for 15 minutes at 50ºC at power ~50W. The peptides and the micelles were mixed and sonicated at 50ºC for 5 min and then allowed to sit at room temperature for 30 minutes. The mixtures were sonicated again and either lyophilized or loading buffer without SDS (12% glycerol, 256 mM 2-mercaptoethanol, 50 mM Tris- HCl, pH 6.8, and a small amount of Brilliant Blue G) was added and a portion of the sample was loaded onto the gel. The samples were either loaded on SDS-PAGE gels prepared without a stacking and spacer gel (Figure 3-1A) or with a stacking and spacer gel (Figure 3-1B,C). The sample in LPPG was prepared in a similar way except that half of the samples were mixed with a final concentration of 0.1 mM α-factor.

[2] Parts of the information in this section was taken, with some modification, from (143).

Circular Dichroism Spectroscopy. For CD analysis of TM1-TM2 and TM6-TM7-CT40 in organic:aqueous solvents or micelles, stock solutions of these peptides were first prepared in 50% TFE:water. Concentration of the peptides was determined by UV spectroscopy at 280 nm using an extinction coefficient of 9570 $M^{-1}cm^{-1}$ for TM1-TM2 and 8230 $M^{-1}cm^{-1}$ for TM6-TM7-CT40 peptide (*171*). Appropriate volumes were lyophilized to yield 0.1 mg of peptide per vial. Solutions of TM1-TM2 and TM6-TM7-CT40 were solubilized in TFE and then water was added to the appropriate concentrations (25%, 50%, 95%; v:v). Blanks were prepared in a similar manner but without peptide. Solutions (20 mM) of detergents [SDS, dodecylphosphocholine (DPC), 1-palmitoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (LPPG), 1-myristoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (LMPG), 1-stearoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (LSPG), 1-palmitoyl-2-hydroxy-*sn*-glycero-3-phosphocholine (LPPC), 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC), dodecyl maltoside (DDM), and hexadecylphosphocholine (HDPC)] in 20 mM phosphate buffer, pH 5.6, were sonicated at 50ºC for 15 min using an S3000 unit (Misonix, Farmingdale, NY) equipped with a 2.5 inch Cup Horn sonicator operated at output power ~50 W. One half of the obtained solution was used as a blank for CD measurement, while the other half was used to solubilize the lyophilized peptide powder. The peptide solutions in detergents were sonicated as described above for 15 min or until the solution clarified. The detergent/lipid-to-peptide weight ratios were between 100:1 and 200:1.

The CD spectra of the peptides were recorded on either an AVIV model 62-DS or model 410 CD instrument (AVIV Associates, Lakewood, NJ). Quartz cuvettes with path length of 0.2 mm or 1 mm were used. All spectra were obtained by averaging 3–5 scans in a spectral window from 260 to 185 nm at an interval of 1 nm, slit width 1 or 2 nm, with a 5 s integration time at each wavelength. CD spectra on blanks, corresponding to the different media without peptide,

were collected at the same instrumental conditions and subtracted from the spectra containing the protein. CD intensities are expressed as mean residue ellipticities (deg cm^2 $dmol^{-1}$).

NMR Sample Preparation. Uniformly labeled [^{15}N]-TM1-TM2 and [^{15}N]-TM6-TM7-CT40 peptides (0.6–0.7 mg) were dissolved in 330 μl of either TFE-d_2:water (1:1 or 1:4), DMSO:water(±0.1%TFA) (1:1) or 60-500 mM detergent solution (d_{25}-SDS, d_{38}-DPC, LPPG, LPPC, LSPG, DPC, DDM, HDPC) in 20 mM phosphate buffer, pH 6.5, prepared in an H_2O/D_2O (9:1) mixture. Buffer solution also contained 0.02% NaN_3 and traces of 2,2-dimethyl-2-silapentane-5- sulfonate sodium salt (DSS) used to reference ^{1}H chemical shifts. The samples in DMSO were prepared as follows: 2.5 mg of [^{15}N]-TM1-TM2 was solubilized in 640 μL DMSO-d_6, vortexed and split into two tubes, 320 μL each. H_2O (80 μL) with or without 0.1% TFA was added to each tube and 3 μL of DSS was also added. ^{15}N chemical shifts were referenced indirectly (*172*). Final peptide concentration was ~0.2 mM and the peptide:detergent molar ratio was 1:500. Before dissolving the peptide, the detergent solution in buffer was sonicated for 10 min at 50ºC using a S3000 Misonix sonicator. The peptide/detergent mixtures were also sonicated for 10 min at 50ºC and then transferred to a Shigemi NMR tube (Shigemi, Allison Park, PA). Nitrogen-proton heteronuclear single quantum correlation ([^{15}N,^{1}H]-HSQC) NMR spectra were recorded at 45ºC on a four-channel Varian 600 MHz NMR spectrometer (Varian NMR Instrument, Palo Alto, CA) equipped with a z-axis pulsed-field-gradient and a Varian 5-mm [^{15}N,^{13}C,^{1}H] triple resonance probe.

3c. Results

Domain Mixing Analysis. Mutational analysis (*145*) and computer modeling (*8*) have indicated that, in the bundled Ste2p, TM1 and TM2 have the potential to interact with TM7. For example, the Konopka group predicted that T50 and A61 would have packing moments with TM7 and that N84 of TM2 would interact with S293 at TM7 (*8*). SDS-PAGE analysis has been used to identify TM domains that interact (*173-176*). Since I had purified TM1-TM2 and TM6-TM7-CT40 we attempted to look for domain interactions between the peptide fragments in SDS and LPPG. SDS was chosen due to its historical use as described above. On the other hand, LPPG was used due to the observation that TM1-TM2 exhibited significantly increased stability in the presence of LPPG micelles (see below). SDS-PAGE gels were run without stacking and spacer gels (Figure 3-1A) or with these gels (Figure 3-1B-D). The bands appeared to be sharper in the gels that contained stacking and spacer gels, so those conditions were used. A comparison was also made to determine if the samples should be run immediately (Figure 3-1A-B) or if they should be lyophilized before loading (Figure 3-1C). It appeared that the peptides behaved similarly under both conditions but the gel with the clearest bands was that of the lyophilized samples.

SDS-PAGE analysis has been useful in determining interactions between two peptides (*176*). The expected MW of TM1-TM2 is 8.7 kDa and of TM6-TM7-CT40 is 11.7 kDa. Homodimers would have MWs of 17.4 or 23.4 kDa, respectively, and the presence of a heterodimer would be indicated by a band with a MW of 20.4 kDa. In both concentrations of SDS, the TM1-TM2 peptide appeared to run as a monomer as judged by the presence of a band at about 12.5 kDa, the TM6-TM7-CT40 appeared as mostly the dimeric form as judged by the presence of a band at about 24.4 kDa with a smear below ranging from about 16.5 kDa to the

band at 24.4 kDa. The MWs were calculated based on relative mobility analysis (*177*). When mixed, these peptides did not appear to interact with each other under the conditions tested as observed by the presence of two bands which both correspond to the positions of the monomeric peptides. The observed MW increase for TM1-TM2 in all of these experiments may be explained by the interaction of the SDS with the TM peptide as described by the Deber group (*178*). Further testing using the LPPG micelles led to very strange results (Figure 3-1D). The samples were prepared as described above with a 20 mM LPPG stock solution instead of SDS, but it appears that the LPPG micelles do not run very well on SDS-PAGE gels probably due to their large size (Figure 3-1D, lanes 2-4). Furthermore, when α-factor was added to the LPPG-peptide mixtures to 0.1 mM the solution became extremely cloudy and the gel analysis was even worse (Figure 3-1D, lanes 5-7). Sonication and lyophilization did not help the α-factor mixture go into solution. Interactions between the α-factor and the detergent may have disrupted the micelle thereby causing the 2TM peptide fragments to aggregate. Overall, under the conditions analyzed, it appears that TM1-TM2 and TM6-TM7-CT40 do not interact in SDS.

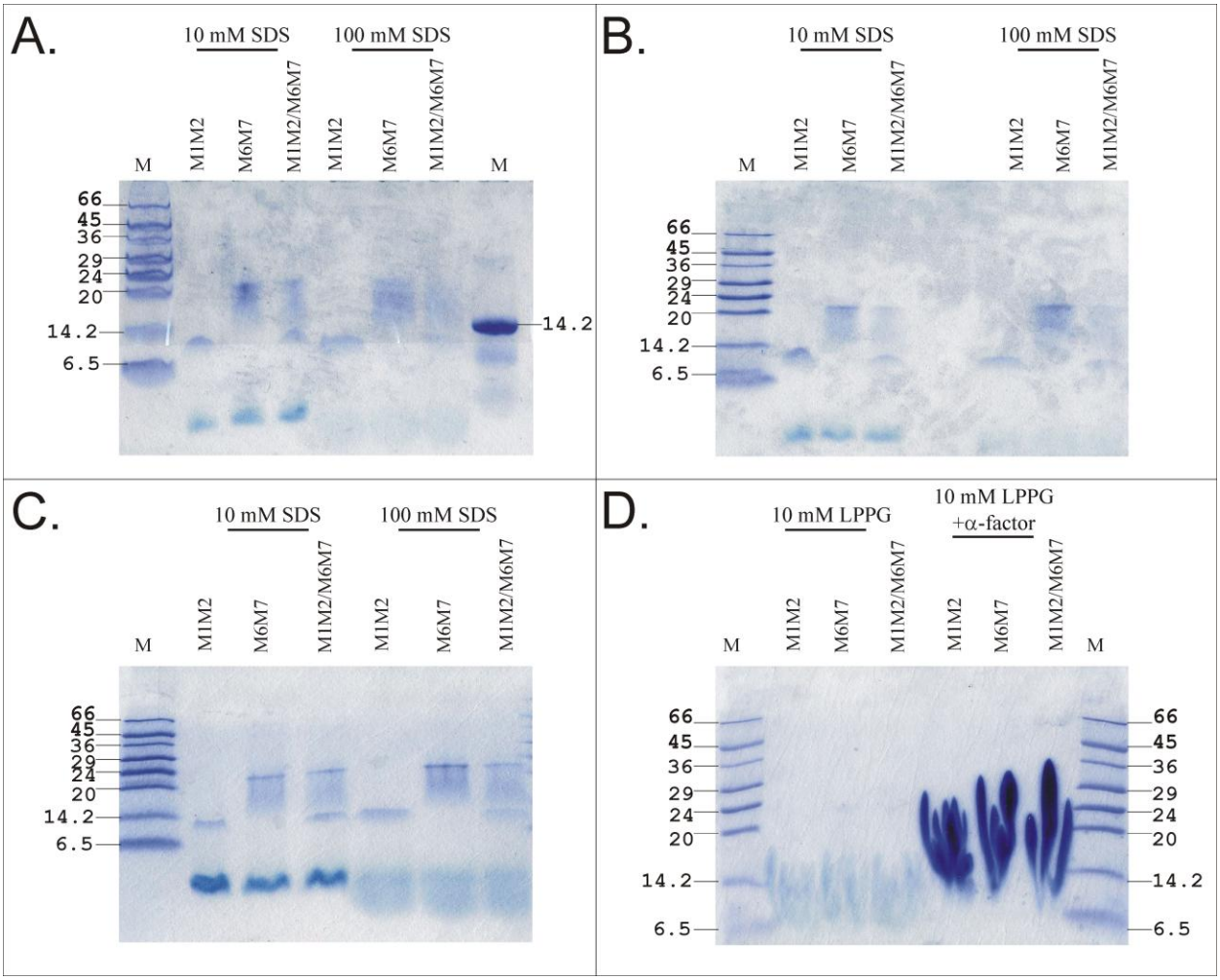


Figure 3-1. SDS-PAGE analysis of domain interactions. TM1-TM2 or TM6-TM7-CT40 peptides were prepared in either 10 mM or 100 mM SDS to 25 μM final concentration. The peptide samples were mixed with loading buffer and analyzed by SDS-PAGE. A) SDS-PAGE gel prepared without stacking or spacer gels. B) SDS-PAGE gel prepared with stacking and spacer gels. C) Samples lyophilized after mixing and incubation then resuspended in loading buffer and run on an SDS-PAGE gel prepared with stacking and spacer gels. D) SDS-PAGE analysis of TM1-TM2 and TM6-TM7-CT40 in LPPG micelles. The MWs of the individual peptides are as follows: TM1-TM2 = 8.7 kDa and TM6-TM7-CT40 = 11.7 kDa. Homodimers would have MWs of 17.4 or 23.4 kDa, respectively. The presence of a heterodimer would be indicated by a band with a MWof 20.4 kDa. To save space in the Figure the peptides are indicated as M1M2 or M6M7.

Secondary structure prediction analysis. Using computer-based secondary structure prediction tools, one can determine the predicted percent helicity of the peptides to be studied. These predictions can also be compared to the biochemical and modeling data by the Dumont and Konopka groups (*8, 123, 145*). To a large extent the predictions are based on hydropathy analysis and the transmembrane regions are predicted to be helical. The rhodopsin-templated model (*8*) predicted that, in the context of the full receptor, TM1 would extend from residues 46-72 and that TM2 would be from 81-105 (Table 3-1). Assuming that all residues not in the TM domain are random, the predicted helicity of TM1-TM2 would be about 65% (Table 3-1). Using a computer neural network analysis program Protein network from HeiDelberg (PHD), the predicted helicity of the TM1-TM2 peptide was 60% α-helical with 11.25% extended strand and 28.75% random coil (Figure 3-2A, Table 3-1). This prediction program is mainly used to predict hydrogen bonds between residues. The helical content predicted by the PHD analysis of this peptide sequence includes a predicted α-helix at the N-terminal extracellular tail (residues 39-51), which had not been predicted in the Ste2p modeling analysis. Other differences between the templated model and the PHD analysis are the TM helix boundaries. PHD defines TM1 to be from 54 to 74 which is 6 residues less than the templated model and extends into the intracellular loop 1 (IL1). PHD predicted TM2 to encompass residues 90 to 103. This TM domain would be only 14 residues and therefore too short to cross the membrane. The placement of this helix is about 11 residues further downstream from the first residue of helix 2 in the Eilers' model (*8*). A second computational method, Membrane Protein Explorer (MPex) (*170*), was used to analyze this peptide fragment (Figure 3-2B, Table 3-1). In this program, one can use a translocon prediction algorithm which predicts TM1 to contain residues 52-74 which start near the N-terminus of TM1 and continues into IL1. MPex predicted TM2 to be from residues 80 to 102

which is very similar to that predicted in the templated model. Overall the predictions of the MPex translocon analysis appear to more closely match the TMs of the templated model though the MPex predicted helicity (58%) was lower than that of the templated model (Table 3-1). One drawback to using the translocon analysis algorithm is that it specifically searches for TM helices, so the N-terminal helix would not be predicted *a priori*.

Analysis of TM6-TM7-CT40 from the templated model indicated that, in the context of the full receptor, TM6 would extend from residues 245 to 269 and TM7 would include residues 274 to 300 (Table 3-1). NMR analysis of a peptide containing TM7 and 40 residues of the C-terminal tail in both TFE:water and DPC micelles has shown the presence of a short helix (Helix 8) that is similar to the one found in rhodopsin and would increase the helicity values slightly (*115, 117*). When determining helical content of the templated model, I did not include this C-terminal tail and I assumed that the non-TM residues were random which resulted in a percent helicity of 48% (Table 3-1). Hydropathy analysis of TM6-TM7-CT40 by PHD predicts 29.36% α-helicity, 21.1% extended strand and 49.54% random coil (Figure 3-3A, Table 3-1). Though a decrease in helicity is not unexpected due to the disordered nature of the CT tail (*115, 117*)(Bhuiyan, unpublished data), I believe that the PHD prediction of this peptide is incorrect. The PHD algorithm predicts that TM6 does not begin until residue 259, which is in the middle of this TM domain, and that this helix is only 8 residues long. A helix of this length would be too short to cross a membrane. Conversely, PHD predicts a longer helix than the templated model by predicting that TM7 encompasses residues 276-302. One prediction by PHD that appears to be biologically relevant is a "break" in TM7 at a Pro residue. The presence of a Pro residue here, in TM6, and in many TM helices of almost all GPCRs is indicative of the inherent flexibility of these signal transducing proteins. The translocon TM analysis by MPex assigned TM6 between

residues 247-269 and TM7 between 277-299 which would indicate about 42% helicity (Figure 3-3B, Table 3-1). As with TM1-TM2, the translocon analysis appears to match the modeled Ste2p helix predictions.

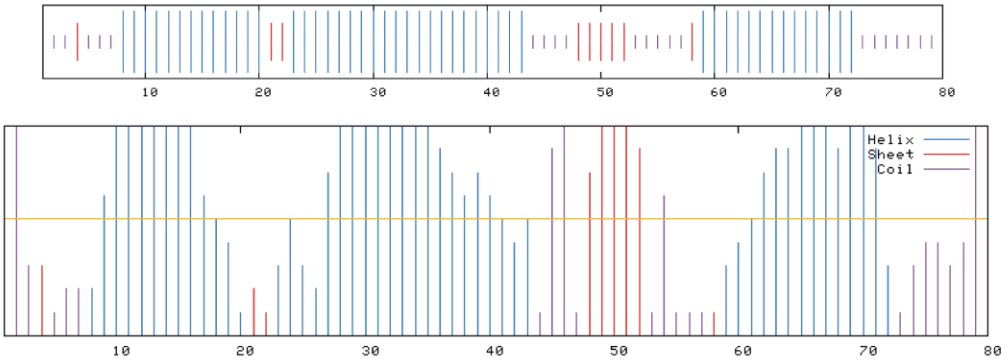
Table 3-1. Transmembrane helix assignments and percent helicity of 2TM fragments.

				% helicity
TM1-TM2	Nterminal helix	TM1	TM2	
Templated model	n/p[a]	46-72	81-105	65%
PHD	39-51	54-74	90-103	60%
MPex	n/p	52-74	80-102	58%
TM6-TM7-CT40	TM6	TM7	Cterminal helix	
Templated model	245-269	274-300	n/p	48%
PHD	259-266	276-302	n/p	29%
MPex	247-269	277-299	n/p	42%

[a]n/p not predicted

A.

```
        40        50        60        70        80        90       100       110
        |         |         |         |         |         |         |         |
GNGSTITFDELQGLVNSTVTQAILFGVRSGAAALTLIVVWITSRSRKTPIFIINQVSLFLIILHSALYFKYLLSNYSSVT
CCCceccchHHHHHHHHHHhheehHhhHHHHHHHHHHHHHHHHhHcCCcEEEEEcCcccehhHHHHHHHHHHHhccccccC
```



Residues with a scale reliability index of prediction of 5 and over (uppercase letters) are predicted at better than 82%.

B.

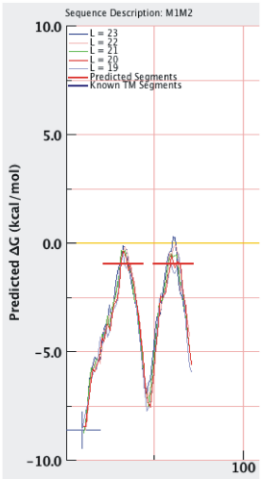
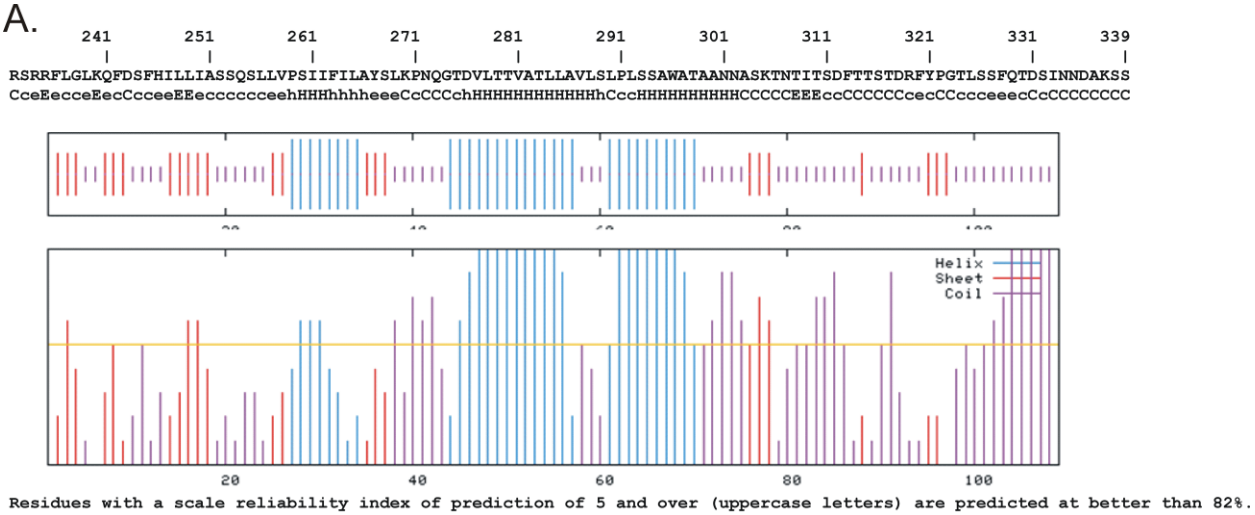


Figure 3-2. Hydropathy analysis of TM1-TM2 by PHD and MPEx. A) Secondary structure prediction using Profile network from HeiDelberg (PHD) (*169*). The sequence of the 80-residue TM1-TM2 peptide is at the top and the predicted secondary structure is on the line underneath it. Cc: random coil; Hh: α-helix; Ee: extended strand. The top box is a visualization of the predictions above it and the bottom box is an indicator of reliability scores. B) Membrane Protein Explorer (MPEx) analysis Translocon TM output indicated two TM domains by red lines at a ΔG of about -1.

A.



Residues with a scale reliability index of prediction of 5 and over (uppercase letters) are predicted at better than 82%.

B.

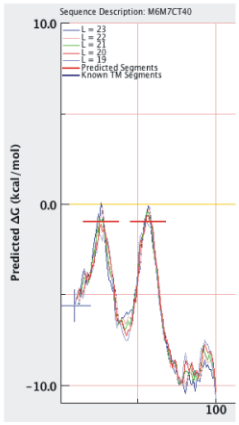


Figure 3-3. Hydropathy analysis of TM6-TM7-CT40 by PHD and MPEx. A) Secondary structure prediction using Profile network from HeiDelberg (PHD) (*169*). The sequence of the 109-residue TM6-TM7-CT40 peptide is at the top and the predicted secondary structure is on the line underneath it. Cc: random coil; Hh: α-helix; Ee: extended strand. The top box is a visualization of the predictions above it and the bottom box is an indicator of reliability scores. B) Membrane Protein Explorer (MPEx) analysis Translocon TM output indicated two TM domains by red lines at a ΔG of about -1.

Far UV CD analysis in organic aqueous mixtures. The 2TM peptide fragments were analyzed in TFE:water ranging from 95% to 25% TFE (v:v). This organic:aqueous mixture was used as it is often considered to be a membrane mimetic (Figure 3-4) (*179*). It is known that TFE often induces polypeptides to form α-helices (*180, 181*). The presence of these are indicated in the CD spectrum by a maximum at 192 nm and a minimum at 208 nm which are due to exciton splitting of the $\pi \rightarrow \pi^*$ transition as well as a minimum at 222 nm which reflects the $n \rightarrow \pi^*$ transition. TM1-TM2 in TFE:water ranging from 25-95% exhibits mean residue molar ellipticities at 222 nm ranging from -19700 to -23800 deg $cm^2 dmol^{-1}$ which suggest the formation of highly helical peptides (Figure 3-4A, Table 3-2). This data can also be used to determining the f_H, or fraction

Table 3-2. Fraction helicity of 2TM peptides in organic:aqueous media.

	Mean Residue Ellipticities (deg $cm^2 dmol^{-1}$)		Helicity Calculations	
	208 nm	222 nm	f_H[a]	222:208[b]
TM1-TM2				
25% TFE	-21700	-19700	0.57	0.91
50% TFE	-23000	-20500	0.60	0.89
95% TFE	-26400	-23800	0.71	0.90
TM6-TM7-CT40				
25% TFE	-960	-3300	0.03	3.4
50% TFE	-12900	-10900	0.28	0.84
95% TFE	-14400	-12700	0.34	0.88

[a] The fraction helicity (f_H) is determined by the equation $f_H=([\theta]_{222}+2340)/-30300$ as described in (*164*).
[b] The ratio of mean residue elliptcities at 222 and 208 nm can be used to distinguish between helices and coiled coils (*165*).

helicity, and it can be analyzed in terms of the ratio of 222 nm to 208 nm to discriminate between monomeric helices and coiled-coils (Table 3-2). This latter point is difficult to ascertain from the graph alone. The highest f_H calculated for TM1-TM2 was in 95% TFE:water (0.71). The percent

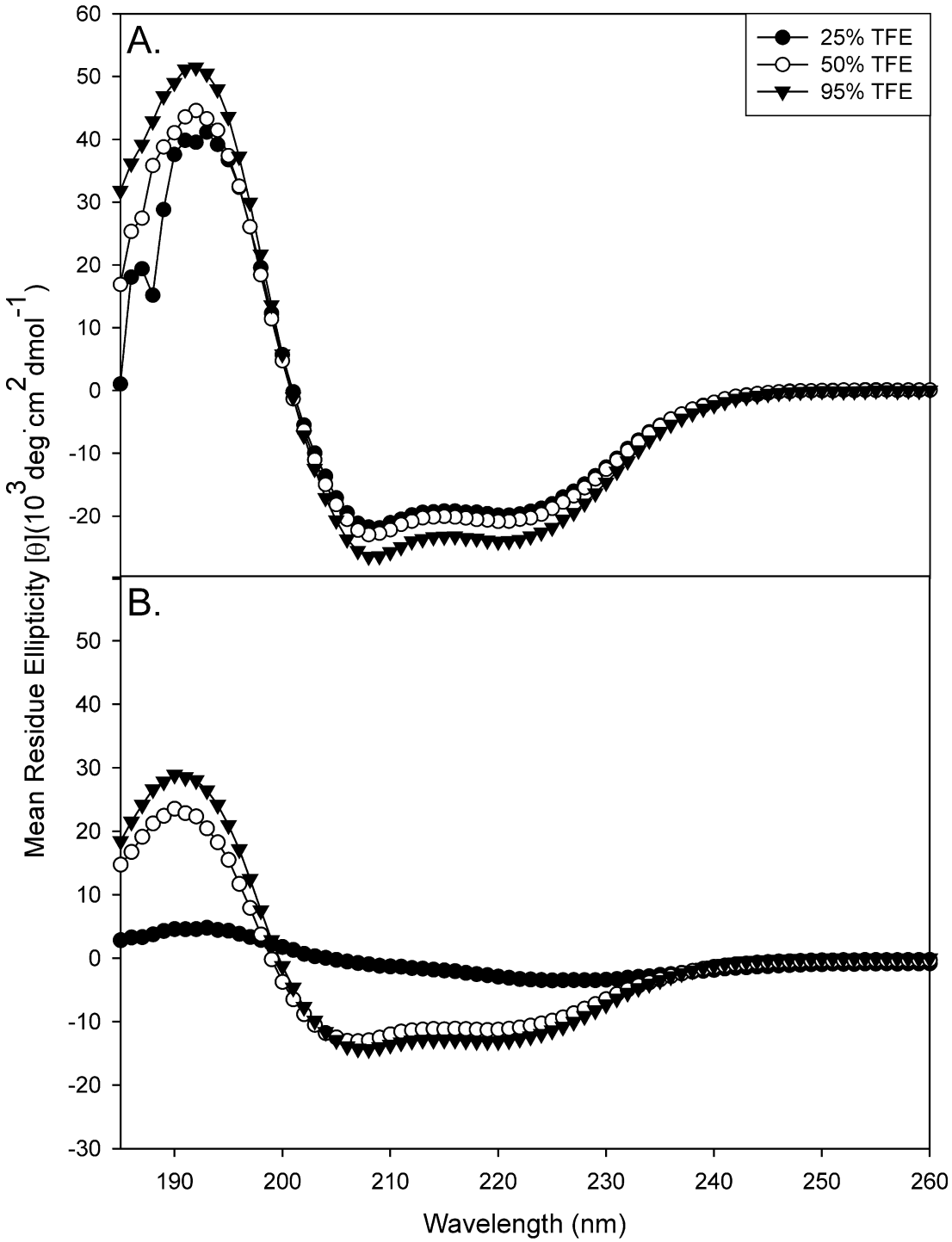


Figure 3-4. Circular dichroism of 2TM fragments in TFE:water. Purified 2TM peptides TM1-TM2 (A) and TM6-TM7-CT40 (B) were solubilized to 25 μM final concentration in TFE:water: 25% TFE:75% water(0.1%TFA), 50% TFE:50% water(0.1%TFA), or 95% TFE:5% water(0.1%TFA) as indicated in the legend at the top right corner. Analysis by CD was performed using a 1 mm cuvette at 25°C as described in the Methods.

helicities of TM1-TM2 found in 50% and 25% TFE were about 60%. The ratio of the 222 nm and 208 nm bands found in all cases (~0.9) indicates that the helical content of this peptide is mostly monomeric as opposed to being a coiled-coil, which should exhibit a ratio of ~1.2 (*165*). HSQC analysis of [^{15}N]-TM1-TM2 will be performed in 25% and 50% TFE because they appear to have similar secondary structure characteristics and because the percent helicity is consistent with those predicted by our hydropathy and modeling analyses (see above).

CD analysis of TM6-TM7-CT40 in TFE:water revealed significantly lower molar ellipticity values than found for TM1-TM2 under the same conditions, -3300 at 25% TFE to -12700 deg cm^{2}dmol^{-1} at 95% TFE (Figure 3-4B, Table 3-1). TM6-TM7-CT40 had the highest helicity in 95% TFE:water, but the CD curve in 50% TFE/50% water was very similar; the decrease in helicity was small as observed by f_H values of around 0.3 (Table 3-2). These experimental f_H values are similar to the percent helicity as determined by the hydropathy analysis described above. In 25% TFE:75% water the peptide appeared to lose nearly all of its helical conformation. Though the mean residue molar ellipticity of TM6-TM7-CT40 in 50% TFE was lower than that of TM1-TM2, the reduction in helicity can be explained by the contribution to the optical activity from the long and mostly disordered C-terminal tail (*115, 117*) which would decrease the helical content.

Far UV CD analysis of TM1-TM2 in micellar environments. Transmembrane peptides may also form α-helical structures in the presence of detergent micelles which are thought to more closely mimic the plasma membrane than an organic:aqueous environment (*167*). In order to examine the secondary structure of the 2TM peptides a large molar excess of detergent (1000-fold) was used to create micelles. This large excess ensures that only one polypeptide is present

in each micelle and should minimize aggregation. TM1-TM2 exhibited minima at 208 nm and 222 nm and a split $\pi \rightarrow \pi^*$ transition in the presence of a variety of detergents (Figure 3-5). On the basis of the mean residue ellipticities TM1-TM2 is highly helical (>50%) in LSPG, LPPG

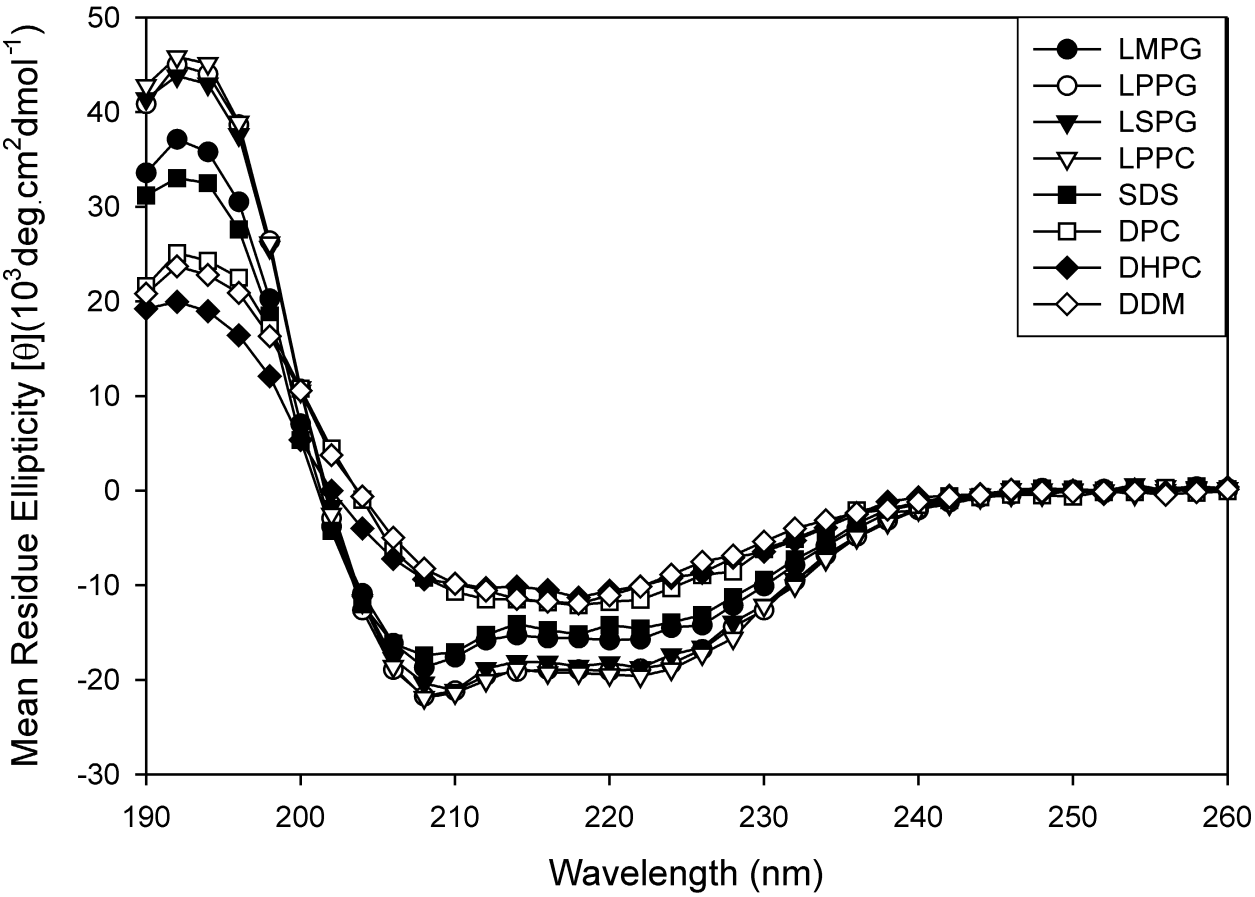


Figure 3-5. Circular dichroism of TM1-TM2 in micellar environments. Purified TM1-TM2 was solubilized in micelles prepared in 20 mM detergent solutions of LMPG (1-myristoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol)), LPPG (1-palmitoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol)), LSPG (1-stearoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol)), LPPC (1-palmitoyl-2-hydroxy-*sn*-glycero-3-phosphocholine), SDS (sodium dodecyl sulfate), DPC (dodecylphosphocholine), DHPC (1,2-dihexanoyl-*sn*-glycero-3-phosphocholine), and DDM (dodecyl maltoside) in 20 mM sodium phosphate buffer pH 5.6 as described in the Methods. The final peptide concentration in all detergent solutions was 25 μM. The symbols used are indicated at the top of the graph.

and LPPC micelles (Figure 3-5, Table 3-3). The TM peptide is also helical in LMPG and SDS though the amount of helicity has decreased to about 41-44%. Conversely, in the presence of DPC, DHPC or DDM micelles, TM1-TM2 did not have the distinct minima that are signatures of an α-helical structure (Figure 3-5), had reduced mean residue ellipticities and the f_H has decreased significantly ≤30% (Table 3-3). Upon further analysis, it appears that the percent helicity can be grouped based on the acyl chain length of the detergent. The detergent-peptide complexes with mean residue ellipticity values at 222 nm more negative than -18000 (deg cm^2dmol^{-1}) contain detergents with acyl chains containing 16-18 carbons, followed by detergents with 12-14 carbons which resulted in CD spectra with mean residue molar ellipticity values at 222 nm between -14000 and -16000 (Table 3-3). The detergent with the 12 C acyl chains had a smaller head group. The group of detergents that resulted in the lowest peptide helicity contains 6-12 carbon atoms in the fatty acyl group and larger head groups (Table 3-3). Though this data appears to indicate that acyl chain length of the detergent plays a major role in the helicity exhibited by TM1-TM2, it is not the only factor that affects helical content. The size of the head group may work in conjunction with the acyl chain length to affect the peptide-micelle interaction. Specifically, comparison of SDS and DPC or DDM, where the acyl chain for both is 12 carbons long, shows a marked difference in the percent helicity of TM1-TM2 (Table 3-3). SDS has a sulfate head group which is smaller than the phosphocholine head group of DPC or the maltose head group of DDM and this changes the cone shape of the detergent and therefore the shape of the micelle which apparently influences peptide integration into the micelle and its helicity. Similar affects have been observed by small-angle X-ray scattering, NMR and CD (*182-184*). Interactions between the peptide and the headgroup could also play a role in

determining peptide-detergent stability (*167*). Thus for TM1-TM2 a critical chain length occurs at approximately 12 carbon atoms.

It has been observed in some cases that the use of salt in a micellar environment stabilizes the detergent-peptide complex (*86, 185, 186*). For the same set of detergents analyzed above, the micelles were prepared in 20 mM sodium phosphate pH 5.6 containing 100 mM NaCl (Figure 3-6). The concentration of NaCl was kept below 150 mM to ensure that there was minimal interference with the CD analysis because the chloride ions can absorb strongly below 200 nm (*161*). In the presence of 100 mM NaCl the micellar environments in which the TM1-TM2 peptides exhibited the greatest ellipticity and well defined 208 nm and 222 nm minima were LPPG and LSPG (Figure 3-6, Table 3-3). LPPC micelles in the presence of salt exhibited a decreased the mean residue molar ellipticity for TM1-TM2 (below -16000 deg cm^2dmol^{-1}) so that it is now grouped with the second group which included LMPG and SDS. The mean residue molar ellipticity of the LMPG also decreased to -13000 deg cm^2dmol^{-1}, but it still is in the middle group. The largest salt effects were observed with LMPG and LPPC; the helical content as judged by f_H was significantly decreased (Table 3-3). Conversely, the f_H of TM1-TM2 in SDS appeared to increase slightly in the presence of NaCl. The helicity of the TM1-TM2 peptide in the presence of salt and DPC, DHPC or DDM micelles was dramatically lower than that of the other groups with the DDM micelles showing the largest salt effect.

Under the micellar conditions tested, the helical content of TM1-TM2 appeared to be closest to the predicted helicities (Table 3-1) in LPPG and LSPG both with and without salt. Accordingly these media would be good candidates for NMR analysis. I also decided to examine SDS (moderate helicity) and DPC (low helicity) in order to evaluate the use of CD as a screening tool in comparison with HSQC analysis. One note that should be mentioned is that during the

analysis of the data I observed that the f_H based on the molar ellipticity at 222 nm can vary by about 0.05. This variation was taken into account during the analyses presented in this chapter. Ultimately I used CD as a qualitative indicator of sample helicity and the appropriateness of the conditions for further analysis.

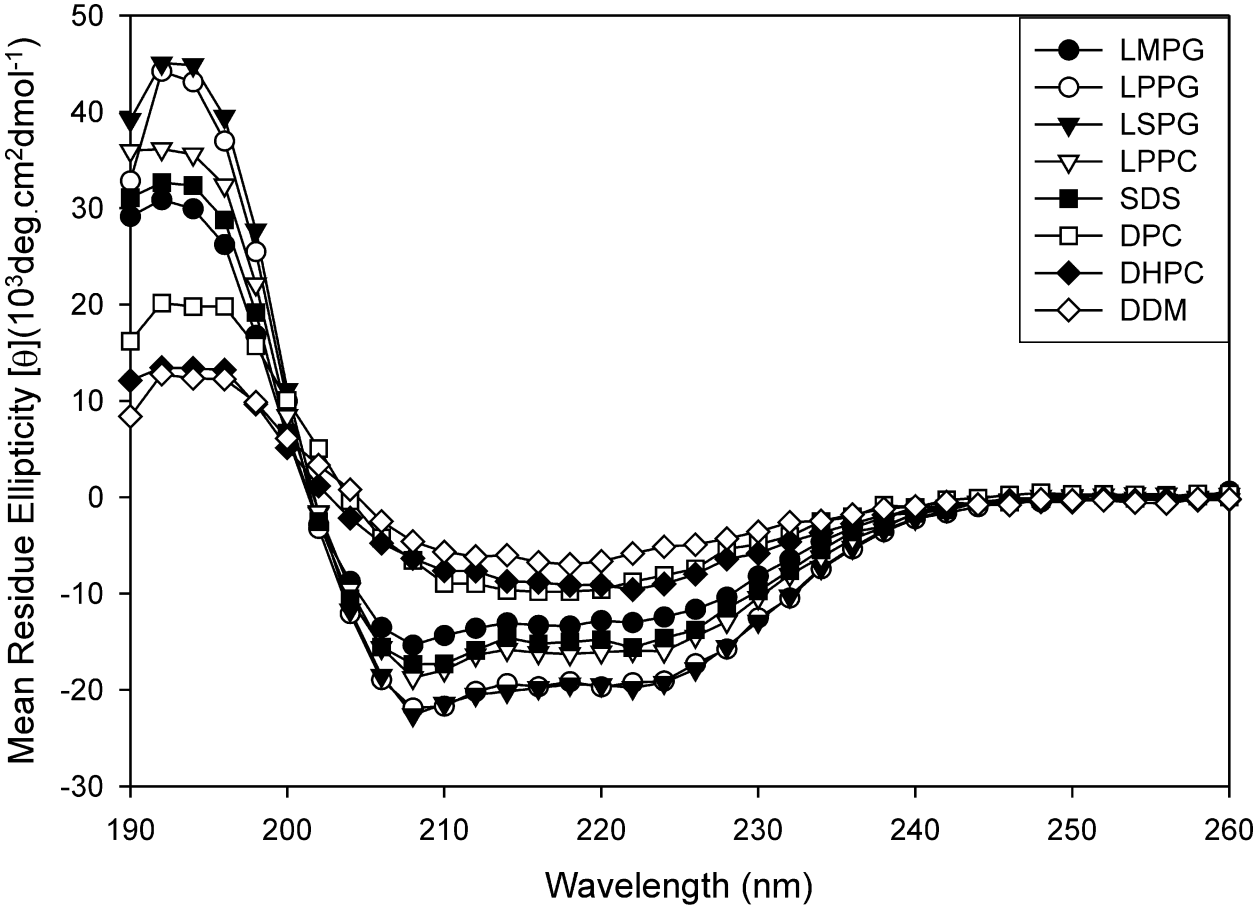


Figure 3-6. Circular dichroism of TM1-TM2 in micelles with 100 mM NaCl. Purified TM1-TM2 was solubilized to 25 μM final concentration in micelles prepared in detergent solutions containing 100 mM NaCl in 20 mM sodium phosphate buffer pH 5.6 with 20 mM LMPG, LPPG, LSPG, LPPC, SDS, DPC, DHPC, and DDM as described in the Methods. The symbols used are indicated at the top of the graph.

Table 3-3. Fraction helicity of TM1-TM2 in a micellar environment.

	Mean Residue Ellipticities[a] (deg cm^2dmol^{-1})		Helicity Calculations		Mean Residue Ellipticities (NaCl)[b] (deg cm^2dmol^{-1})		Helicity Calculations (NaCl)[b]		Detergent Information	
	208 nm	222nm	f_H	222:208	208 nm	222 nm	f_H	222:208	Acyl chain length	Head Group
LMPG	-18700	-15700	0.44	0.84	-15400	-13000	0.35	0.84	14	Phosphoglycerol
LPPG	-21700	-18800	0.54	0.87	-21900	-19300	0.56	0.88	16	Phosphoglycerol
LSPG	-20400	-18700	0.54	0.92	-22600	-19800	0.58	0.88	18	Phosphoglycerol
LPPC	-21900	-19600	0.57	0.89	-18700	-15900	0.45	0.85	16	Phosphocholine
SDS	-17400	-14600	0.41	0.84	-17300	-15600	0.44	0.90	12	Sulfate
DPC	-9200	-11600	0.31	1.26	-6500	-8800	0.21	1.35	12	Phosphocholine
DHPC	-9400	-10100	0.26	1.07	-6300	-9600	0.24	1.52	6	Phosphocholine
DDM	-8200	-10200	0.26	1.23	-4600	-5800	0.11	1.26	12	Maltoside

[a]All detergents were in 20 mM sodium phosphate pH 5.6.
[b]All detergents were in 100 mM NaCl, 20 mM sodium phosphate pH 5.6.

Far UV CD analysis of TM1-TM2 in mixed micellar environments. The micelles described above are generally good membrane mimetics, but the plasma membrane contains a variety of different lipids and head groups (*187, 188*). For example, the yeast plasma membrane is mostly composed of 16 to 18 carbon acyl chain fatty acids, glycerides and phospholipids that are saturated or contain one double bond and the headgroups are mostly either zwitterionic or negatively charged (phosphatidylethanolamine, phosphatidylinositiol, phosphatidylserine, and phosphatidylcholine) (*187*). Several groups have begun to use mixed micelles as a tool to increase stability of a peptide/protein or to better mimic the native membrane environment (*168, 182, 189-192*). In some cases, the use of 2 different detergents has been shown to "rescue" the secondary structure or initial [^{15}N,^{1}H]-HSQC of a peptide/protein (*168*). It has been shown by using paramagnetically labeled detergents in a mixed micelle system that the second detergent is in fact incorporated and that the nature of the detergent affects the protein in the mixed micelle environment (*193*). TM1-TM2 was analyzed in mixed micelles in the manner that was described above. Initial analysis was performed using three different detergents that contained two different headgroups (Figure 3-7). LPPG was mixed with either HDPC (a 16C acyl chain with a phosphocholine headgroup) or with DPC (a 12C acyl chain with a phosphocholine headgroup). The helicities of TM1-TM2 in LPPG and HDPC micelles were similar, mean residue ellipticity at 222 nm -17100 and -18100 deg cm^{2}dmol^{-1}, respectively, indicating that the main characteristic of the detergent necessary for helical formation is the length of the acyl chain (Figure 3-7A, Table 3-4). When mixed in varying ratios, f_H values were all approximately 50% or greater. The mixed micelle with the highest helical content was LPPG:HDPC mixed in a 4:1 ratio (f_H=0.6) (Table 3-4) which appeared to be slightly greater than either LPPG (f_H=0.49) or HDPC (f_H=0.52) alone. As stated above, the error in the f_H is about 0.05. Nevertheless, the results indicate that the

head groups may also play a role in determining the TM1-TM2 conformation because the acyl chains in both HDPC and LPPG contained 16C.

When the same analysis was performed on mixed micelles containing LPPG and DPC, a "rescue" of the helical secondary structure was observed (Figure 3-7B). As judged from the CD spectrum, which was a broad featureless negative band from 230 to 210 nm with a maximum mean residue ellipticity at 222 nm of -11200 deg-cm^2-dmol^{-1}, for TM1-TM2 in DPC micelles, this peptide contains a low (29%) helical conformation. However, when DPC is mixed with LPPG at ratios from 1:4 to 4:1, the helicity increases to ~45%. However, in no case was the helicity of TM1-TM2 in LPPG/DPC mixtures as high as that of LPPG:HDPC (4:1). In fact, the helicity of this peptide in all of the LPPG:DPC mixed micelle samples is less than that in LPPG or HDPC alone (Figure 3-7, Table 3-4).

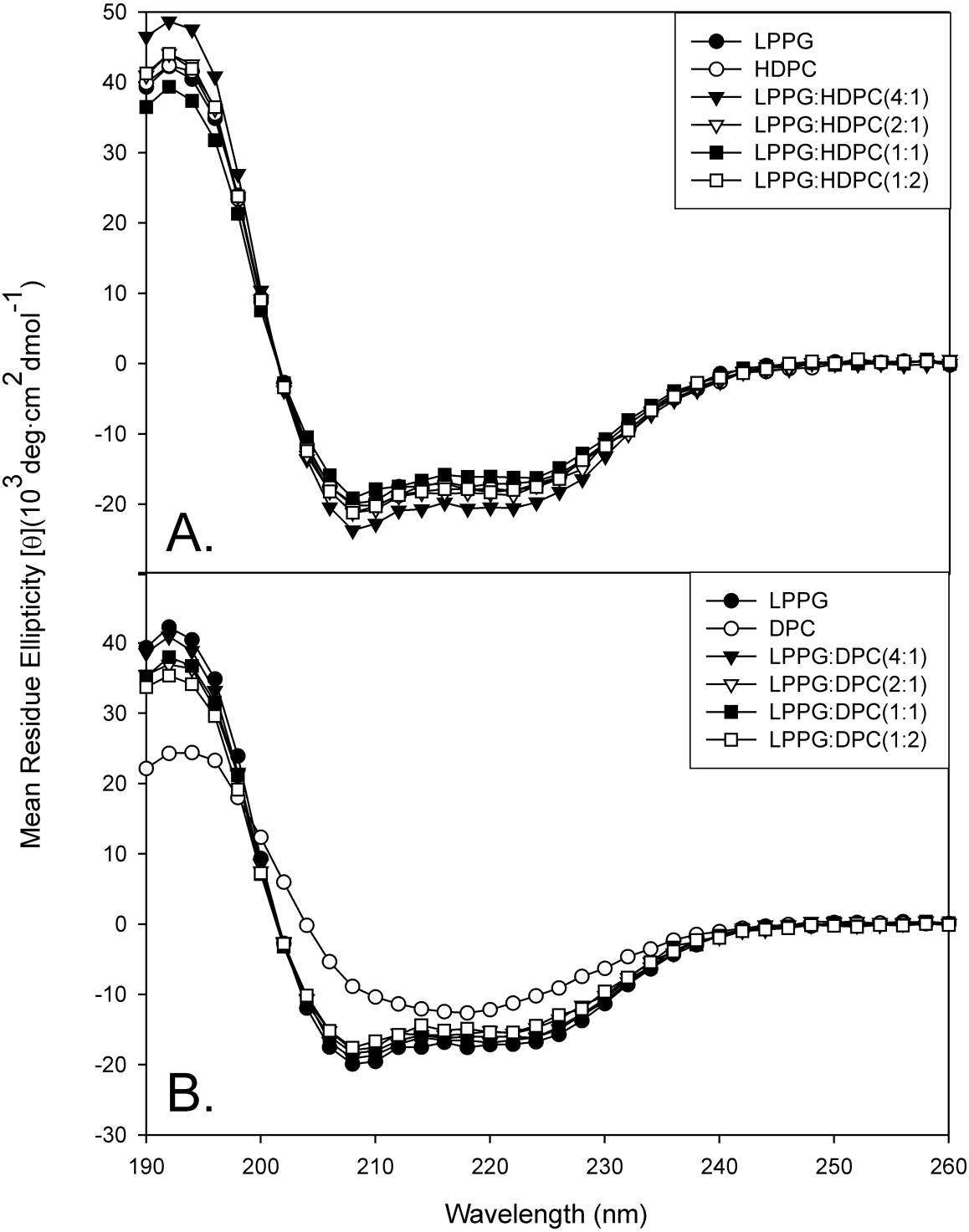


Figure 3-7. Circular dichroism analysis of TM1-TM2 in mixed detergent micelles. . A) Analysis of 25 μM TM1-TM2 in each detergent alone (LPPG, HDPC (heptadecylphosphocholine)) and in the presence of mixed detergents in varying ratios as described on the graph. B). Analysis of 25 μM TM1-TM2 in LPPG or DPC alone or in the presence of mixed detergents in varying ratios as described on the graph. All peptide-micelle complexes were prepared in 20 mM sodium phosphate buffer, pH 5.6.

Table 3-4. Fraction helicity of TM1-TM2 in mixed micellar environments.

	Mean Residue Ellipticities (deg $cm^2 dmol^{-1}$)		Helicity Calculations		Detergent Information	
	208 nm	222 nm	f_H	222:208	Acyl chain length	Head Group
LPPG	-19900	-17100	0.49	0.86	16	Phosphoglycerol
HDPC	-20200	-18100	0.52	0.90	16	Phosphocholine
DPC	-8900	-11200	0.29	1.26	12	Phosphocholine
LPPG:HDPC (4:1)	-23700	-20600	0.60	0.87	16/16	Phosphoglycerol:choline
LPPG:HDPC (2:1)	-21200	-18800	0.54	0.89	16/16	Phosphoglycerol:choline
LPPG:HDPC (1:1)	-19200	-16200	0.46	0.84	16/16	Phosphoglycerol:choline
LPPG:HDPC (1:2)	-21200	-18000	0.52	0.85	16/16	Phosphoglycerol:choline
LPPG:DPC (4:1)	-19100	-16500	0.47	0.86	16/12	Phosphoglycerol:choline
LPPG:DPC (2:1)	-17900	-15600	0.44	0.87	16/12	Phosphoglycerol:choline
LPPG:DPC (1:1)	-18400	-16000	0.45	0.87	16/12	Phosphoglycerol:choline
LPPG:DPC (1:2)	-17600	-15400	0.43	0.88	16/12	Phosphoglycerol:choline
LSPG	-20100	-17300	0.49	0.86	18	Phosphoglycerol
DDM	-5000	-6600	0.14	1.32	12	Maltoside
HDPC	-18800	-16800	0.48	0.89	16	Phosphocholine
DPC	-8700	-10200	0.26	1.17	12	Phosphocholine
LSPG:DDM (3.7:1)	-20500	-18100	0.52	0.88	18/12	Phosphoglycerol:maltoside
LSPG:HDPC (3.8:1)	-21000	-18000	0.52	0.86	18/16	Phosphoglycerol:choline
LSPG:DPC (6.3:1)	-17500	-15600	0.44	0.89	18/12	Phosphoglycerol:choline

As my analysis was being performed, a paper was published in JACS which related the size of the micelle to the stability of the protein-micelle complex (*192*). For each protein tested, micelles composed of different detergents could have similar influence on protein structure provided that the size of the micelle was kept constant. Based on sample stability during NMR analysis (see below) we concluded that the optimal micelle size is that of the LPPG micelle-peptide complex for TM1-TM2. In order to make mixed micelles based on the JACS paper I had to increase the size of the detergent containing the glycerol head group to LSPG. I then used the formula in the paper: $L(\chi_A)=L_B(L_A-L_B)\chi_A$, where L is the distance from headgroup to headgroup across the micelle core, A and B are two different detergents, and χ is the mixing ratio for detergents A and B. This mixing ratio was used to determine the required ratio of LSPG to three other detergents (DDM, HDPC, and DPC) to generate mixed micelles with similar sizes to that of the LPPG micelle for this analysis.

The results with mixed micelles prepared as above are summarized in Table 3-4. DPC and DDM alone supported little to almost no helical content for TM1-TM2. TM1-TM2 in HDPC and LSPG are highly helical (Figure 3-8). When the detergents are mixed to generate mixed micelles of the same size, the helicity is more similar to that found in LSPG and HDPC than in DPC or DDM alone (Figure 3-8 A and C; Table 3-4). My conclusion from the CD analysis on TM1-TM2 in various detergents and mixed detergents was that LPPG alone and both LSPG:DDM and LSPG:HDPC were the best media for supporting a TM1-TM2 structure that would mimic that in the native receptor and in membranes. This conclusion will be checked by NMR analysis as described in a later section of this chapter. It should be noted that the presence of two detergents does not necessarily mean that the two detergents are mixed in a pure micelle. When mixed with a peptide, the different types of detergents may interact with different regions

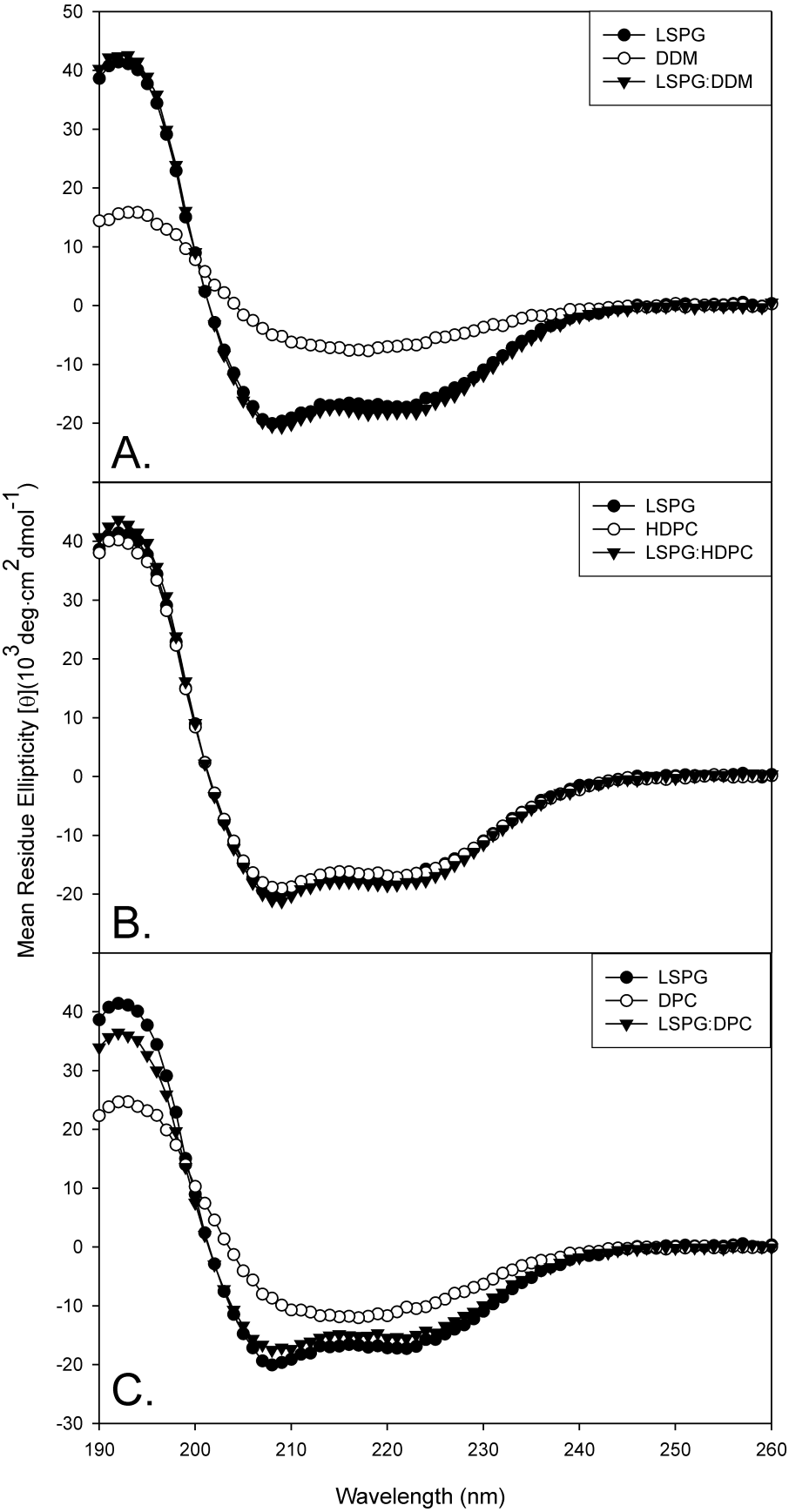


Figure 3-8. Circular Dichroism of TM1-TM2 in mixed micelles of the same predicted diameter. Micelles that should have the same size of micelles composed of LPPG were designed using the equation: $L(\chi_A)=L_B+(L_A-L_B)\chi_A$, where L is the distance from headgroup to headgroup across the micelle core, A and B are two different detergents, and χ is the mixing ratio for detergents A and B. It was determined that the molar ratios of the detergents used to prepare the putative mixed micelles were as follows: A) LSPG:DDM (3.7:1), B) LSPG:HDPC (3.8:1) and C) LSPG:DPC (6.3:1). Micelle-peptide preparations were prepared as described in the Methods section.

of the peptide with varying affinities (*168*). This may influence the actual detergent species that affect peptide conformation and stability. Since the goal of this research was to define conditions that could be used to conduct high resolution NMR analysis of the Ste2p fragments, I decided not to analyze the mixed detergent micelles in detail. I will return to this question briefly in the perspectives section of the dissertation.

Far UV CD analysis of TM6-TM7-CT40 in micellar environments. Given the level of helicity and the good sample stability, as determined by a lack of precipitation out of solution, for TM1-TM2 in micelles, I also investigated TM6-TM7-CT40 in these membrane mimetics. As observed above in the TFE:water analysis, the overall helicity of TM6-TM7-CT40 was less than that of TM1-TM2 under comparable conditions (Tables 3-2 and 3-5). The highest helicity for TM6-TM7-CT40 was observed in LPPG micelles, but TM6-TM7-CT40 was also somewhat helical in SDS micelles (Figure 3-9, Table 3-5). When analyzed in DPC micelles, the helicity decreased, but some helical content remained (Figure 3-9, Table 3-5). These results differ from the helicity of TM1-TM2 in DPC micelles because the TM1-TM2 peptide lost a large amount of its helical content with a significant decrease in f_H and a shift in the 222:208 ratio from below 1 to 1.26. In the case of TM6-TM7-CT40, this peptide retained about 50% of the helicity relative to the peptide in other detergent micelle complexes and the 222:208 ratio remained below 1. As described above, the overall decrease in helicity as compared to TM1-TM2 was not unexpected and the helical content in LPPG and SDS is comparable to the predicted helicity by the hydropathy analysis (Table 3-1). My conclusion was that both LPPG and SDS micelles could be good candidates for analysis by HSQC. Analysis of TM6-TM7-CT40 in DPC will also be performed as a control.

Table 3-5. Fraction helicity of TM6-TM7-CT40 in a micellar environment.

	Mean Residue Ellipticities (deg $cm^2 dmol^{-1}$)		Helicity Calculations		Detergent Information	
	208 nm	222 nm	f_H	222:208	Acyl chain length	Head Group
LPPG	-12500	-12400	0.33	0.99	16	Phosphoglycerol
SDS	-11600	-9800	0.25	0.85	12	sulfate
DPC	-7700	-7200	0.16	0.94	12	Phosphocholine

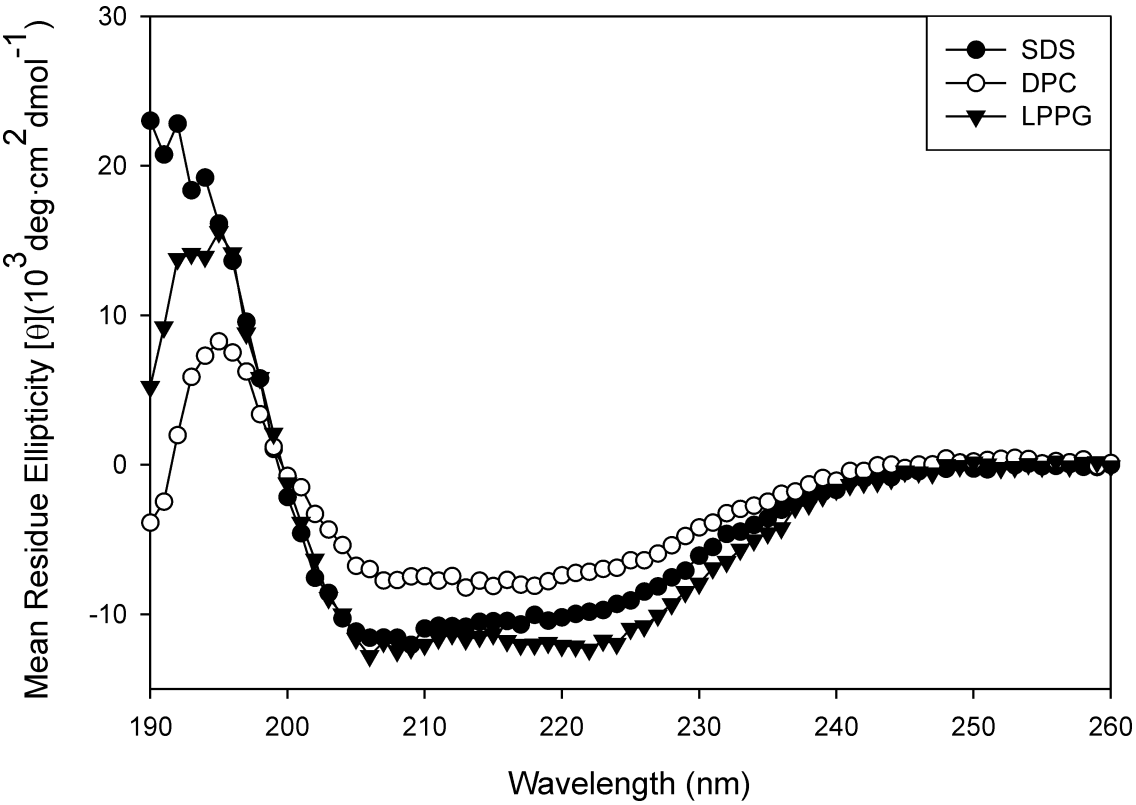


Figure 3-9 Circular Dichroism of TM6-TM7-CT40 in micelles. Purified TM6-TM7-CT40 was solubilized to 25 μM final concentration in micelles prepared in detergent solutions of 20 mM sodium phosphate buffer pH 5.6 with 20 mM SDS, DPC, or LPPG as described in the Methods. The symbols used are at the top of the graph.

[^{15}N,^{1}H]-HSQC analysis of TM1-TM2 in TFE:water. [^{15}N,^{1}H]-HSQC analysis can be used as an indicator of a membrane protein's structural and chemical stability in various environments. The number and width of the correlation peaks as well as their dispersion in this 2-dimensional (2D) experiment are indicators of whether the sample would be useful in 3D and 4D experiments. The purified ^{15}N-labeled TM1-TM2 peptide was solubilized in TFE:water(0.1% TFA)(1:1, v:v). The peptide was soluble at concentrations up to 1 mM and at could be studied at temperatures up to 45°C.

The HSQC spectrum of [^{15}N]-TM1-TM2 in TFE:water(0.1% TFA)(1:1,v:v) showed a peak dispersion of the backbone residues of about 1.8 ppm, from about 7.4 ppm to 9.2 ppm (Figure 3-10). This is a peak dispersion that would be consistent with a peptide that is helical and contains redundant aliphatic amino acids. There were 76 N to H correlation peaks out of the expected 78 and they were sharp peaks of equivalent intensities. Moreover the sample was stable for an extended period and the spectrum measured after a month with storage at room temperature had not changed (data not shown). These observations taken together indicate that this membrane mimetic would be a useful one to use in the structure analysis of TM1-TM2.

Analysis of [^{15}N]-TM1-TM2 was also performed in TFE:water(0.1% TFA)(1:3, v:v) at 45°C (Figure 3-11). When compared to the spectrum in Figure 3-10 the peaks appeared to be broader and not as well resolved as those in 50% TFE. Though the peak dispersion was similar to the sample in TFE:water (1:1), the peak number was decreased. Furthermore, the sample was not stable over time and by 13 hours virtually none of the peaks were still visible. Therefore this sample condition is not suitable for 3D and 4D NMR analysis.

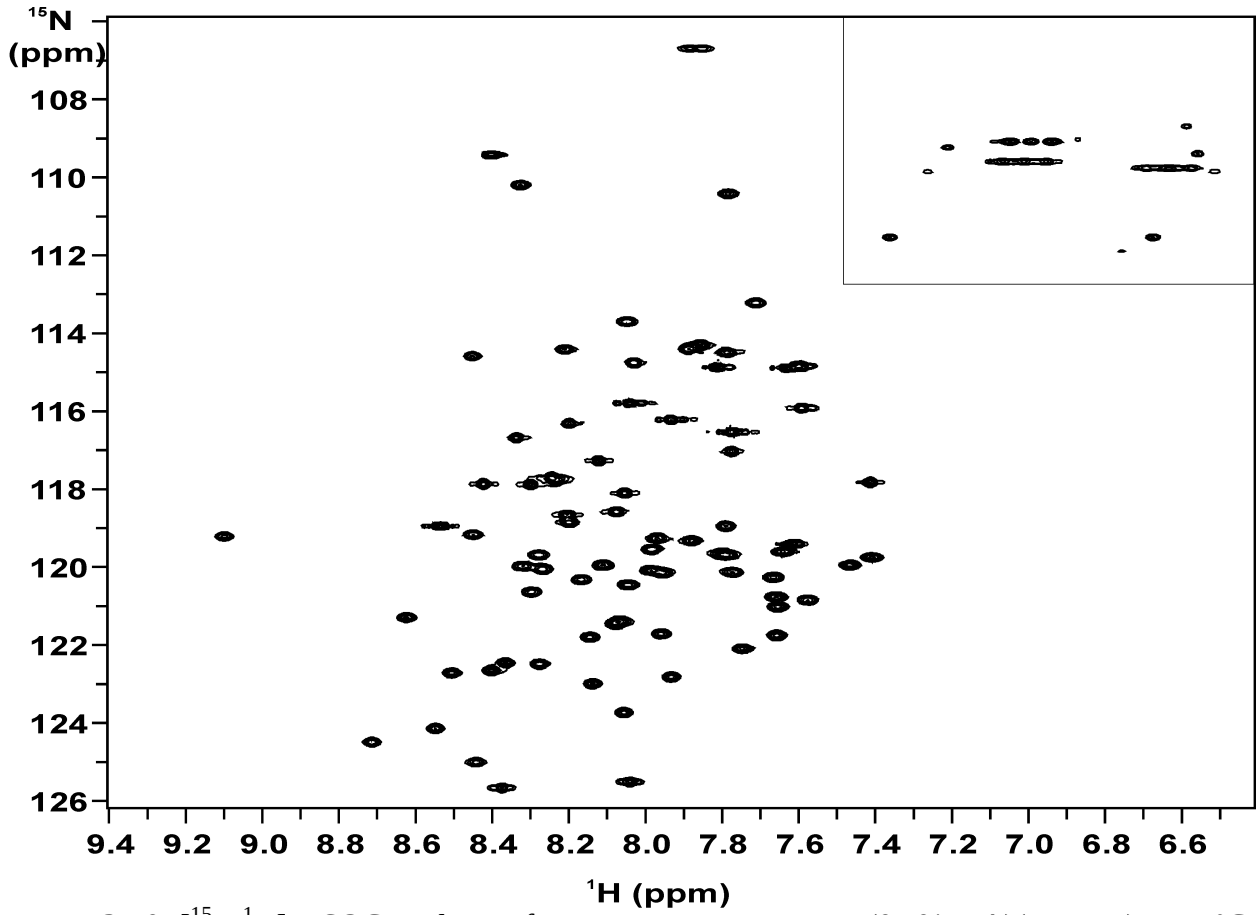


Figure 3-10. [^{15}N,^{1}H]-HSQC analysis of TM1-TM2 in TFE:water(0.1%TFA)(1:1, v:v) at 45°C. ^{15}N-TM1-TM2 was solubilized in TFE:water(0.1% TFA) (1:1, v:v) to a final concentration of 0.25 mM. The sample was heated to 45°C prior to acquisition of the HSQC spectrum. The side chain amide crosspeaks are in the boxed region.

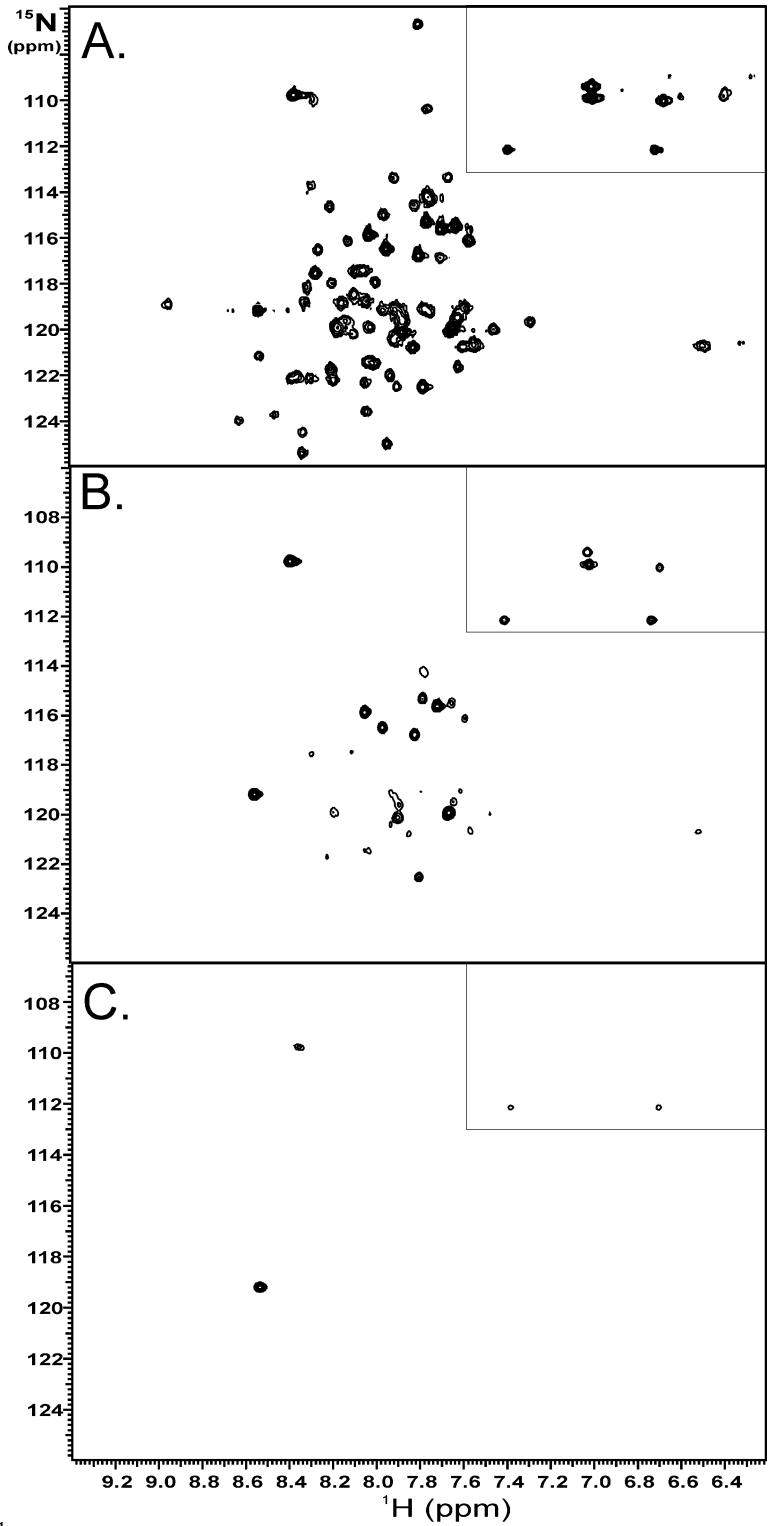


Figure 3-11. [^{15}N,^{1}H]-HSQC analysis of TM1-TM2 in TFE:water(0.1%TFA) (1:3, v:v) at 45°C. NMR analysis was performed on a 0.25 mM ^{15}N-TM1-TM2 sample solubilized in TFE:water(0.1%TFA) (1:3, v:v). The side chain residues are in the box in the upper right corner. HSQC analysis was performed over time A) 30 min after sample preparation, B) 3 h after sample preparation, and C)13 h after sample preparation.

[^{15}N,^{1}H]-HSQC analysis of TM1-TM2 in dimethyl sulfoxide. A second organic:aqueous membrane mimetic medium that has been used in NMR is dimethyl sulfoxide (DMSO) (*194-202*). This sample could not be analyzed using CD due to its spectroscopic properties (*203*). During the sample preparation, it appeared that the peptide did not fully go into solution in the DMSO. One-dimensional ^{1}H and [^{15}N,^{1}H]-HSQC NMR at 25°C were acquired on the samples that were prepared with and without TFA as described in the Methods section. The HSQC analysis of TM1-TM2 in DMSO showed only 10 backbone crosspeaks after an acquisition time of 46 minutes. Based on this finding, it was determined that DMSO was not a useful membrane mimetic in which to study the structure of TM1-TM2.

[^{15}N,^{1}H]-HSQC analysis of 2TM peptides in micellar environments. The helical nature of the TM1-TM2 peptide in micelles such as LPPG, SDS and LPPC based on the CD data described above led me to analyze the HSQC fingerprint in these micelle environments. As a control, I have also chosen to analyze the TM1-TM2 peptide in DPC micelles as this detergent showed poor peptide helicity by CD analysis. The peptide-micelle complexes were prepared as described in the Methods and the [^{15}N,^{1}H]-HSQC analysis was performed at 45°C (Figure 3-12). The backbone peak dispersion in LPPG and SDS was about the same, 1.5 ppm. Though the peaks were sharper in LPPG micelles than in SDS micelles, each displayed similar numbers of peaks, about 74 out of 78 expected peaks. The main difference between the two micelle environments was that the line widths were much better in LPPG micelles than in SDS. Furthermore, the TM1-TM2 in LPPG micelles was stable for at least one month. In fact, during a comprehensive study of detergent micelles with integral membrane proteins, the Girvin group determined that the most successful membrane mimetic was LPPG (*182*). TM1-TM2 in LPPC

micelles had high helical content based on CD, but when analyzed by [^{15}N,^{1}H]-HSQC, the peaks were broad and overlapped. Though the peak dispersion was similar to LPPG, I believe that this would not be a useful micelle environment for structure analysis. Finally, NMR of [^{15}N]-TM1-TM2 in DPC micelles resulted in 5 peaks with very low signal:noise ratio and a precipitate was observed after a short time indicating an unstable peptide-micelle complex (data not shown). Based on my observations the CD analysis of the helicity of the TM1-TM2 peptide correlates reasonably well with the data obtained by initial HSQC experiments. The membrane mimetic media that gave the helicity that was the most similar to predicted helical characteristics also gave the sharpest and best resolved HSQC spectrum. Therefore, CD would be a good screening technique for NMR conditions.

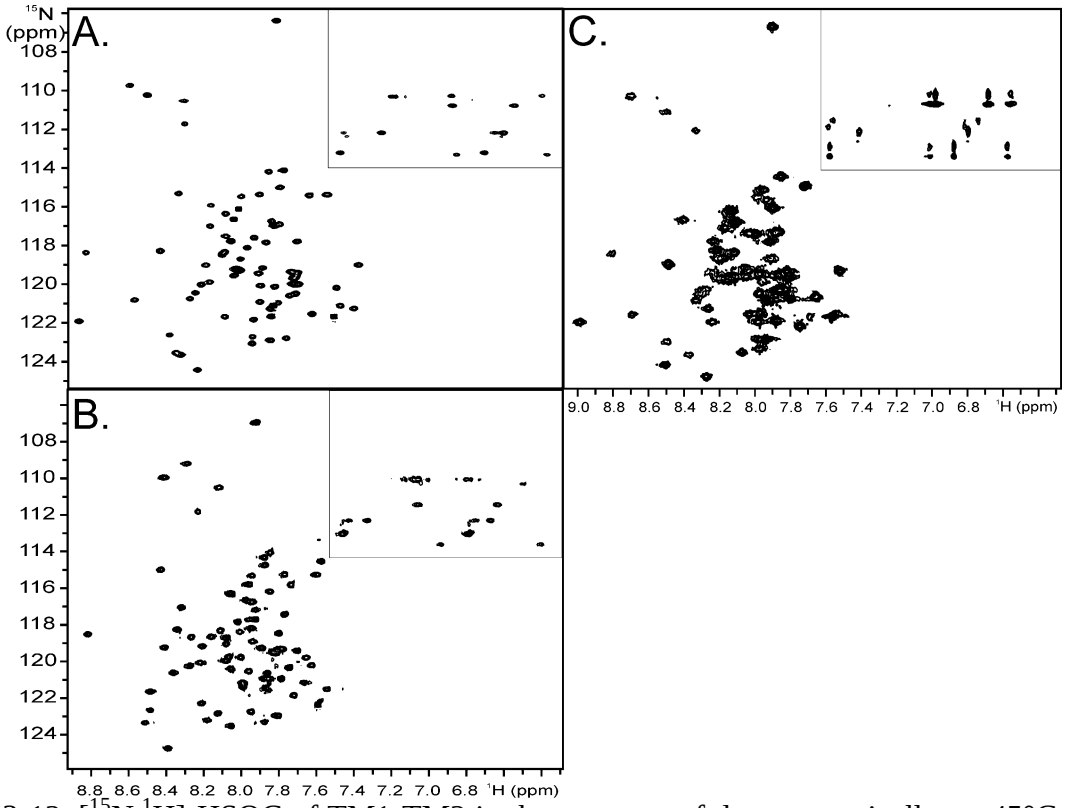


Figure 3-12. [^{15}N,^{1}H]-HSQC of TM1-TM2 in the presence of detergent micelles at 45°C. TM1-TM2 (1 mg) was solubilized in A) 60 mg LPPG, B) 60 mg SDS, or C) 40 mg LPPC in 20 mM sodium phosphate buffer, pH 6.5:D_2O (9:1) as described in the methods. The HSQC analysis was performed at 45°C for about 2 hours. The sidechain crosspeaks are in the boxes in the top right of each spectrum.

[^{15}N,^{1}H]-HSQC analysis of [^{15}N]-TM1-TM2 peptides in mixed micelles. Analysis of TM1-TM2 by circular dichroism indicated that the peptide was highly helical in the presence of detergents containing two different headgroups. Therefore HSQC spectra of [^{15}N]-labeled TM1-TM2 was analyzed in LSPG:DDM, LSPG:DPC and LSPG:HDPC (Figure 3-13). The peak dispersion of [^{15}N]-TM1-TM2 in these mixed micelles appeared to be about 1.6 ppm which is similar to that of LPPG. In all cases, though, the resolution was poor and the peaks were broadened. Comparisons of peak position could be made to TM1-TM2 in LPPG indicating that the TMs may be in similar conformations. However, due to the peak broadening and overlap in the middle of the spectra, these systems would not be as useful for 3D and 4D NMR analysis.

[^{15}N,^{1}H]-HSQC analysis of [^{15}N]-TM6-TM7-CT40 peptide in micelles. The analysis of the C-terminal 2TM peptide in micelles was indicative of a less structured peptide than TM1-TM2. The peak dispersion in both LPPG and SDS micelles was about 1 ppm (Figure 3-14A,B) whereas peak dispersion for TM1-TM2 in these same micelles was 1.5 ppm. TM6-TM7-CT40 is a 109-residue peptide that contains 2 Pro residues, therefore the number of peaks expected is 106. In LPPG micelles, 93 backbone peaks were observed, whereas in SDS micelles 106 backbone peaks were observe (Figure 3-14A,B). The peak numbers were higher in SDS, but this peptide had sharper peaks that were of more equal intensity in LPPG than in SDS. Under both conditions the crosspeaks were not as sharp as the TM1-TM2 crosspeaks in the same media (see Figure 3-12 for comparison). Stability of TM6-TM7-CT40 in micelles was poor which would also make it more difficult to analyze by higher dimensional NMR experiments. The HSQC analysis of TM6-TM7-CT40 in the presence of DPC micelles resulted in fewer than expected peaks (~40 peaks) (Figure 3-14C). Even when the experiment was performed with increased

sensitivity, the peak dispersion and peak number were not improved (Figure 3-14D). The TM6-TM7-CT40 peptide is not a good candidate for structure determination in detergents at this point. Additional attempts at medium and sample preparation optimization are necessary. My focus will be on the structure determination of TM1-TM2 in TFE:water.

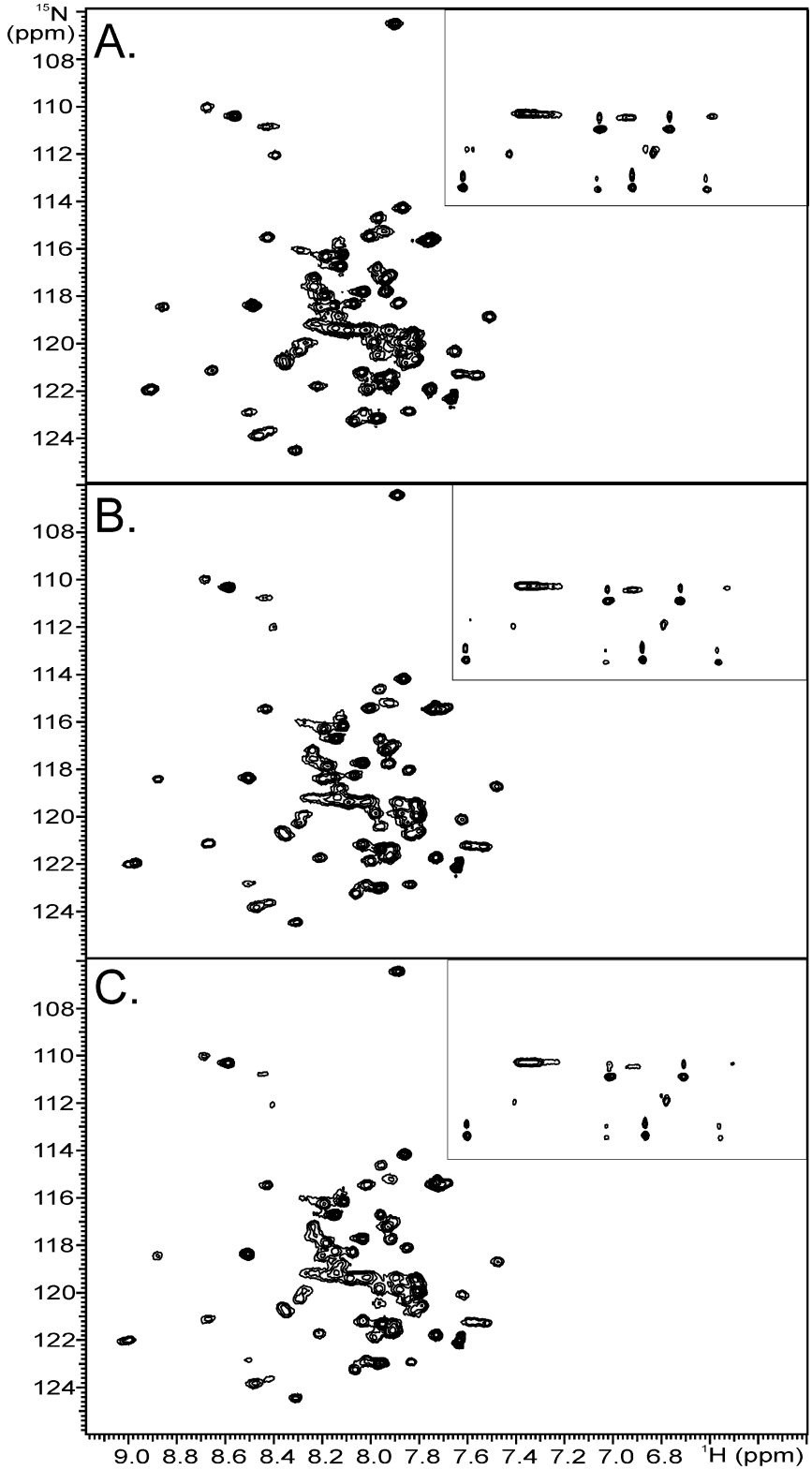


Figure 3-13. [^{15}N,^{1}H]-HSQC of TM1-TM2 in mixed detergent micelles at 45°C. TM1-TM2 was solubilized to 0.25 mM in 60 mM mixed micelles in sodium phosphate buffer pH 6.5:D_2O (9:1). The ratio of the detergents was determined by the formula in (*192*). A) The LSPG:DDM ratio was 3.7:1. B) The LSPG:DPC ratio was 6.3:1. C) The LSPG:HDPC ratio was 3.8:1. The samples were prepared as described in the Methods section.

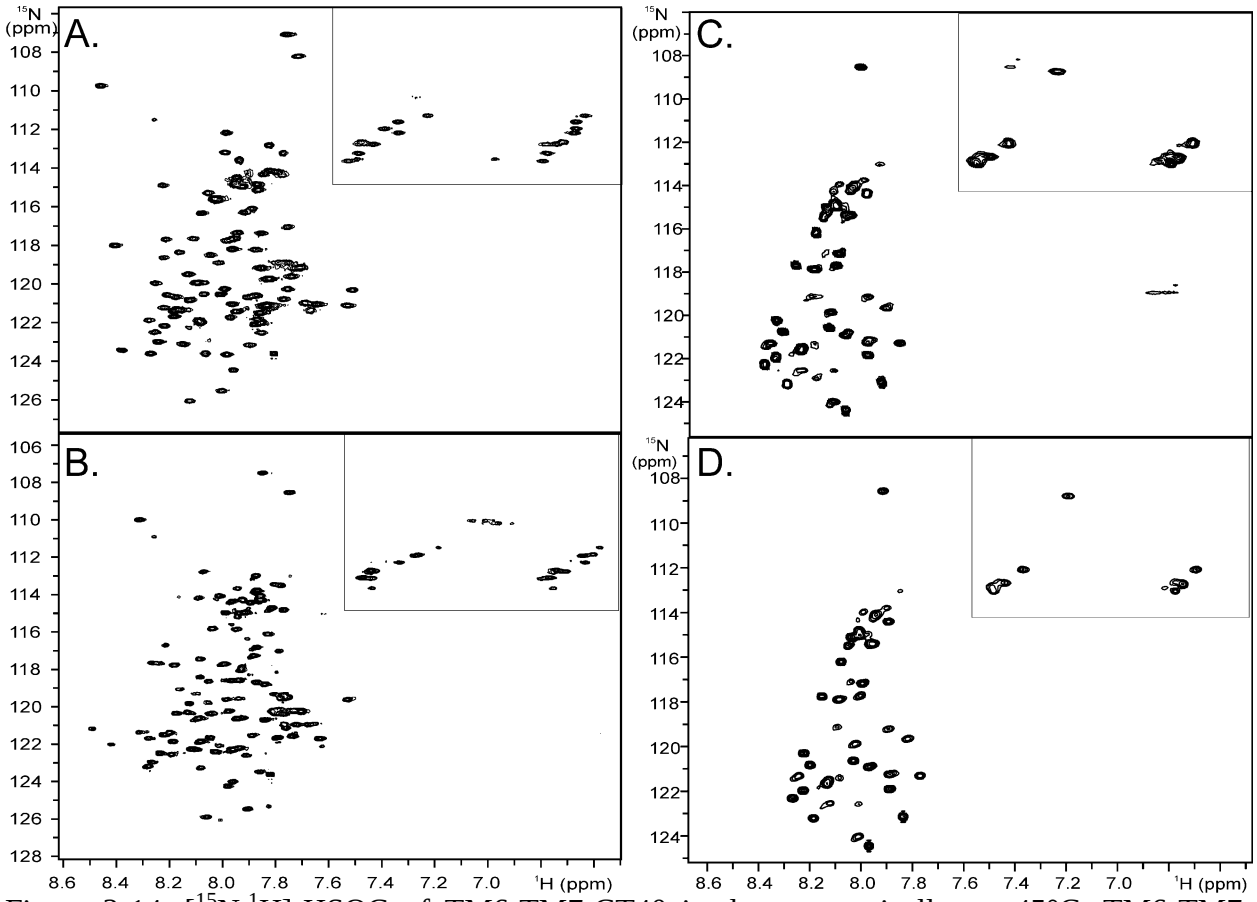


Figure 3-14. [^{15}N,^{1}H]-HSQC of TM6-TM7-CT40 in detergent micelles at 45°C. TM6-TM7-CT40 was dissolved in the presence of A) 25 mg LPPG, B) 25 mg SDS, or C and D) 25 mg DPC micelles in 20 mM sodium phosphate buffer, pH 6.5:D_2O (9:1) to give a final peptide concentration of 0.2 mM as described in the methods. The HSQC analysis was performed at 45°C for about 2 hours. The pulse sequence for the TM6-TM7-CT40 sample in D) was a sensitivity-enhanced HSQC pulse sequence. The sidechain crosspeaks are in the boxes at the top right of each spectrum.

3d. Summary

Initial biophysical analysis is a useful tool to determine aspects of protein conformation and stability and can be used as a screen for conditions for additional structure analysis. For example, groups have been using HSQC fingerprinting as a tool to determine useful solvents and/or conditions for a detailed NMR structure analysis of various polypeptides. If the peak number, peak resolution and/or peak dispersion are poor in these HSQC analyses the sample is thought to be unacceptable for 3D and 4D NMR experiments. I have taken this analysis a step backward in that I developed a screening technique by CD analysis which requires much less peptide. Furthermore, the analysis is done within 20-40 minutes which makes it much more efficient than HSQC analysis. CD does not require isotopically labeled peptide which makes it more economical. There appeared to be a correlation between the ability of peptide fragments of a GPCR to form highly helical structures as judged by CD and the usefulness of a given membrane mimetic solvent for NMR analysis of the polypeptide. LPPG micelles and TFE:water(0.1%TFA) (1:1,v:v) were determined to be appropriate membrane mimetics to conduct a structure analysis of the TM1-TM2 fragment; Ste2p(G31-T110;M54L,C59S,M69V,M71I),based on the CD and HSQC data presented above. Conversely, even though the CD of the 2TM fragment Ste2p(R231-S339;M250A,C252S,M294A) appeared to be helical in micellar environments the HSQC analysis resulted in broad peaks that were not well resolved and the samples were not stable over time. Most significantly all media identified by CD screening to be poor helix supporters gave very poor HSQC spectra with very few crosspeaks. Thus CD can be used to rule out these media saving isotopically labeled peptide.

Overall I believe that the use of circular dichroism as an initial screening technique can help to define the conditions for a high resolution NMR analysis of membrane peptide

fragments. The peptides that have high helical content as observed by mean residue molar ellipticities, f_H and 222:208 comparisons appeared to present HSQC spectra with good peak dispersion and high peak counts. The caveats to this analysis are that since the NMR requires a much higher concentration of peptide, the stability data obtained by CD spectroscopy may not always be extrapolated to the NMR conditions. This appears to be the case for TM6-TM7-CT40 for which the correlation between CD helicity and HSQC quality was not as good as for TM1-TM2. Further study by Katrina Caroccia has shown that stability of CD samples can be studied by incubating the sample in a heat block over time periods as long as a week. These stability analyses have been used to narrow down even further the number of conditions to test for with HSQC NMR.

Chapter 4

Structure Determination of Ste2p(G31-T110) in Trifluoroethanol:Water an Organic:Aqueous Membrane Mimetic

4a. Introduction

Solution NMR is a powerful tool for the structure determination of biopolymers, though it is limited by the isotropic tumbling rate of the macromolecular complex under investigation. It is also useful when characterizing protein dynamics and/or protein-ligand interactions. Methods are under development to better study large macromolecular complexes in solution NMR (reviewed in *204*), but high resolution structures of large IMPS, especially full length GPCRs, remain elusive (*85, 90-92*). These proteins are hard to express with isotopic labels, difficult to purify and must be examined in the presence of membrane mimetics such as detergent micelles or lipid bicelles. The large protein/micelle or bicelle complexes are often too large to tumble efficiently, resulting in broad lines and poorly resolved NMR spectra. One strategy to circumvent this problem is to study fragments of IMPs and GPCRs to learn more about the folding of these proteins and to gain insights into their secondary and tertiary structure.

To apply this strategy it is important to work with fragments that are large enough to reflect the structure and function of the intact receptor and which may aid in deciphering protein-ligand interactions. As described in the previous chapters, the study of GPCR fragments appears to be a valid method because they have been shown to reconstitute to form functional proteins both *in vitro* and *in vivo* (*123, 125, 205-207*). Polypeptide fragments of Ste2p, including each single transmembrane domain (TM), loops and the C-terminal tail have been previously studied

by solution NMR (*113, 115, 117, 208*) (Bhuiyan unpublished data) and have provided details on the structural tendencies of discrete regions of this GPCR.

Organic aqueous media has been used for biophysical analysis of hydrophobic peptides for some time (*179-181, 209, 210*). Such media have useful spectroscopic properties and are ideal for circular dichroism and NMR analyses. In NMR analyses, the mobility of an IMP fragment in an organic-aqueous solvent is thought to be much faster relative to that in micelles and bicelles. Although media such as TFE:water are not considered to be biologically relevant by some (*182, 211*), they have been used to determine IMP structures because one can acquire highly resolved spectra (*87, 89, 115, 212*). Moreover, some fragments of both transporters and receptors exhibit long-range connectivities in TFE and chloroform:methanol:water and appear to assume folded structures that may be relevant to the biologically active state (*89, 212*).

As the protein size increases, determination of NOE connectivities between side chains becomes increasingly difficult. This is especially true for IMPs which are predominantly α-helical and which are composed of high percentages of Ala, Leu, Val and Ile residues. In these proteins, the inter-helical interactions tend to be predominantly van der Waals interactions which would require assignment in the methyl region of these aliphatic residues. However, side chain analysis is hindered in larger proteins and protein/micelle complexes because of the chemical shift degeneracy of the multiple methyl groups that are a major component of the transmembrane regions of the protein and because the ability to perform higher dimension NMR is limited by the tumbling rate. A slower tumbling rate results in significant line broadening and a decrease in the signal:noise ratio by affecting the relaxation parameters of the sample. In order to improve relaxation parameters, the polypeptide can be expressed in a perdeuterated background which enables better resolution of the backbone peaks and reduces T2 relaxation pathways, and thereby

aids in the acquisition of 3D and 4D experiments. However, when a peptide is perdeuterated, the invisibility of the side chain deuterons results in loss of valuable spectroscopic information including the determination of long-range NOE connectivities from side chain protons. Because fragments of GPCRs are mostly α-helical little information on long-range contacts comes from connectivities between the backbone atoms, as would be the case in β-sheets, and interactions between helices generally involve side chain groups. The faster isotropic tumbling in organic:aqueous media will allow the level of deuteration to be decreased thus allowing observation of the side chain protons.

In this chapter, I will describe the structure determination of Ste2p(G31-T110; TM1-TM2) in TFE:water(0.1% TFA) (1:1, v:v). Due to the chemical shift degeneracy of the aliphatic amino acids, special labeling techniques such as selectively protonated methyl groups in an otherwise deuterated background were employed. This labeling scheme along with ^{15}N and ^{13}C labels allowed structural characterization of this peptide fragment in an organic-aqueous membrane mimetic. I will also be comparing the structure determined here to the structure determined in LPPG micelles (*213*).

4b. Materials and Methods

NMR sample preparation. NMR samples containing 0.5-4 mg of isotopically labeled TM1-TM2 were prepared in 350 μL TFE-d_2:water(0.1% TFA) (1:1, v:v) or TFE-d_3:D_2O(0.1% TFA-d) (1:1, v:v) depending on the experiment performed as indicated in the results section. The peptide was quantitated by UV as described in the Materials and Methods section of Chapter 3 and then aliquoted and lyophilized. The samples were prepared by the addition of 175 μL TFE and then 175 μL water(0.1% TFA). The samples generally went into solution, but sometimes the samples needed to be sonicated and/or filtered prior to the NMR experiments. Some samples were exchanged in TFE:D_2O(0.1% TFA) prior to NMR analysis. These samples were solubilized in a large excess of TFE:D_2O(0.1% TFA) and incubated at room temperature overnight or at 50°C for 1 hour to exchange a portion of the amide protons to deuterons.

NMR Measurements. The NMR experiments were performed on a Varian 600 MHz spectrometer at the College of Staten Island, CUNY, a Bruker Avance 700 MHz spectrometer at the University of Zurich, Switzerland, or a Bruker 900 MHz spectrometer at the New York Structural Biology Center (NYSBC) in New York City.

NMR Experiments. The experimental parameters are defined in Table 4-1. Sample stability and initial [^{15}N,^{1}H]-HSQC fingerprinting analysis was performed at 45ºC, 35ºC, 25ºC and 15ºC (*214*). The backbone resonances were assigned by performing HNCO (*215-218*), HN(CA)CO (*94*), HNCA (*215-218*),and HNCACB (*217-219*) and experiments at 45ºC and 25ºC. The ^{15}N-resolved NOESY-HSQC (*214, 220*) and the ^{15}N-resolved TOCSY-HSQC (*214, 220*) were performed at 45ºC and 25 ºC to assign side chain proton chemical shifts and an HCCH-TOCSY

experiment was performed at 25ºC to assign carbon and proton side chain resonances(*101*, *221*). The methyl protons and carbons of the Ile, Leu and Val residues were assigned using HMCMCBCANH (*106*) and ct-[^{13}C,^{1}H]-HSQC (*222*) experiments performed on [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-TM1-TM2 at the University of Zurich by the Zerbe group. Further side chain analysis was performed with ^{13}C-edited TOCSY and ^{13}C-edited NOESY experiments at 25ºC. Unlabeled TM1-TM2 was used in NOESY experiments (*223*) acquired with 300 and 150 msec mixing times and TOCSY experiments (*224*, *225*) with 60 and 25 msec mixing times. The [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-TM1-TM2 peptide was also used in NOESY-ct-[^{13}C,^{1}H]-HSQC (*214*, *226*, *227*) and ct-[^{13}C,^{1}H]-HSQC-NOESY-ct-[^{13}C,^{1}H]-HSQC (*228*) to determine long-range connectivities between the methyl groups of Ile, Leu or Val residues.

Dynamics of [^{15}N]-TM1-TM2 were analyzed by performing [^{15}N,^{1}H]-HSQC with varying relaxation delays to determine T2 relaxation at 45ºC and 25ºC. [^{15}N,^{1}H]-HSQC of [^{15}N]-TM1-TM2 was measured with or without saturation to determine heteronuclear NOEs (*229*). [^{15}N]-TM1-TM2 was solubilized in TFE-d_3:D_2O(0.1% TFA-d) and [^{15}N,^{1}H]-HSQC experiments were performed over time to follow the backbone H-D exchange rates.

Structure calculations. The distance constraints were obtained from the NOESY experiments described above. NOEs were assigned and extracted in NMRView 5 (*230*). The backbone resonances were used to determine dihedral angle constraints using TALOS (*98*). The calculated structures were determined using DYANA (*109*). One hundred randomized structures were calculated and the 20 conformers with the lowest target functions were analyzed as a bundle and the structure was analyzed using Molmol (*231*).

Table 4-1. NMR experiemtnal parameters used in structure determination of TM1-TM2 in TFE:water(0.1%TFA) (1:1, v:v)

NMR Experiment	Complex Points	Sweep Width	Instrument
[^{15}N,^{1}H]-HSQC at 45ºC	2048 (^{1}H)*256 (^{15}N)	6010 (^{1}H)*900(^{15}N)	600 MHz
[^{15}N,^{1}H]-HSQC at 25ºC	4096 (^{1}H)*512 (^{15}N)	6010 (^{1}H)*1500(^{15}N)	600 MHz
NOESY-[^{15}N,^{1}H]-HSQC	2048(^{1}H)*256(^{1}H)*64(^{15}N)	8012(^{1}H)*6600(^{1}H)*12068(^{15}N)	600 MHz
TOCSY-[^{15}N,^{1}H]-HSQC	5195(^{1}H)*7299(^{15}N)*2838(^{1}H)	6010(^{1}H)*6000(^{13}C)*900(^{1}H)	600 MHz
HNCA	2048(^{1}H)*512(^{13}C)*128(^{15}N)	6010(^{1}H)*4526(^{13}C)*900(^{15}N)	600 MHz
HNCACB	2048(^{1}H)*512(^{13}C)*128(^{15}N)	6010(^{1}H)*8000(^{13}C)*900(^{15}N)	600 MHz
HN(CA)CO	2048(^{1}H)*256(^{13}C)*32(^{15}N)	6010(^{1}H)*2500(^{13}C)*900(^{15}N)	600 MHz
HNCO	2048(^{1}H)*512(^{13}C)*128(^{15}N)	6010(^{1}H)*2000(^{13}C)*900(^{15}N)	600 MHz
HCCH-TOCSY	2048(^{1}H)*256(^{13}C)*256(^{15}N)	3613(^{1}H)*3600(^{13}C)*12001(^{15}N)	600 MHz
^{13}C-edited TOCSY	2048(^{1}H)*256(^{13}C)*512(^{13}C)	7225(^{1}H)*12019(^{13}C)*1199(^{13}C)	600 MHz
^{15}N-edited NOESY	2048(^{1}H)*256(^{15}N)*256(^{1}H)	8012(^{1}H)*6000(^{15}N)*10200(^{1}H)	600 MHz
NOESY (300 and 150 msec)	256(^{1}H)*2048(^{1}H)	6010(^{1}H)	600 MHz
TOCSY (60 and 25 msec)	256(^{1}H)*2048(^{1}H)	6010(^{1}H)	600 MHz
NOESY-[^{13}C,^{1}H]-HSQC	2048(^{1}H)*256(^{1}H)*256(^{13}C)	8012(^{1}H)*6000(^{1}H)*10200	600 MHz
[^{15}N,^{1}H]-HSQC for T2 relaxation	2048(^{1}H)*512(^{15}N)	6009(^{1}H)*900(^{15}N)	600 MHz
NOE-[^{15}N,^{1}H]-HSQC	2048(^{1}H)*256(^{15}N)	6010(^{1}H)*900(^{15}N)	600 MHz
[^{15}N,^{1}H]-HSQC for H-D exchange	4096(^{1}H)*512(^{15}N)	6010(^{1}H)*1500(^{15}N)	600 MHz
NOESY-ct-[^{13}C,^{1}H]-HSQC	256(^{1}H)*128(^{13}C)*128(^{1}H)	8091(^{1}H)*3622(^{13}C)*3622(^{1}H)	900 MHz
ct-[^{13}C,^{1}H]-HSQC-NOESY-ct-[^{13}C,^{1}H]-HSQC	1024(^{1}H)*106(^{13}C)*104(^{1}H)	1796(^{1}H)*3622(^{13}C)*1800(^{1}H)	900 MHz
HMCMCBCANH_val/ile	2048(^{1}H)*40(^{15}N)*36(^{13}C)	9765(^{1}H)*1998(^{15}N)*2641(^{13}C)	700 MHz
HMCMCBCANH_leu	2048(^{1}H)*40(^{15}N)*60(^{13}C)	9765(^{1}H)*1998(^{15}N)*2641(^{13}C)	700 MHz
ct-[^{13}C,^{1}H]-HSQC	2048(^{1}H)*400(^{13}C)	9328(^{1}H)*8803(^{13}C)	700 MHz

4c. Results[3]

Assignments of backbone nuclei of the TM1-TM2 peptide at 45°C. Sample stability is very important when performing structural studies on a protein and/or peptide fragment by NMR as it takes several days to acquire the vital set of 2D and 3D spectra. Conditions were determined above (see Chapter 3) which indicated that TFE:water would be a good membrane mimetic solvent to use for the structure determination of TM1-TM2. When performing multidimensional NMR, it is important to use as high a concentration of protein as possible without compromising sample or structural stability. The higher the concentration the faster the acquisition time, due to better signal to noise ratios, and the better the sensitivity of multidimensional NMR experiments with multiple magnetization transfers. If the acquisition time is shorter, more experiments can be run for a single sample preparation. If the concentration is too high, there may be intermolecular interactions or increased viscosity both of which could affect tumbling rates and T2 relaxation parameters. In general, 1 mM is a good concentration to use for multidimensional NMR studies (*232*). We attempted to prepare a 1 mM sample of [^{15}N]-TM1-TM2 in TFE:water for analysis in a [^{15}N,^{1}H]-HSQC experiment at 45°C. The sample went into solution very well and the HSQC spectrum collected after being held at 45°C for one day exhibited sharp peaks using a 60 min acquisition time (Figure 4-1A). The sample was incubated at 45°C in a heat block for 6 days and the experiment was performed again using the same acquisition conditions. There was a loss of peak number and intensity in the one week old sample (Figure 4-1B). In contrast, a [^{15}N]-TM1-TM2 sample at 0.5 mM in TFE:water exhibited a well resolved [^{15}N,^{1}H]-HSQC spectrum with all of the initial peaks and about the same intensity over an extended period of time, at least 22

[3] Note that in this section I will be using TFE:water to define TFE:water(0.1%TFA) (1:1, v:v) unless specified otherwise.

days (Figure 4-2), which means that we would be able to run 3D and 4D NMR experiments at 45°C.

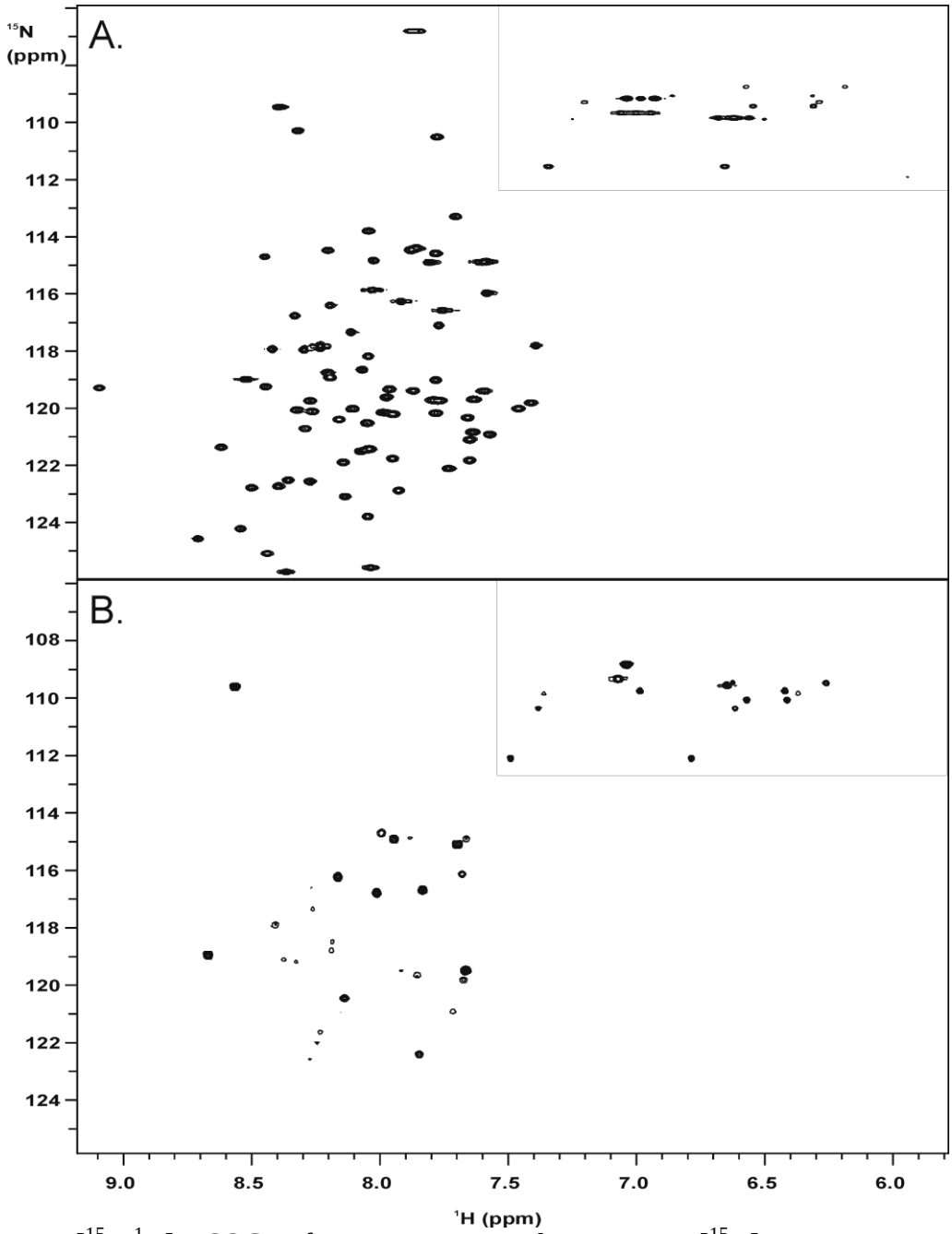


Figure 4-1. [^{15}N,^{1}H]-HSQC of an approximately 1 mM [^{15}N]-TM1-TM2 sample in TFE:water(0.1%TFA) (1:1, v:v). Stability of the NMR sample was examined over a one week period with incubation at 45°C for six days. A) [^{15}N,^{1}H]-HSQC analysis one day after sample preparation. B) [^{15}N,^{1}H]-HSQC analysis six days after sample preparation. The boxed areas contain the side chain NH crosspeaks.

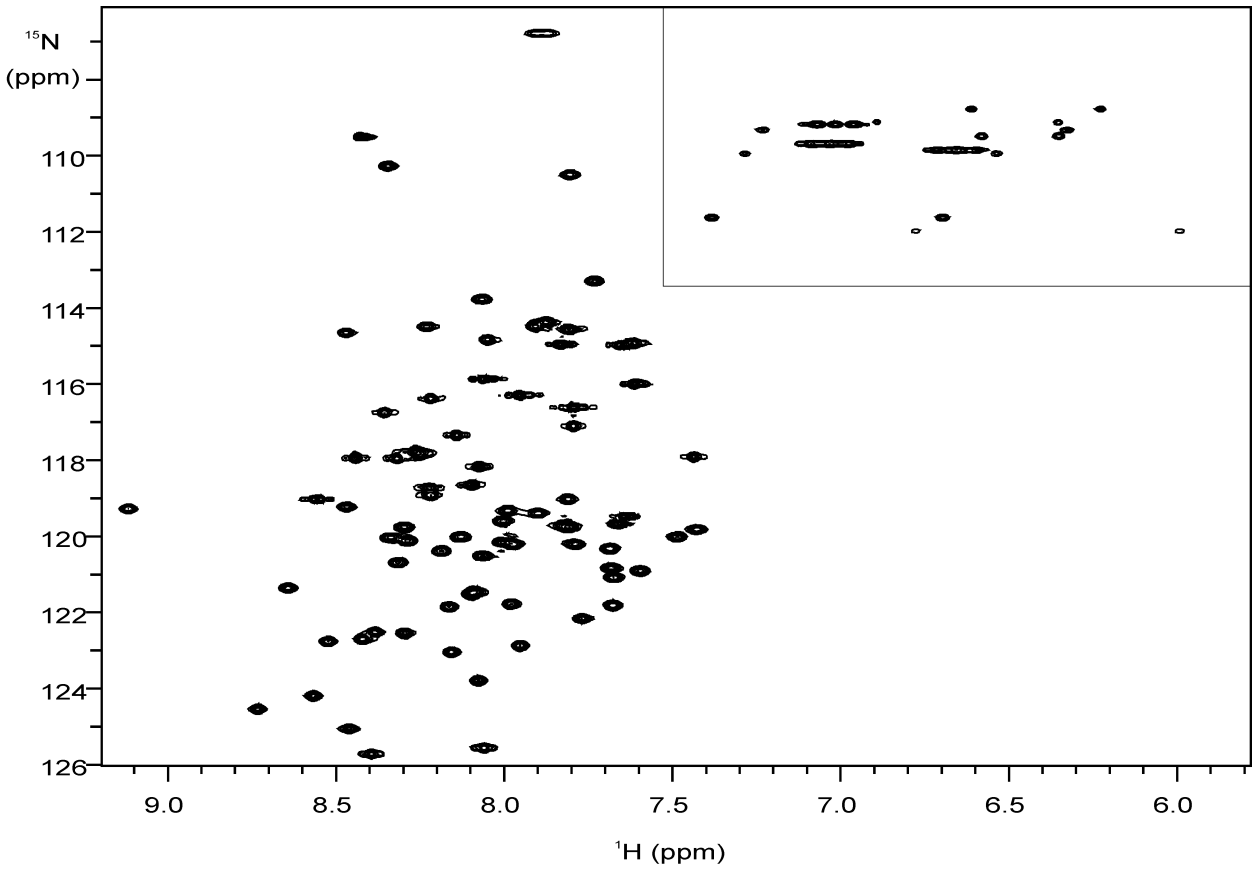


Figure 4-2. [^{15}N,^{1}H]-HSQC of an approximately 0.5 mM ^{15}N-TM1-TM2 sample in TFE:water(0.1%TFA) (1:1, v:v). Stability of the NMR sample was examined after 22 days with incubation at 45°C for six days. The boxed areas contain the side chain NH crosspeaks.

At the initial stages of an NMR analysis, TM1-TM2 selectively labeled with [^{15}N]-amino acids were used to assign peaks to certain types of residues. As described in Chapter 2, I labeled TM1-TM2 with [^{15}N]-Ala, [^{15}N]-Val, [^{15}N]-Ile, [^{15}N]-Leu, [^{15}N]-Ser and [^{15}N]-Phe (*143*). These peptides were dissolved in TFE:water to approximately 0.2 mM and [^{15}N,^{1}H]-HSQC spectra were run at 45°C (Figure 4-3). The analysis was also performed at 25°C with similar results (data not shown). In all cases the HSQC spectrum revealed the number of crosspeaks expected for polypeptide labeled with a given type of amino acid (Ala 5; Ile 9; Leu 11; Phe 5; Val 7; Ser 10). In some cases scrambling of the label was observed but the expected residues were easily assigned based on their relative intensities. Table 4-2 presents the chemical shifts for residues in

each amino acid labeling pattern. Initial backbone assignments were made using the ^{15}N chemical shifts determined from the selectively labeled peptides above and using 3D HSQC-NOESY and 3D HSQC-TOCSY for assignments. We also observed that the spectra could be overlayed with the HSQC for the uniformly [^{15}N]-labeled TM1-TM2 with very little change in the chemical shifts. This indicated that the multiple sample preparations resulted in the same final environments for the GPCR fragment. Ultimately, however the actual resonance assignments were made and the above assignments were verified using 3D experiments such as HNCO and HN(CA)CO.

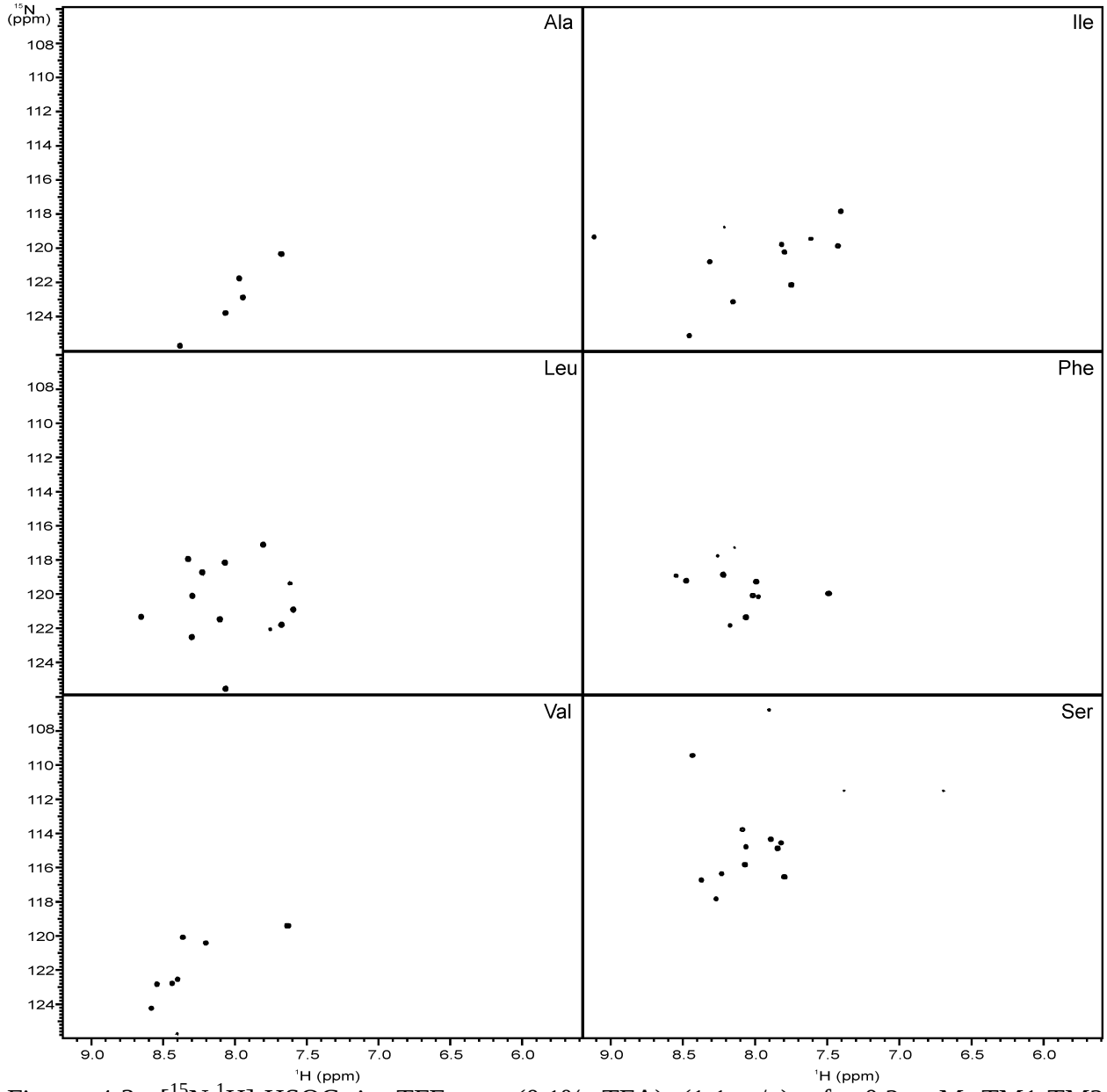


Figure 4-3. [^{15}N,^{1}H]-HSQC in TFE:water(0.1% TFA) (1:1 v/v) of ~0.2 mM TM1-TM2 selectively labeled with [^{15}N]-amino acids. The HSQC experiments were performed at 45°C with similar results observed at 25°C. The ^{15}N-labeled amino acid is indicated at the top of each spectrum.

Table 4-2. ^{15}N and ^{1}H Chemical shifts of TM1-TM2 selectively labeled with different [^{15}N]-amino acids in TFE:water(0.1% TFA) at 45°C.

[^{15}N-Ala]-TM1-TM2	N	HN
A52	122.55	7.948
A61	123.48	8.07
A62	121.43	7.975
A63	120.04	7.685
A96	125.37	8.386

[^{15}N-Ile]-TM1-TM2	N	HN
I36	121.85	7.759
I53	120.48	8.322
I67	119.45	7.822
I71	119.01	9.126
I80	117.54	7.417
I82	119.61	7.439
I83	119.84	7.807
I91	122.78	8.157
I92	124.76	8.46
V109	119.2	7.621

[^{15}N-Leu]-TM1-TM2	N	HN
L102	122.26	8.296
L103	118.46	8.227
L41	121.21	8.1
L44	125.33	8.061
L54	119.77	8.283
L64	116.85	7.803
L66	120.6	7.595
L88	121.53	7.673
L90	117.6	8.316
L93	121.05	8.64
L97	117.89	8.07
I36	121.85	7.759
V109	119.2	7.621

[^{15}N-Phe]-TM1-TM2	N	HN
F38	121.1	8.065
F55	118.96	8.461
F81	119.72	7.486
F89	119.01	7.98
F99	118.63	8.217
N32	118.77	8.548
D39	117.41	8.253
Q42	116.97	8.125
Y98	119.91	7.97
Y106	119.88	8.015
Y101	121.61	8.167

[^{15}N-Val]-TM1-TM2	N	HN
V109	119.2	7.621
V45	122.46	8.531
V49	122.22	8.384
V57	122.51	8.424
V68	120.06	8.184
V69	123.88	8.561
V86	119.77	8.349
A96	125.37	8.386

[^{15}N-Ser]-TM1-TM2	N	HN
S104	114.05	7.883
S107	114.61	7.833
S108	116.3	7.78
S34	115.64	8.059
S47	116.43	8.36
S59	113.53	8.071
S73	117.59	8.247
S75	114.26	7.8
S87	114.45	8.042
S95	116.06	8.218
G43	106.52	7.886
G33	109.22	8.417

The HN(CA)CO/HNCO experiments can be used to assign the carbonyl C for the *i* and *i*-1 residues (*216, 217, 233, 234*). The carbonyl experiments are better to use in the sequential assignment of amino acids than the HNCA/HNCACB experiments due to the high sensitivity of the HNCO experiment and the interresidue connectivities observed during the HN(CA)CO experiment (*232*). The HNCA and HNCACB experiments are used to assign the amide proton, nitrogen, C_α and C_β of the *i* and *i*-1 residues of a given amino acid (*216, 219, 233*) (Figure 4-4). In order to increase the peak sharpness and resolution, the sample that was labeled selectively at the protons of Ile, Leu and Val methyl groups ([^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-TM1-TM2) was used to run these experiments. The only proton that was necessary for the 3D experiments was the amide proton which exchanges back from a deuteron during purification. Using these experiments along with the [^{15}N,^{1}H]-HSQC on [^{15}N]-TM1-TM2 as well as the selectively labeled peptides, we were able to assign 99% of the backbone residues for this peptide. The nuclei that were not assigned were the HN and N of the first Gly residue and the N of the Pro residue. The amino group of the first Gly residue cannot be observed by HSQC based NMR due to the fast rate of exchange of the protons with the water in the solvent. The amide nitrogen of the Pro backbone would not be observed during these experiments because there is no amide proton for the magnetization to be transferred through. The assigned [^{15}N,^{1}H]-HSQC spectrum of [^{15}N]-labeled TM1-TM2 in TFE:water is shown in Figure 4-5. All chemical shifts assigned at 45°C can be found in the Appendix (Supplementary Tables S1 and S2).

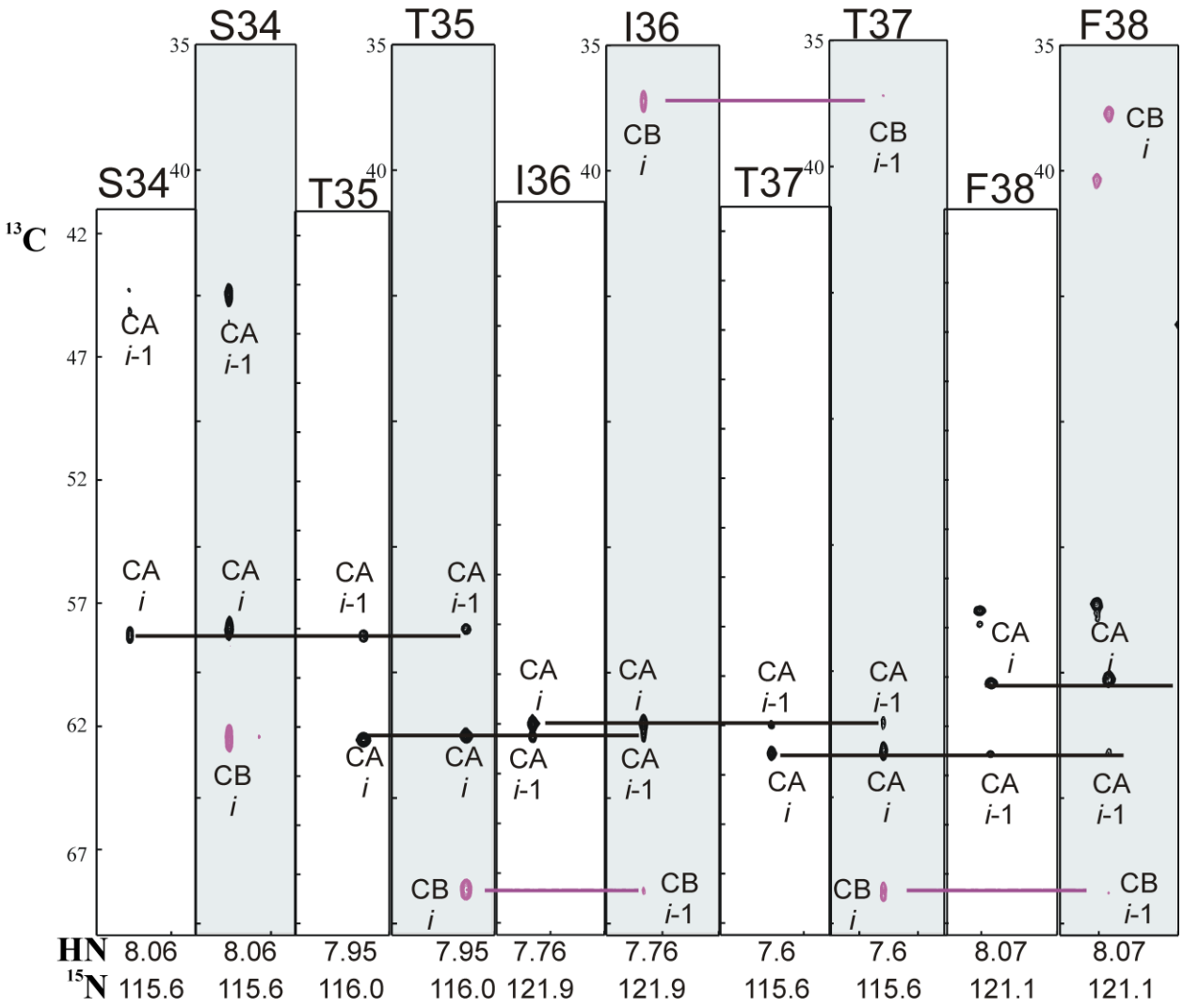


Figure 4-4. Strips of HNCA and HNCACB of [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-TM1-TM2 in TFE:water(0.1% TFA) (1:1, v:v) at 45°C. The HNCA (white) and HNCACB (gray) spectra were analyzed for assignment of CA (black) and CB (pink) chemical shifts. The *i* and *i*-1 residue for each chemical shift is indicated and the CA and CB chemical shifts are connected by lines.

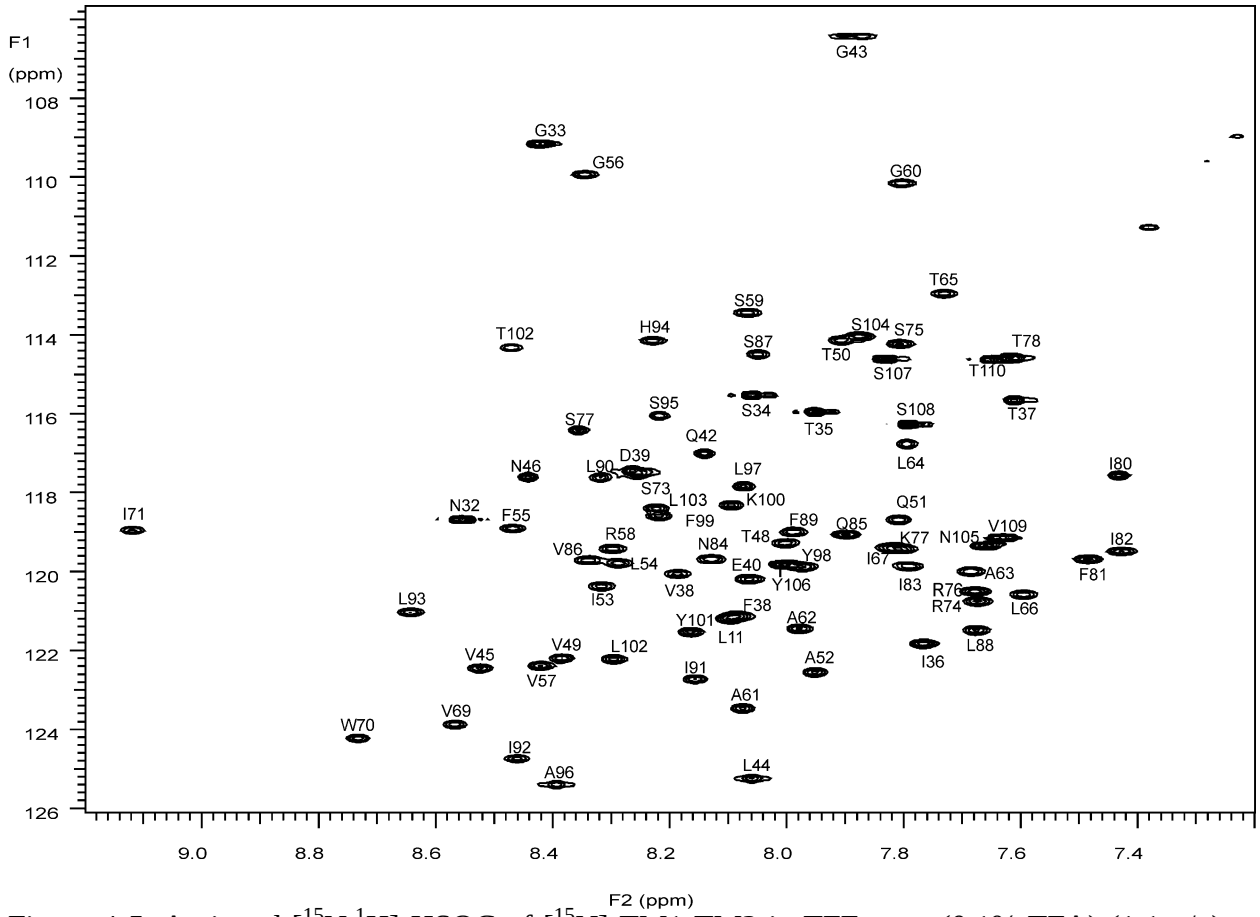


Figure 4-5. Assigned [^{15}N,^{1}H]-HSQC of [^{15}N]-TM1-TM2 in TFE:water(0.1% TFA) (1:1 v/v) at 45°C. The backbone amide crosspeaks were assigned based on the analysis of 3D HSQC-TOCSY, HSQC-NOESY, HNCO/HN(CA)CO and HNCA/HNCACB NMR spectra.

Assignments of sidechain protons of the TM1-TM2 peptide at 45°C. The side chain protons were assigned at this temperature using ^{15}N-resolved HSQC-NOESY and HSQC-TOCSY 3D NMR experiments (*214, 220*). NOESY experiments are through space experiments and could give information concerning through space connectivities within a residue and/or between residues. TOCSY experiments are used to assign protons of the side chain in through bond fashion and are useful for assignments within residues. Only 40% of the observable side chain nuclei were assigned before we decided that performing the experiments at 25°C might increase stability of a tertiary structure. We expected that at the lower temperature hydrogen bond and/or

van der Waals contact formation between the two helices would increase and therefore increase our chances of observing interhelical connectivities.

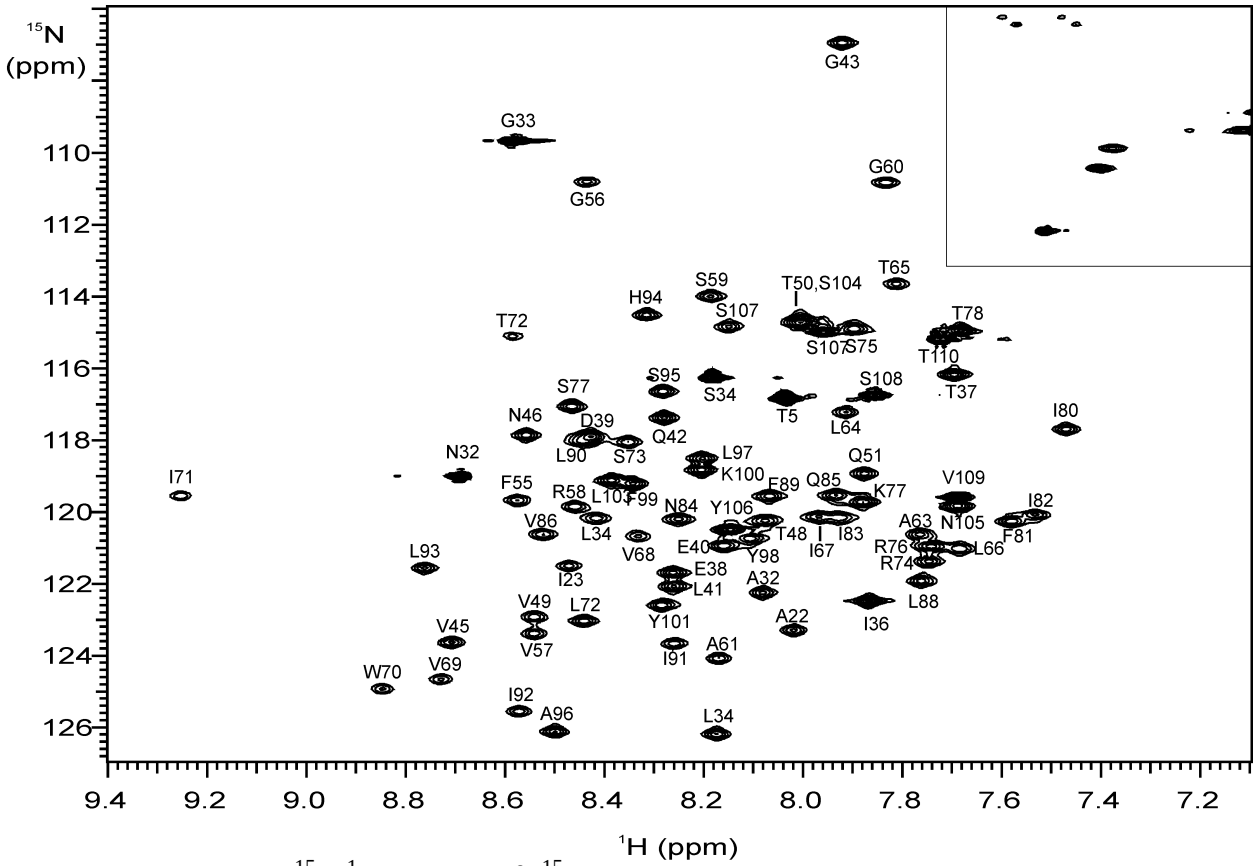


Figure 4-6. Assigned [^{15}N,^{1}H]-HSQC of [^{15}N]-TM1-TM2 samples in TFE:water(0.1% TFA) (1:1 v/v) at 25°C. The backbone amide crosspeaks were assigned based on the already assigned HSQC spectrum at 45°C analysis. The boxed region contains side chain amide crosspeaks.

Assignments of backbone nuclei of the TM1-TM2 peptide at 25°C. We performed the same experiments described above at 25°C and used the assignments made at 45°C as a guide for making assignments at 25°C in TFE:water. The first experiment we ran was the [^{15}N,^{1}H]-HSQC at 25°C with a 0.5 mM sample of the [^{15}N]-TM1-TM2 to determine if the peak dispersion, peak sharpness and peak number would still be useful in 3D experiments (Figure 4-6). In fact, the HSQC spectrum was still reasonably sharp and contained the same number of peaks as found at

45°C. Therefore, we proceeded with the assignment experiments. The assignments of the NH peaks from the 45°C HSQC spectra were related to the 25°C HSQC spectrum by tracking known peaks with temperature (see Figure 4-17 below). The HNCA/HNCACB and HN(CA)CO/HNCO analyses were run with a 0.5 mM sample of the [^{15}N,^{13}C]-TM1-TM2 and were used to corroborate the backbone assignments in conjunction with the [^{15}N,^{1}H]-HSQC data. We were able to assign 97% of the backbone nuclei. In addition to the same nuclei that were not assigned at 45°C, I was also unable to assign the CO carbon of Thr78 and the CA and CB carbons of the Pro79. All chemical shifts assigned at 25°C can be found in the Appendix (Supplementary Tables S3 and S4).

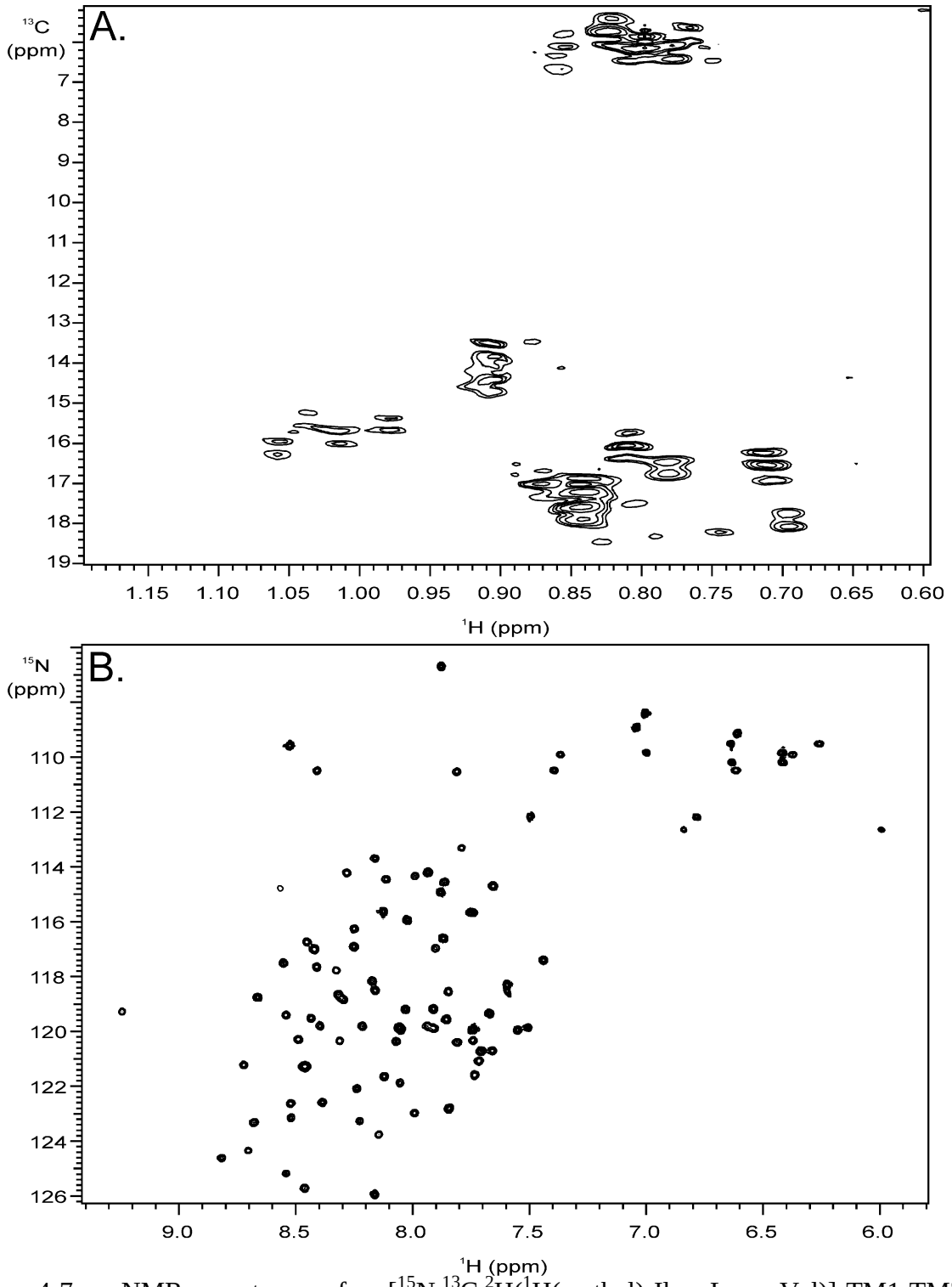


Figure 4-7. NMR spectrum of [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-TM1-TM2 in TFE:water(0.1%TFA) (1:1 v/v). A) ct-[^{13}C,^{1}H]-HSQC analysis of the methyl region chemical shifts observed at 45°C. B) [^{15}N,^{1}H]-HSQC spectrum of the amide groups of the same sample at 45°C.

Assignments of sidechain nuclei of the TM1-TM2 peptide at 25°C. As described above, ^{15}N-resolved NOESY and TOCSY experiments were used to begin the assignments of the sidechain nuclei at 25°C. These experiments were used to assign ~50% of sidechain protons.

Many of the chemical shifts that remained unassigned were those of the aliphatic residues Ile and Leu. These methyl containing residues generally are difficult to assign due to their high redundancy in the TM domains of integral membrane proteins and very small chemical shift differences in similar structural environments (*106, 144, 235*). To assign methyl groups of these residues, which are critical in the determination of long-range interhelical connectivities, a sample was prepared using selective methyl group protonation of a single methyl group on Ile, Leu and Val in an otherwise perdeuterated background (see Chapter 2). The labeling of the sample was analyzed by ESI-MS and 1D ^{1}H NMR (see Chapter 2) as well as by ct-[^{13}C,^{1}H]-HSQC and [^{15}N,^{1}H]-HSQC (Figure 4-7). To assign the methyl carbons, two HMCMCBCANH experiments were performed on a peptide sample prepared in TFE-d_2:water (*106, 213*) and were acquired at the University of Zurich by Dr. Zerbe. The spectra were analyzed by correlating the HN chemical shift assigned above with the C chemical shift of the methyl group. Two experiments were run because the Val experiment was centered at a different chemical shift than the Leu and Ile residues. My original assignments of the Val methyl groups were used to compare our chemical shift data with that acquired in Zurich as well as at the NYSBC. The shifts were off slightly due to the type of spectrometers used to acquire the data, but the spectra could be adjusted during processing (Figure 4-8A). Further analysis of this spectrum allowed the complete assignments of the methyl ^{13}C of the Leu residues (Figure 4-8B). Similar analysis was performed on the Ile HMCMCBCANH spectrum resulting in the assignment of the

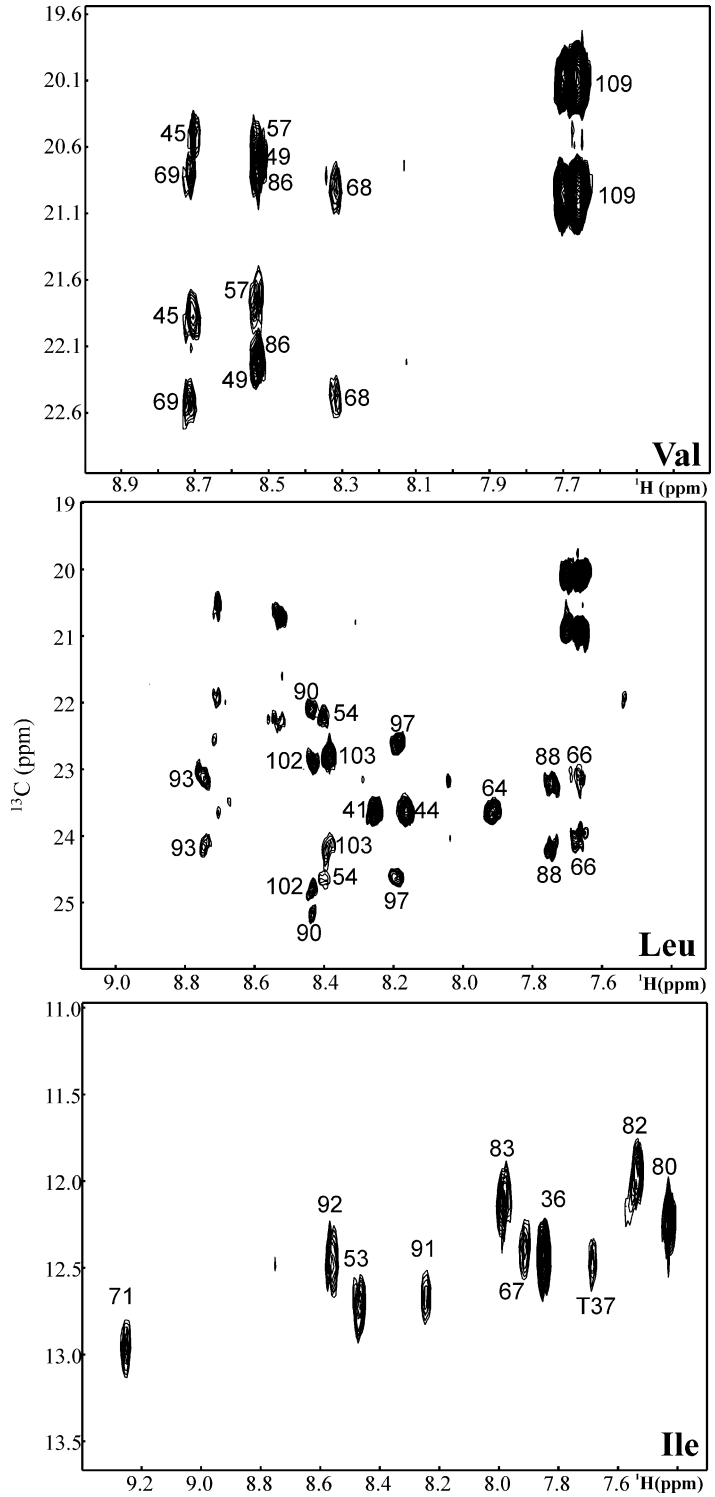


Figure 4-8. Assignment of the methyl carbons using [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-TM1-TM2 in TFE:water(0.1% TFA) (1:1 v/v). The HMCMCBCANH experiments were performed to assign the aliphatic residues. A) Val assignments. B) Leu assignments. C) Ile assignments. The spectra were acquired at the University of Zurich by Dr. Oliver Zerbe.

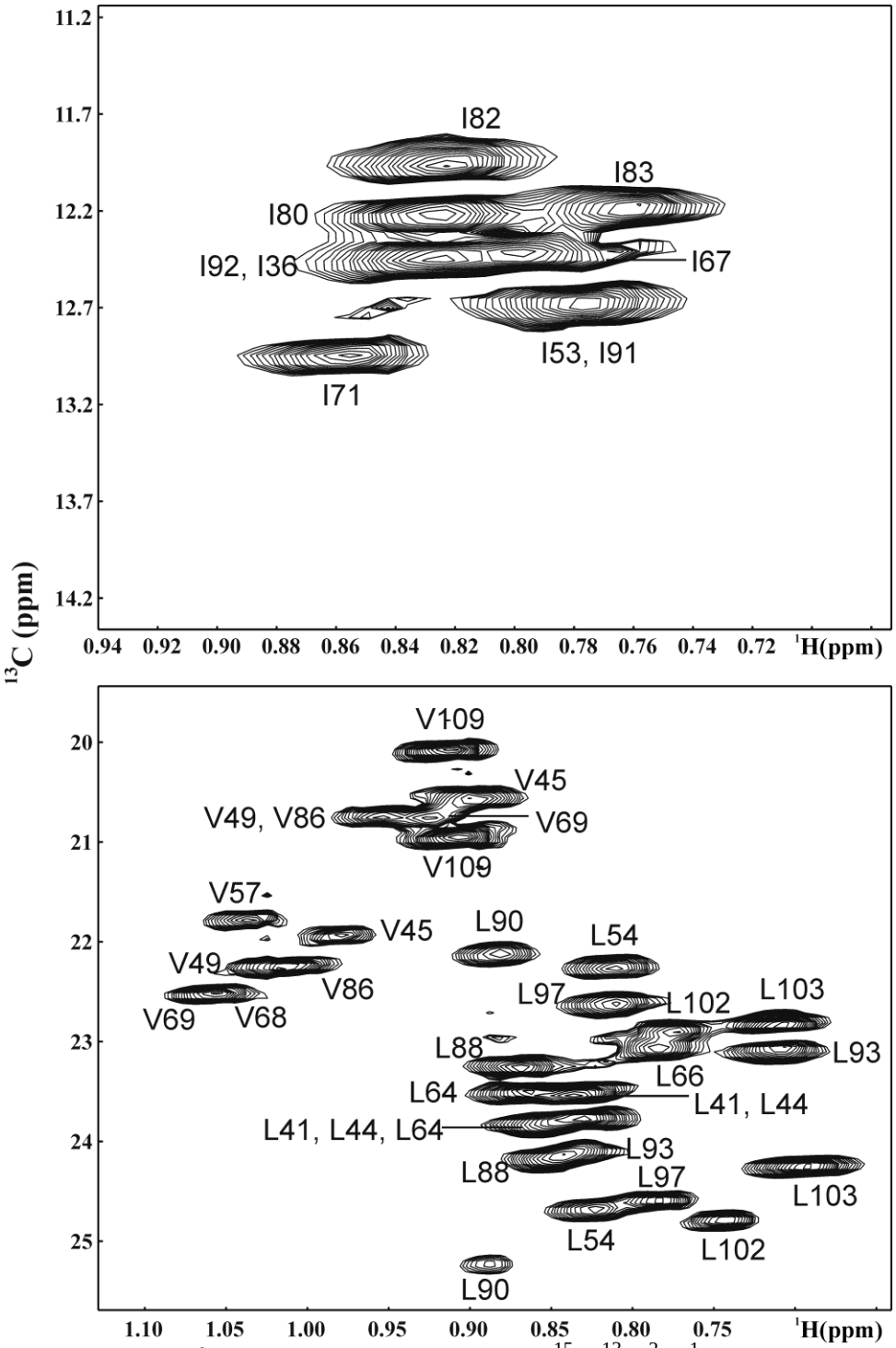


Figure 4-9. Assignment of the methyl protons using [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-TM1-TM2 in TFE:water(0.1% TFA) (1:1 v/v). The ct-[^{13}C,^{1}H]-HSQC experiment was performed to assign the aliphatic residues. Top panel: Ile C-H crosspeaks. Bottom panel: Val and Leu C-H crosspeaks. The spectrum was acquired at the University of Zurich by Dr. Oliver Zerbe.

methyl C of all of the Ile residues (Figure 4-8C). Using the carbon chemical shifts from the HMCMCBCANH experiments we were able to assign the methyl protons for most of the protonated methyl groups by comparing these assignments to a ct-[^{13}C,^{1}H]-HSQC spectra acquired on the same NMR magnet (Figure 4-9). The ct-[^{13}C,^{1}H]-HSQC spectrum contained all of the expected residues and the peaks were fairly well separated with overlap of only a few crosspeaks. Due to the isotopic labeling pattern of the precursor used in the biosynthesis of the protonated Val and Leu residues, both of the methyl groups in the branch were labeled (Figures 4-8 and 4-9 and (*213*)). Each particular Leu and Val residue was randomly labeled in one methyl group, but the overall set of labeled peptide molecules had both methyl groups labeled. This is the explanation for the two peaks per residue in the Val and Leu HMCMCBCANH spectra (Figure 4-8A,B) as well as the double peak count for Val and Leu in the ct-[^{13}C,^{1}H]-HSQC (Figure 4-9). These experiments as well as those performed previously resulted in the assignment of 83% of the nuclei in the Leu and Ile residues.

For the assignments of side chain chemical shifts other than the methyl groups of Ile, Leu and Val, 3D ^{13}C-edited COSY and TOCSY experiments were attempted. Experiments that are used for such assignments include ct-HMQC, HCCH-COSY, HCCH-TOCSY and CCH-TOCSY. All of these experiments do not rely on the HN for magnetization transfer and are run in D_2O with samples in which the amide proton has been exchanged for a deuteron. This aids acquisition of spectra that are clean of water peak artifacts in the Hα chemical shift region. The peptide was exchanged in TFE:D_2O and the amount of exchange was determined by [^{15}N,^{1}H]-HSQC analysis (Figure 4-10). Based on 45 remaining residues present in the HSQC, the exchange was not complete. Further H-D exchange analysis (see below) indicated that the exchange rate for many residues in this peptide is very slow. Some of the residues were still

present even after 1 month in deuterated media. Exchanged samples used in later experiments were prepared at 50°C to facilitate a faster exchange. Nevertheless, this sample was used in HCCH-TOCSY and ^{13}C-edited TOCSY experiments. These spectra were analyzed and additional side chain assignments were made.

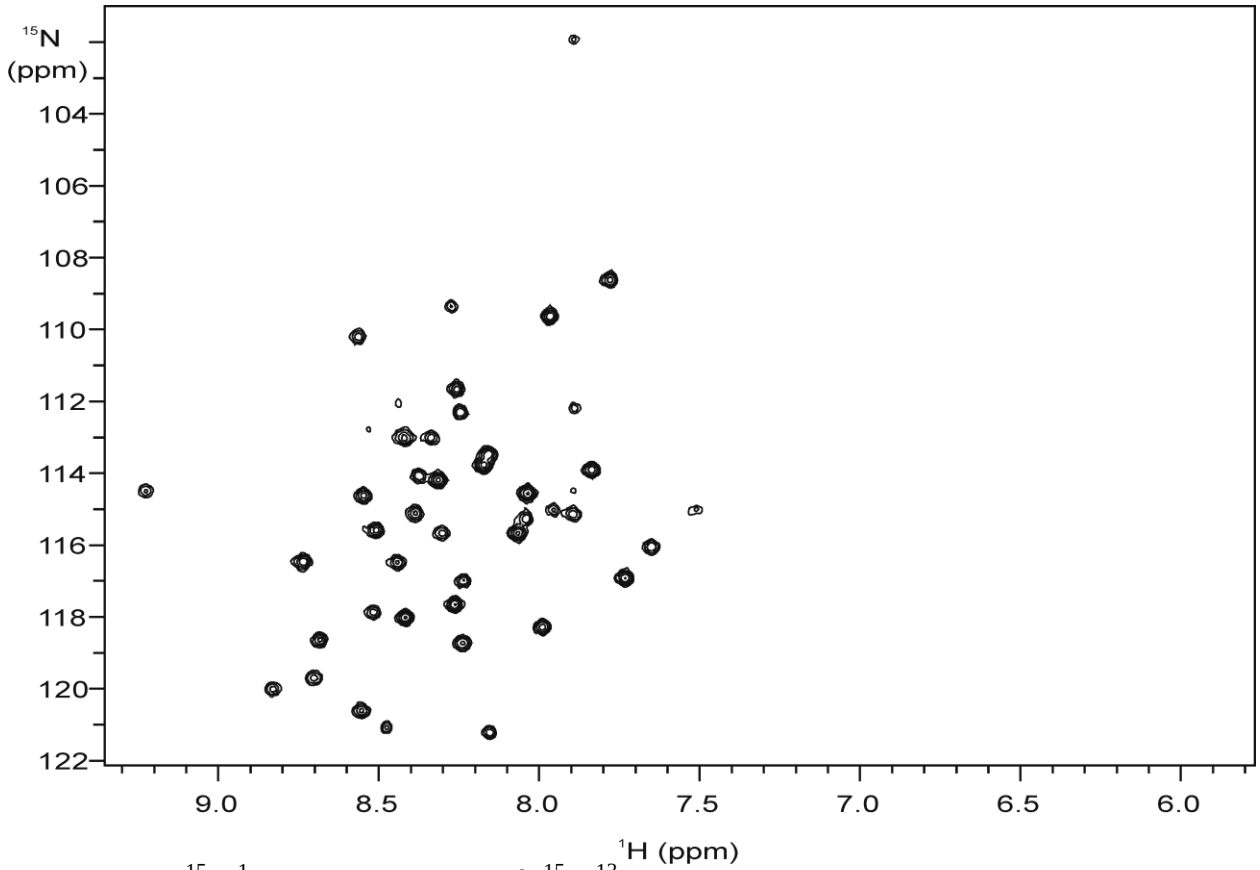


Figure 4-10. [^{15}N,^{1}H]-HSQC analysis of [^{15}N,^{13}C]-TM1-TM2 exchanged in TFE:D_2O(0.1%TFA) (1:1, v:v) overnight at room temperature. The presence of amide crosspeaks indicates that the exchange was not complete after overnight incubation.

Further side chain assignments were made using a ^{15}N-edited NOESY experiment performed with 2 mg of [^{15}N]-TM1-TM2 in TFE-d_2:water. Analysis of this spectrum increased the number of side-chain proton assignments and allowed for short- and medium- range connectivities (±3or 4 residues) to be determined. These connectivities are valuable in secondary structure determinations and are necessary for the calculation of a high resolution structure.

NOESY and TOCSY experiments using unlabeled TM1-TM2 were also employed for refining some sidechain assignments and connectivity data. Different mixing times were used in these experiments in order to elucidate information about different types of residues. The 2D TOCSY experiment was run with mixing times at 60 and 25 msec. The shorter mixing times were used in the observation of resonances coupled through at least 3 bonds whereas the longer mixing times would help to elucidate resonances coupled through more than 5 bonds. There were very few resonances that were separated to be assigned but these experiments were helpful in assigning the Arg and Lys chemical shifts. The 2D NOESY experiments were performed at 300 and 150 msec mixing times to elucidate connectivities originating from side chain resonances such as Glu, Asn and Tyr.

In an attempt to observe additional medium- and long-range connectivities, an exchanged sample was prepared to perform 3D ^{13}C-edited NOESY experiments. Four different ^{13}C-edited NOESY experiments were analyzed: gnoesyChsqc, gnoesyChsqc–D_2O, gnoesyChsqcSE, and gChsqcnoesy. The experiment that had the best peak dispersion and peak intensity, gnoesyChsqc, was run over 3 days. The NOE peaks were assigned and no interhelical connectivities were observed and the connectivities that were observed were similar to the previous ones determined. The NOEs were extracted and added to the constraint files for structure calculations.

The experiments above increased the overall sidechain assignments for TM1-TM2 in TFE:water to 72% of all of the assignable protons. The aliphatic residues were assigned to 91%, but the aromatic residues were more difficult to assign. These residues were only assigned to 53% due to overlap of the resonances. Overall, 84% of all of the expected nuclei were assigned for TM1-TM2 in TFE:water at 25°C.

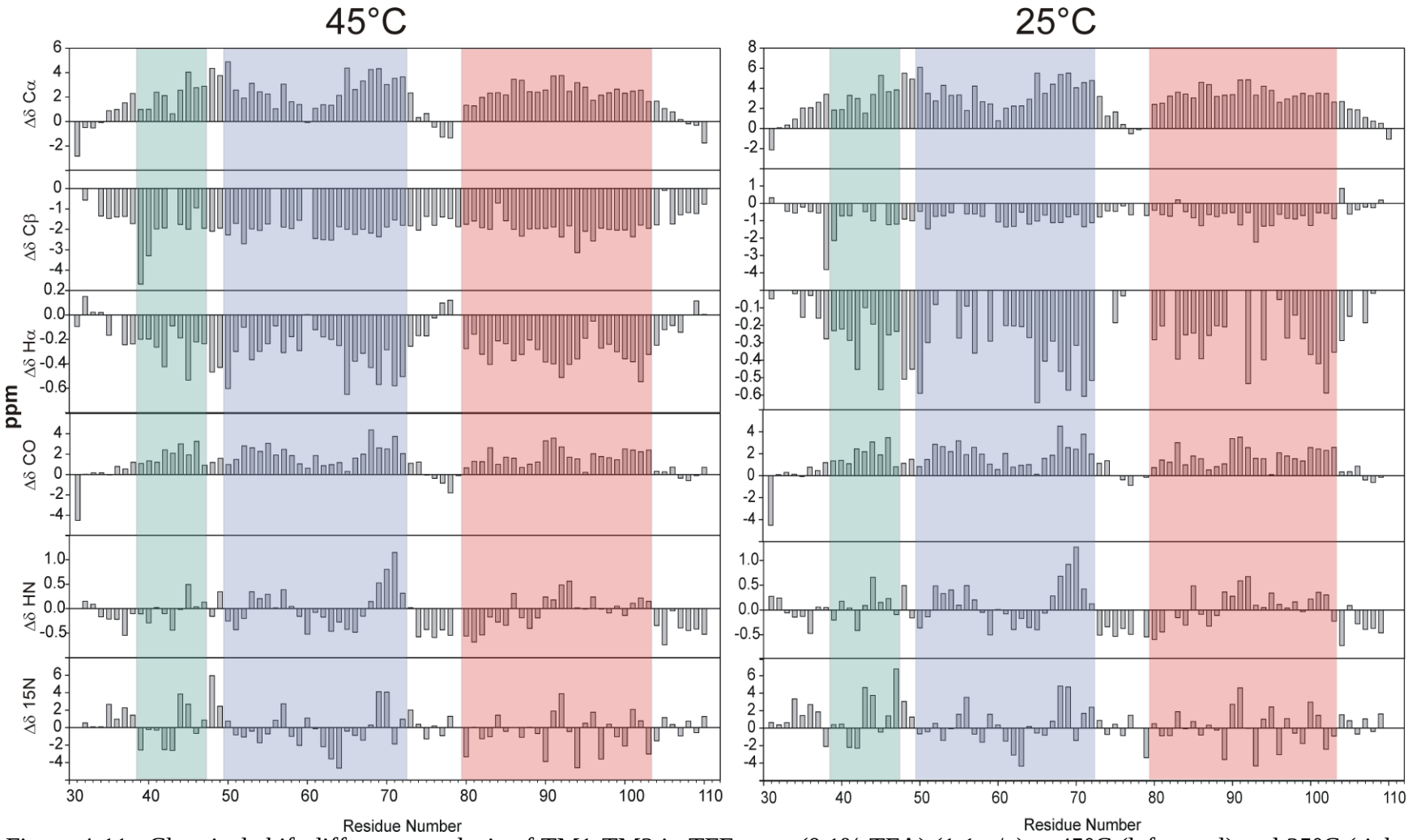


Figure 4-11. Chemical shift difference analysis of TM1-TM2 in TFE:water(0.1% TFA) (1:1 v/v) at 45°C (left panel) and 25°C (right panel). The chemical shifts differences were analyzed by comparison to a random coil database. The boxes indicate the predicted helices from the structure of TM1-TM2 determined in LPPG micelles (*213*):NT helix in green, TM1 in blue, TM2 in red.

Analysis of secondary structure information for TM1-TM2. The secondary structure of regions of the peptide can be assigned by using the chemical shift data derived above. The analysis involves chemical shift difference analysis and considers chemical shift perturbations due to temperature. Chemical shift difference analysis, based on chemical shift indexing (*95-97*), is now widely used for secondary structure analysis at the atom level (*236-240*). The chemical shifts at the NMR nuclei are compared to a library of random coil sequences and the differences are analyzed. When analyzing the Cα and CO chemical shifts, positive deviations from random coil values are indicative of an α-helical conformation whereas negative deviations indicate a β-strand conformation. Conversely, when analyzing the Cβ, Hα, HN and ^{15}N chemical shifts, a negative deviation is indicative of α-helical conformation whereas positive deviations indicate a β-strand conformation. TM1-TM2 appears to be highly helical at both 25°C and 45°C in TFE:water based on the Cα, Cβ, Hα and CO chemical shift differences (Figure 4-11). The helical regions determined in the structure of TM1-TM2 in LPPG micelles (*213*) are indicated by boxes in the figure. The helical regions appear to be similar for both membrane mimetic media and includes the N-terminal helix that was observed in the structure in LPPG micelles and that was predicted by the PHD secondary structure prediction algorithm (see Chapter 3). The differences observed at the Cα atoms appear to be more sensitive than those of the Cβ atoms as judged by changes in the loop region and towards the ends of the peptide which are more obvious in the Cα than Cβ analysis. The difference of the HN and ^{15}N atoms from random coil was not as strongly correlated as the other chemical shifts as observed in the bottom two graphs at both 45°C and 25°C where the differences fluctuate from positive to negative. Finally, there does not appear

to be a significant difference between the secondary structure of TM1-TM2 at 25°C when compared to 45°C.

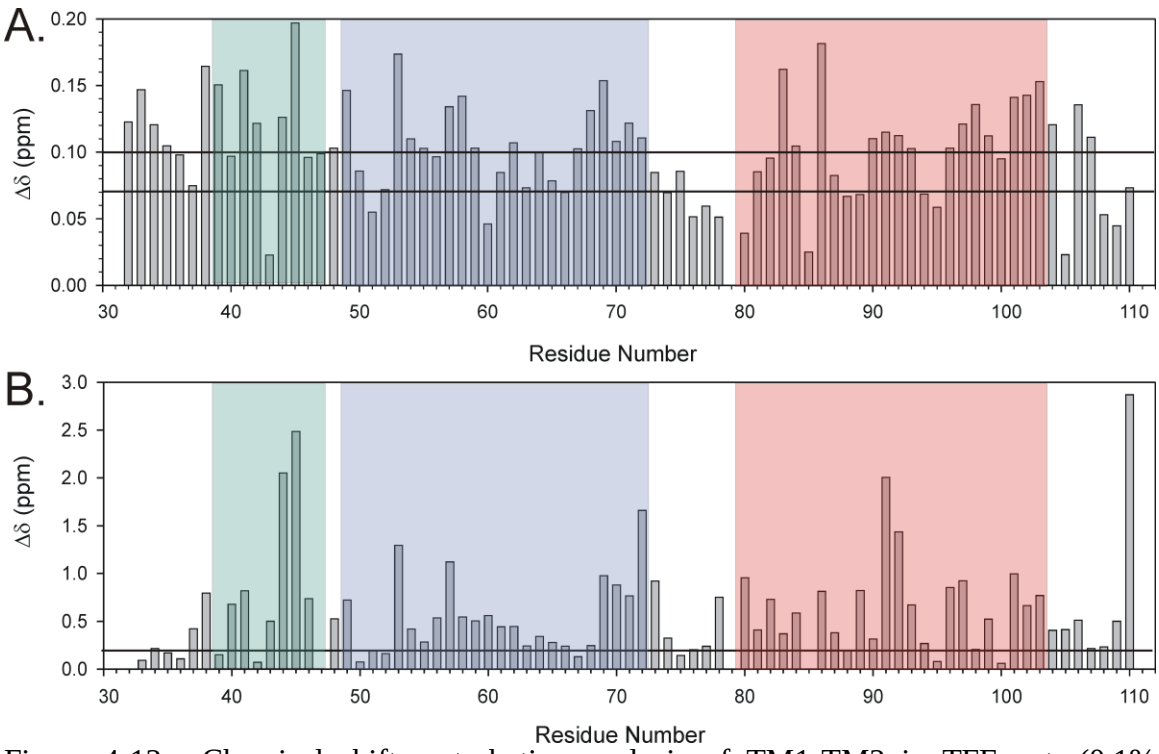


Figure 4-12. Chemical shift perturbation analysis of TM1-TM2 in TFE:water(0.1% TFA) (1:1 v:v). A) Temperture perturbations observed in TFE:water(0.1%TFA). [^{15}N,^{1}H]-HSQC analysis was performed at both 45°C and 25°C and the chemical shift changes were analyzed based on the formula $((\Delta H^2 + (\Delta N/5)^2)/2)^{1/2}$ (*241*). The line at 0.07 is the basis for increased perturbation as defined by (*242*) and the line at 0.1 is the basis for increased perturbation as defined by (*243*). B) Membrane mimetic perturbations. [^{15}N,^{1}H]-HSQC analysis was performed at 45°C in both TFE:water(0.1%TFA) and LPPG micelles and the chemical shift changes were analyzed as described above. The line indicates the largest perturbation observed by temperature in TFE:water(0.1% TFA) (1:1,v:v). The boxes indicate the predicted helices from the structure of TM1-TM2 determined in LPPG micelles (*213*):NT helix in green, TM1 in blue, TM2 in red.

The local environment of individual amino acids can be probed by analyzing the perturbation of the crosspeaks in an HSQC experiment at different temperatures (*241-243*). Chemical shift perturbations determine if a residue is in a structured environment or not. For example, a large perturbation is indicative of a lack of a hydrogen bonding

network that would hold the residue in place, whereas a small perturbation is indicative of a residue which is in a stable hydrogen bonded network as would be seen in an α-helix. In the case of TM1-TM2 when studied at 45°C and 25°C, the perturbations are mainly small with the highest perturbation observed being approximately 0.2 (Figure 4-12A). Although the perturbations observed are larger than the standards in the literature, 0.07 and 0.1 (*242*, *243*), they appear to be relatively small. This is readily apparent when comparing the perturbations of TM1-TM2 in TFE:water and LPPG micelles both at 45°C (Figure 4-12B). The line in Figure 4-12B indicates the highest perturbation observed by changing the temperature in the TFE:water environment and may be used as a baseline for chemical shift perturbations between membrane mimetics (Figure 4-12A). When comparing the two different membrane mimetics at the same temperature, very large perturbations are observed throughout. There is a large perturbation in the N-terminal helix which is hypothesized to interact with the LPPG micelle headgroups but does not have a similar interaction in TFE:water. In TM1, the perturbations are large on the N-terminal side of the GxxxG motif and at the C-terminus of the helix close to the connecting loop. TM2 has large perturbations throughout with the largest perturbation being in the center of the helix. Overall, there does not appear to be a large perturbation in chemical shifts in TFE:water at different temperatures but significant perturbation in chemical shift were found between TFE:water and LPPG micelles at the same temperature. This suggests that the local environment of the amino acids is different in the two membrane mimetics. Using chemical shift indexing, both membrane mimetics favor the same secondary structure. However, both analyses taken together indicate that connectivities between the two helices, if present, may be different.

Dynamics of TM1-TM2 in TFE:water at 45°C and 25°C. Movement of the nuclei in a peptide can give information about the peptide's structure and conformational interactions. The spin-spin relaxation of a peptide can be determined by performing [^{15}N,^{1}H]-HSQC analyses with varying relaxation delays from 0 msec to 210 msec (*244*). The peak volumes in the HSQC spectra were then analyzed and the T2 relaxation time for each residue was determined. If a residue has a relatively longer relaxation time it is presumed to be in a flexible region (*245*). The T2 relaxation time for [^{15}N]-TM1-TM2 at 45°C and 25°C were analyzed (Figure 4-13). Overall, even though the relaxation times appear to be longer at 45°C than at 25°C, the peptide at both temperatures appears to be in a rigid structure except for the N- and C-termini which have relatively long T2 relaxation times. At both temperatures, an increase in relaxation times in the interhelical loop region is observed which would be indicative of a more flexible region. This correlates well with the helices determined in LPPG micelles as indicated by the boxes.

The ^{15}N(^{1}H)-NOE (H-NOE) of a residue can be determined by performing [^{15}N,^{1}H]-HSQC with and without prior saturation of the amide protons (*244*, *245*). This experiment allows the determination of the relative rigidity of different regions of a peptide backbone. The data are plotted as the intensity of the HSQC crosspeak with saturation/ the intensity of crosspeak without saturation. If the intensities are similar then the number will be close to 1 and the residue would be thought to be in a rigid and therefore a structured environment. At 45°C, the H-NOE values are ~0.25 which would indicate that this peptide under these conditions is somewhat flexible (Figure 4-14, open circles). The chemical shift indexing data indicated that helices are present at 45°C, but

the hydrogen bonds are easily perturbed and the helix is not as rigid. At 25°C, the H-NOE values for many residues are closer to 1 (Figure 4-14, filled circles).

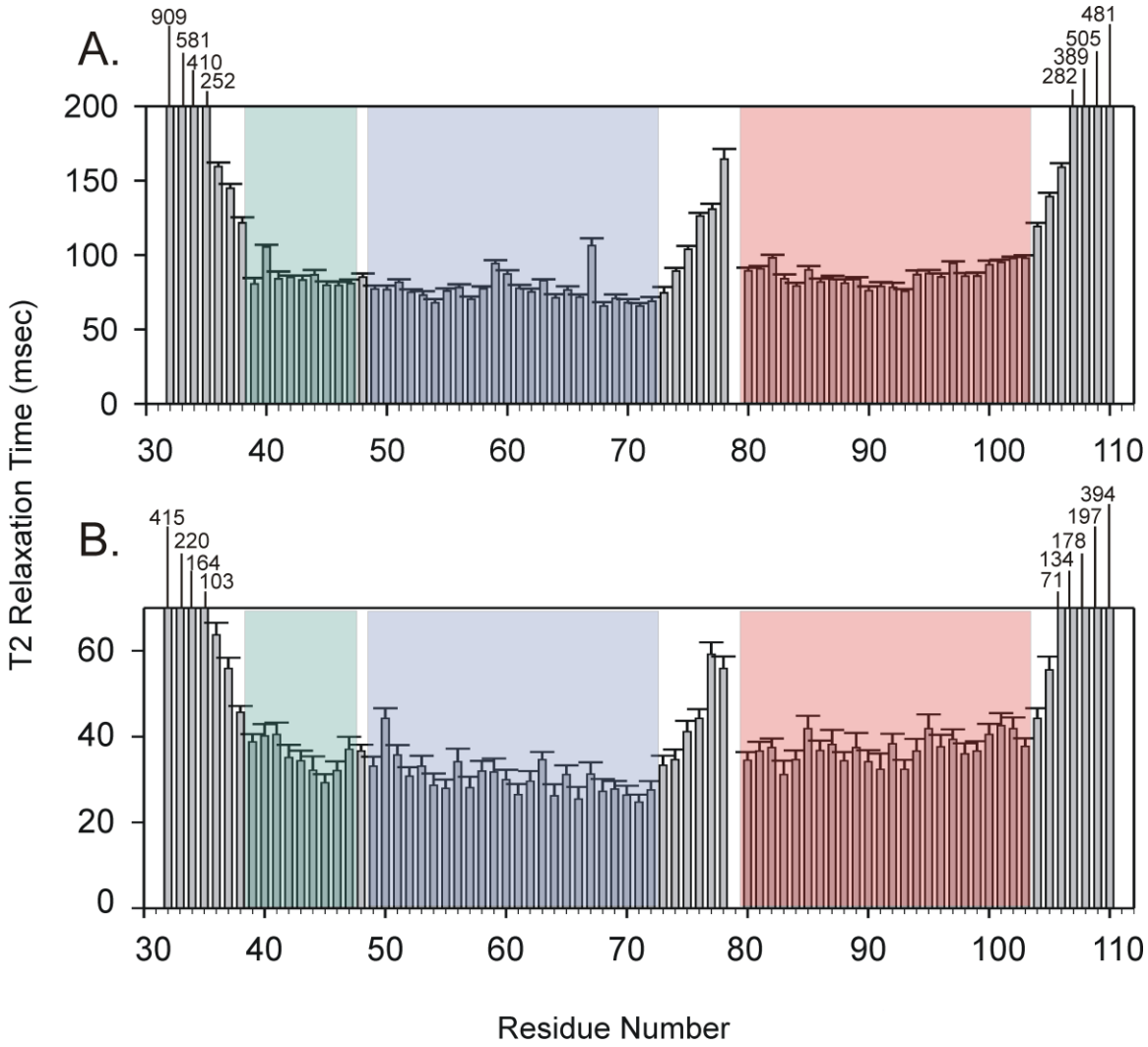


Figure 4-13. T2 relaxation times for TM1-TM2 determined at 45°C (A) and 25°C (B) in TFE:water(0.1%TFA) (1:1 v/v). The T2 relaxation times were determined by performing [^{15}N,^{1}H]-HSQC analyses with varying relaxation delays including 0 msec, 10 msec, 30 msec, 50 msec, 70 msec, 110 msec, 150 msec, and 210 msec. The peak volumes were analyzed and the T2 relaxation time for each residue was determined. The boxes indicate the predicted helices from the structure of TM1-TM2 determined in LPPG micelles (*213*):NT helix in green, TM1 in blue, TM2 in red. The numbers at the top of the graphs are the T2 relaxation times for those longer than 200 msec (A) and 70 msec (B).

The N- and C-termini have very low H-NOE values which would correlate to their lack of secondary structure. The H-NOEs of residues 12-19, 21-31, 35-43, 51-73 are mostly above 0.6 with increased rigidity (H-NOE values >0.75) observed for residues 13-16, 19-21, and 61-63. Increased flexibility within TM1 is observed by lower H-NOE values between residues 32-34 which is just after the GxxxG domain. Furthermore, the loop residues 44 to 49 had HNOE values of less than 0.6 which indicates more flexibility in this region. These data appear to match the H-NOE data determined in LPPG micelles except that the flexible region in TM1 was between residues 26 and 27 which would be within the GxxxG motif itself (see (*213*)).

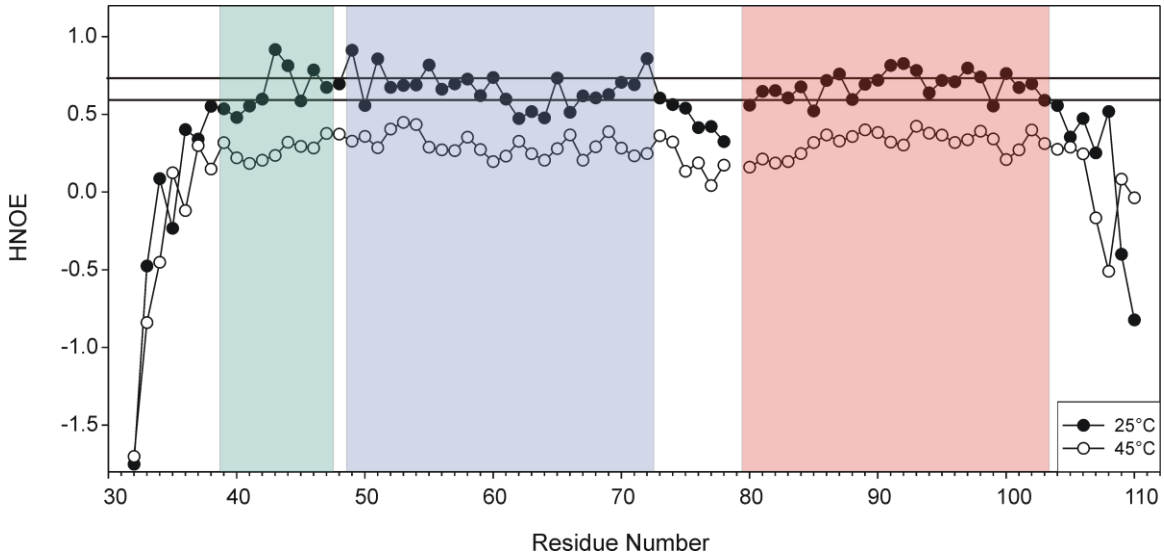


Figure 4-14. H-NOE analysis at 45°C (open circles) and 25°C (filled circles) in TFE:water(0.1%TFA) (1:1, v:v). The H-NOEs were determined by performing [^{15}N,^{1}H]-HSQC with and without saturation before starting the HSQC pulse sequence. The data is plotted as the intensity of the saturated residue/ the intensity of the unsaturated residue. The temperatures for each measurement are indicated by the legend at the bottom. H-NOE values above 0.6 (bottom line) are indicative of rigidity and values above 0.75 (top line) are indicative of increased rigidity. The boxes indicate the predicted helices from the structure of TM1-TM2 determined in LPPG micelles (*213*):NT helix in green, TM1 in blue, TM2 in red.

Hydrogen-deuterium exchange of the amide protons can be used to determine the presence of hydrogen bonds and the accessibility of exchangeable protons to water (*232*, *246-248*). If an intramolecular hydrogen bond is present, it will slow the rate of exchange from hydrogen to deuterium. If there is no hydrogen bond present and/or if the hydrogen bond is to the solvent (i.e. water), then the rate of exchange will be very fast. To study the rate of exchange for TM1-TM2, [^{15}N]-labeled peptide was solubilized to approximately 0.3 mM in fully deuterated TFE-d_3:D_2O(0.1% TFA-d) and a series of [^{15}N,^{1}H]-HSQC experiments were performed at 25°C at multiple time points. The first acquisition began ~15 minutes after sample prep and each HSQC experiment took about 20 minutes to perform. HSQCs were run at 13 different time points (ranging from 15 min to 10 days) (Figure 4-15) and the peak intensities were analyzed to determine the rates of exchange for each residue (Figure 4-16). The first peak to exchange completely was N32 which is in a highly flexible position at the N-terminus of the peptide. Over the next two hours, more peaks began to disappear from the termini (G33, S34, T35, N105, Y106, S107, S108, V109 and T110). The sixth residue of the N-terminus, I36, appears to be exchanging more slowly indicating that this may be the beginning of the N-terminal helix in TFE:water. From residue 37 through residue 104, the exchange rate is relatively slow. The N-terminal helix and TM1 appear to be continuous by this method of analysis as there is no increase in exchange rate between the two. Based on the graph in Figure 4-16, the middle of TM1 exchanges more quickly than either end of the helix. There is faster exchange in the loop region and then the exchange slows down in TM2. The middle of TM2 exhibits very slow exchange which may be indicative of interhelical

connectivities to TM1. A summary of all dynamic properties are presented in Supplementary Table S5 in the Appendix.

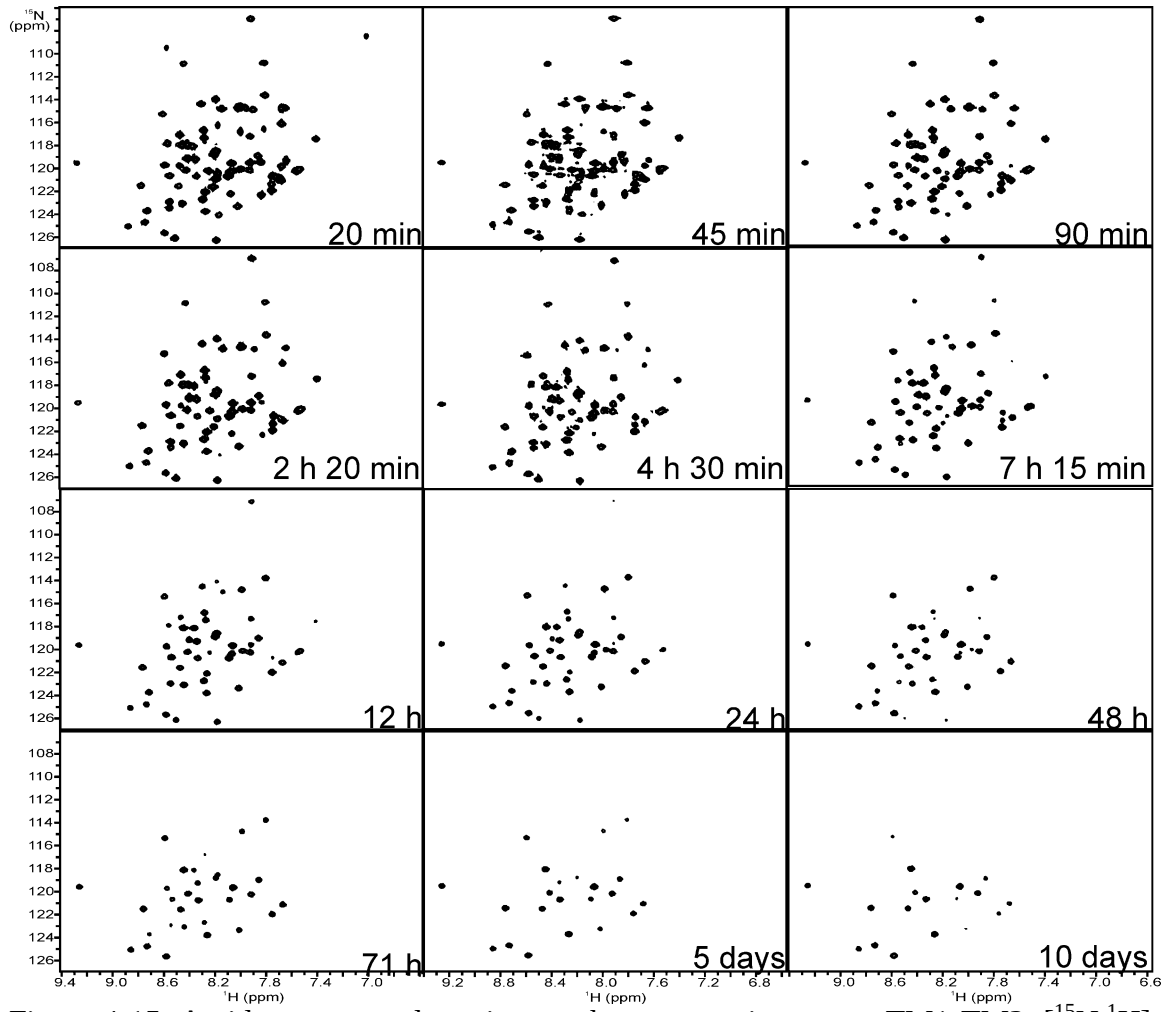


Figure 4-15. Amide proton to deuterium exchange experiments on TM1-TM2. [^{15}N,^{1}H]-HSQC of approximately 0.3 mM [^{15}N]-TM1-TM2 were measured after various times of contact of the peptide with D_2O. The ^{15}N-labeled peptide was solubilized in TFE-d_3:D_2O(0.1%TFA-d) and [^{15}N,^{1}H]-HSQC analysis was performed over time to determine the rate of exchange from N-H to N-D. The analysis was performed at 25°C and the sample was kept at room temperature between the HSQC experiments.

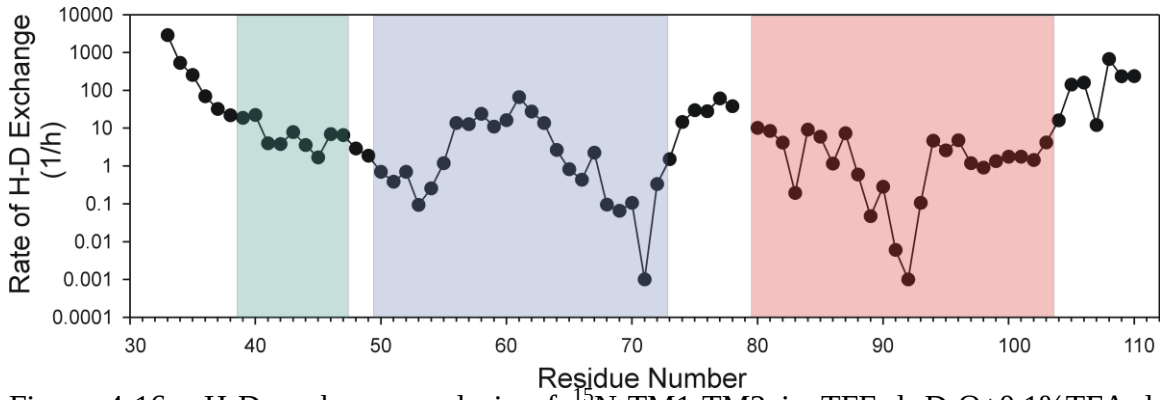


Figure 4-16. H-D exchange analysis of ^{15}N-TM1-TM2 in TFE-d_3:D_2O+0.1%TFA-d. The exchange rates for each amino acid were analyzed by [^{15}N,^{1}H]-HSQC analysis at 25°C at the time points indicated in Figure 4-15. The intensities of the peaks were analyzed and the rates of exchange determined for each residue using NMRViewJ rate analysis. The boxes indicate the predicted helices from the structure of TM1-TM2 determined in LPPG micelles (*213*):NT helix in green, TM1 in blue, TM2 in red.

Temperature coefficient analysis of [^{15}N]-TM1-TM2. In order to determine the effect of temperature on the chemical shifts of [^{15}N]-TM1-TM2 in TFE:water, [^{15}N,^{1}H]-HSQC analyses were performed at 45°C, 35°C, 25°C and 15°C (Figure 4-17). In general, as the temperature is increased the ^{15}N and HN chemical shifts are expected to move towards the high field of the spectrum (*249*). This tendency is observed for most of the crosspeaks, but some move more than others. For example, if we focus on the Gly region of the overlayed spectra, we see that the change in G33 with temperature is very large whereas the changes for G43 and G60 are very small. One interpretation of these results is that G33 is in a highly flexible region of the peptide that contains hydrogen bonds to water molecules, whereas G43 and G60 are in more constrained regions of secondary structure that contain intramolecular hydrogen bonds. This would be consistent with their participation in helical regions of TM1-TM2. G60 is positioned in the predicted TM1 region. G43 is outside of TM1 but is predicted to be in a helical region by the PHD algorithm (Chapter 3) and is in the middle of the N-terminal helix observed in the

structure of TM1-TM2 in LPPG micelles (*213*). The other visible Gly, G56, has chemical shifts that are slightly more spread apart than G60, but not as different as G33, which may indicate increased flexibility. G56 is predicted to be in the center of TM1and it might be involved in a kink in the helix along with G60. The flexibility caused by the kink would correlate to the involvement of residues in this region (R58 and C59; (*19-21*)) in α-factor binding. One caveat to this analysis is that, due to the increased ^{15}N decoupling power needed to perform the HSQC analysis, the lower temperature data may be less accurate due to the extra heat amount generated within the sample volume. We can measure the proton shifts for Ile71 from the 1D ^{1}H spectrum at different temperatures to determine if there is in fact an affect caused by the decoupling. The Ile71 residue is well separated from the other peaks in the ^{1}H spectrum and there is no decoupling during acquisition so this analysis does not require high power. The data is presented in Table 4-3. The changes in chemical shift with increased temperature in the 1D ^{1}H spectra are more linear than in the [^{15}N,^{1}H]-HSQC analysis. This indicates that chemical shift changes measured at the lower temperature of 15°C may not be as accurate in the 2D NMR analysis.

Table 4-3. Chemical shifts of I71 acquired during ^{1}H 1D experiments and during [^{15}N,^{1}H]-HSQC experiments.

Temperature at Acquisition	Chemical shift ^{1}H (ppm)	$\Delta\delta$ ^{1}H (ppm)[a]	Chemical shift [^{15}N,^{1}H]-HSQC (ppm)	$\Delta\delta$ ^{1}H from [^{15}N,^{1}H]-HSQC (ppm)
15°C	9.29	-0.05	9.285/119.62	-0.032
25°C	9.24	-0.058	9.253/119.55	-0.075
35°C	9.182	-0.061	9.178/119.41	-0.06
45°C	9.121	-	9.118/119.28	-

[a]The $\Delta\delta$ ^{1}H (ppm) chemical shift was determined by subtracting the temperature in a given row from the next higher temperature (i.e 45°C-35°C=$\Delta\delta$ in the 35°C row)

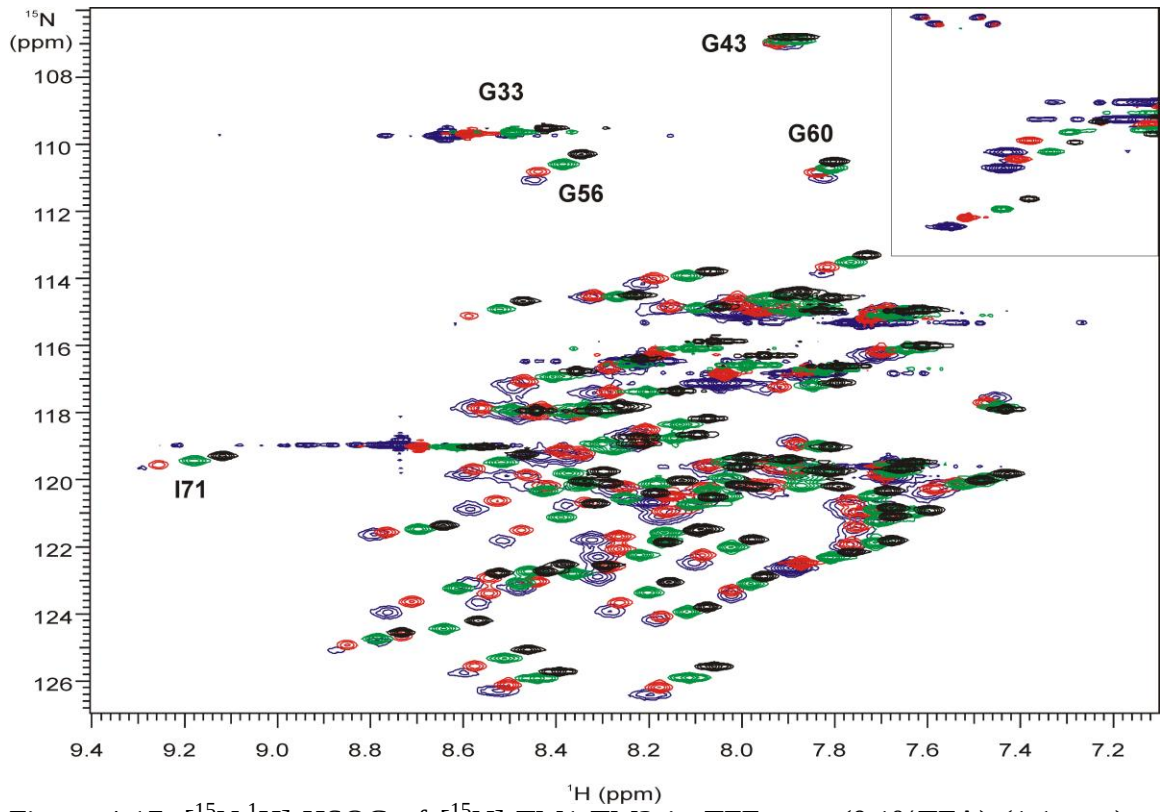


Figure 4-17. [^{15}N,^{1}H]-HSQC of [^{15}N]-TM1-TM2 in TFE:water(0.1%TFA) (1:1, v:v) at various temperatures. The above spectral overlay of [^{15}N,^{1}H]-HSQC experiments were performed on 0.5 mM [^{15}N]-TM1-TM2 solubilized in TFE:water. The peptide was examined by [^{15}N,^{1}H]-HSQC at increasing temperatures 15°C (purple), 25°C (red) 35°C (green) and 45°C (black).

A more quantitative way to analyze the temperature data is to calculate temperature coefficients. Temperature coefficients help to predict secondary structure but may not be accurate if the structure changes abruptly during the course of the experiment (*250*). Since the secondary structure of integral membrane proteins are normally very stable (reviewed in *251*), I do not believe that this caveat would be a problem in my analysis of TM1-TM2. In general, any coefficient more positive than -4.6 ppm/°C can be predicted with >85% certainty to be in a hydrogen bond and residues with coefficients between -5 ppm/°C and -7 ppm/°C can be inferred to be helical. Residues with coefficients more negative than -7 ppm/°C would not be considered to be in a

hydrogen bond network. Based on these criteria large stretches, of TM1-TM2 would be predicted to be helical. However, this treatment was not very successful in distinguishing the hydrogen bonding tendencies of individual residues of TM1-TM2 because even some residues at the chain ends had coefficients more positive than -5 ppm/°C (Figure 4-18). The lack of correlation could be due to the presence of organic media or because the coefficients were generally determined for soluble proteins as opposed to integral membrane proteins.

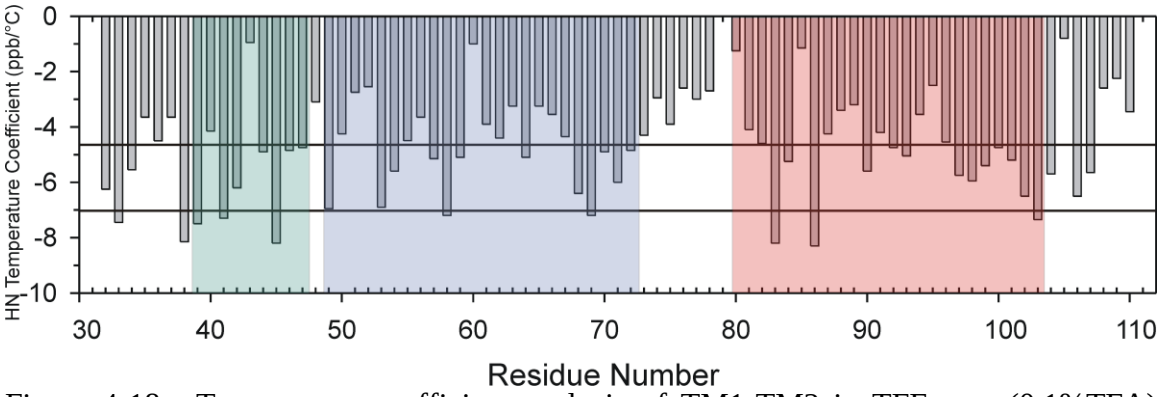


Figure 4-18. Temperature coefficient analysis of TM1-TM2 in TFE:water(0.1%TFA) (1:1, v:v). The temperature coefficients were determined by performing [^{15}N,^{1}H]-HSQC at four different temperatures (45°C, 35°C, 25°C, and 15°C). The boxes indicate the predicted helices from the structure of TM1-TM2 determined in LPPG micelles (*213*):NT helix in green, TM1 in blue, TM2 in red.

Angle constraints determined by TALOS. A program called TALOS can be used to predict backbone dihedral angles of given amino acid residues by performing comparative chemical shift analyses with a subset of proteins from the database (*98*). The program analyzes each residue in the context of a 3 amino acid window and looks for proteins in the database that have similar 3 amino acid chemical shifts. The current database is comprised of mostly soluble proteins which may lead to the misidentification

of a given dihedral angle when the approach is applied to membrane proteins. With the understanding that the angle predictions may not be perfect, the chemical shift data acquired at 45°C and 25°C were analyzed by the TALOS program. The output is shown in Supplementary Tables S6 (45°C) and S7 (25°C) in the Appendix. The number of angles that are defined as "Good" at 45°C is 61 which indicates that these residues are in helical conformations which match the protein structures in the database. If we assume that the rest of the angles which are characterized as "None" and "New" are in a random coil conformation, then we can calculate that at 45°C the TM1-TM2 peptide is 76% helical. It appears that at 45°C there is no distinction between the N-terminal helix and TM1 and that there is a small break in the helix at G60. Based on the TALOS criteria the TM1 helix extends into the intracellular loop as the "Good" angles continue until residue 74. TM2 starts either at residue 80 or 82 because there is an angle defined as "New" at residue F81. The second TM helix extends to residue S104 and the remaining residue conformations are not defined.

The output files that were generated at 25°C tell a slightly different story which may lend itself to describing a different fold in the overall protein structure. The number of angles defined as "Good" at 25°C actually decreases to 81 which would lead to a reduced helical content of 64%. At this temperature there appears to be a shortening of the N-terminal helix (38-41) and a break between this helix and TM1. TM1 begins at L44 and extends through T72 but there is a disruption of the helix at G56 and S59-G60. This could be indicative of a kink formation in this region which is postulated to be biologically relevant (see Chapter 1). Finally TM2 appears to extend from I82 to L103 with no breaks in the helix. These differences may be important in the structure

determination. The stabilization of tertiary structure at 25°C would affect the local structure of TM1 by potentially kinking TM1 which could change the way that TM1 and TM2 interact with each other. If there is a change in the conformation of TM1 this could perturb the relatve conformations of the N-terminal helix and TM1 and there may now be a "break" between this helix and TM1. If the tertiary fold is not stable at 45°C then the N-terminal helix may not be in a different conformation with respect to TM1 and therefore may appear to be continuous.

For the structure calculations at 25°C, the angles that were not labeled as "Good" were removed so that only angles 38-41, 44-55, 57-58, 61-72 and 82-103 were defined in the DYANA processing.

Determination of long-range NOEs between TM1 and TM2. The assignments of the side chain methyl protons of TM1-TM2, obtained on the [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-TM1-TM2, were used to analyze 3D experiments designed to look for NOE connectivities between the selectively protonated residues, i.e NOESY-ct-[^{13}C^{1}H]-HSQC and ct-[^{13}C,^{1}H]-HSQC-NOESY-ct-[^{13}C,^{1}H]-HSQC. These spectra were acquired using a 900 MHz Bruker spectrometer at the NYSBC. The NOE crosspeaks observed in these experiments were assigned and it was determined that they correlated to both inter- and intrahelical connectivities. Analysis of the NOESY-ct-[^{13}C,^{1}H]-HSQC experiment indicated that there were interhelical connectivities between the methyl protons of V69 and I91 as well as intrahelical connectivities in the N-terminal helix between V45 and V49. These connectivities are not observable by ^{15}N-resolved NOESY and were not observed by the ^{13}C-resolved NOESY of the double labeled peptide. Assignments in the

ct-[^{13}C,^{1}H]-HSQC-NOESY-ct-[^{13}C,^{1}H]-HSQC also resulted in the observation of intrahelical NOEs in the N-terminal helix (V45-V49, L41-V45), TM1 (L64-V68) and TM2 (V86-L90) and identified interhelical connectivities between TM1 and TM2. The ct-[^{13}C,^{1}H]-HSQC-NOESY-ct-[^{13}C,^{1}H]-HSQC experiment provided verification of the NOE between V69 and I91 as well as two connectivities between I67 and L88 (Figure 4-19). As described above, this can be explained by the labeling of both methyl groups on the Leu residues.

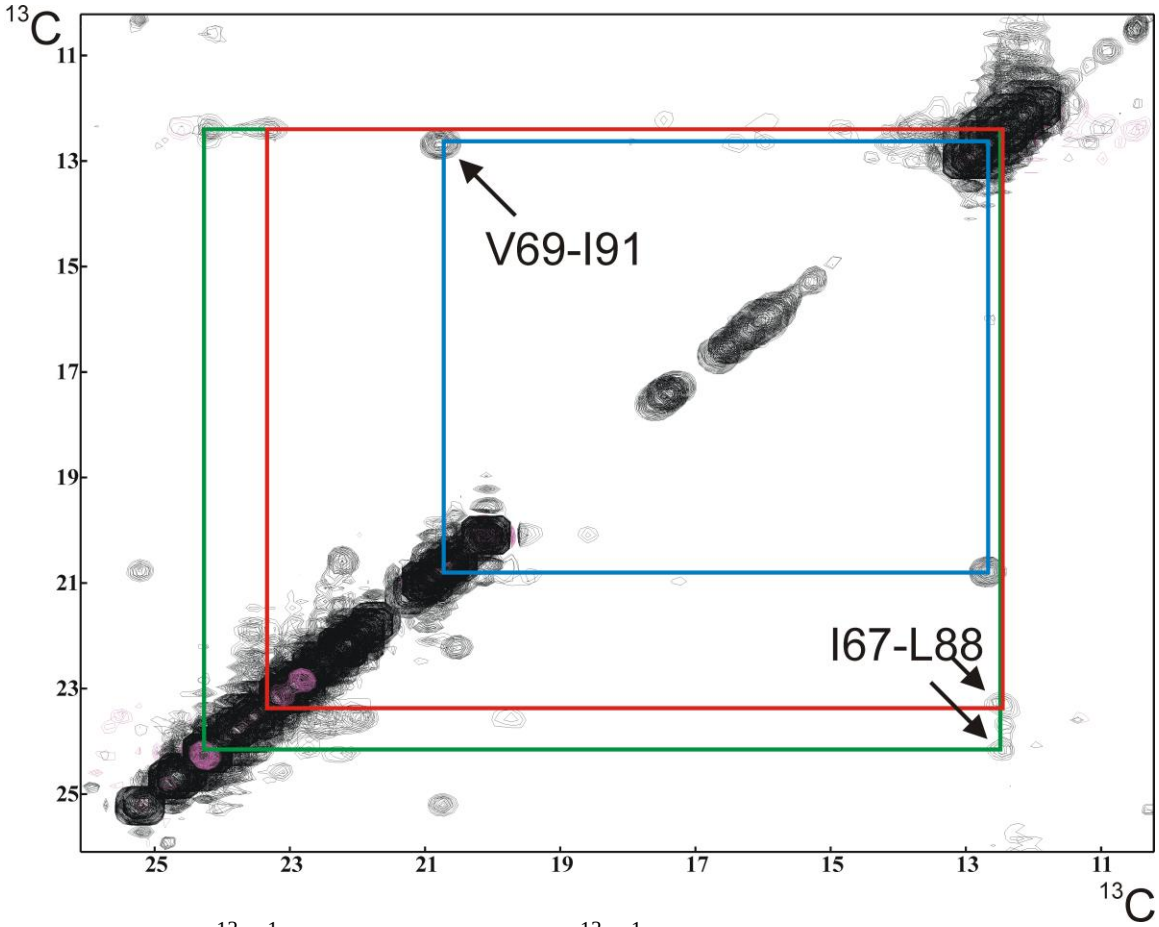


Figure 4-19. ct-[^{13}C,^{1}H]-HSQC-NOESY-ct-[^{13}C,^{1}H]-HSQC NMR analysis with all of the ^{1}H peaks in the same plane. Long-range connectivities were assigned using [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-labeled TM1-TM2 peptide. The interhelical connectivites that were determined are indicated by boxes. The red and green boxes show connectivites between I37 and L58. The blue box indicates connectivites between V39 and I61.

Use of a paramagnetic label on a Cys containing TM1-TM2 peptide to determine more long-range constraints. Paramagnetic labeling can be used to define long-range constraints between regions of a protein. The nitroxide radical that is a part of the "spin-label" can affect the T2 relaxation of residues within ~15 Å of this moiety resulting in broadening and loss of intensity (*252-254*). This effect can be determined by performing [^{15}N,^{1}H]-HSQC analysis on the non-spin labeled and spin-labeled polypeptide followed by a comparison of the peak intensities. If the peak intensity decreases at a residue that is far away in the primary sequence of the peptide then it can be said that this residue is close in space due to the presence of a tertiary fold. As described in the following chapter, I was able to successfully express and purify Met-His$_6$-Ste2p(G31-T110:M54L,C59,M69V,M71I; DE-TM1-TM2-Cys59) in quantities large enough to attach the spin-label and analyze by HSQC (Figure 4-20).

In the comparisons of the peptide with and without the spin-label, it was observed that the majority of peaks on TM1 were broadened to the point where they could no longer be observed by NMR (Figure 4-20, blue circles; Figure 4-21). This result was expected because the radical can affect residues within 10-15Å and the C59 residue is in the center of the helix and the average length of a transmembrane helix is about 26 Å (*255*). Some peaks on TM2 were also affected by the presence of the spin-label indicating that they were also within 10-15Å of the nitroxide radical. The side chain NH groups at residues N84 and Q85 decreased in intensity and were, therefore, within 15 Å of the nitroxide radical. The intensities of the amide backbone crosspeaks of the S95 and A96 residues were also decreased (Figure 4-20, red circles; Figure 4-21). Some of the

residues could not be correlated to the HSQC of TM1-TM2 without the Met-His6-tag and those are indicated by the asterisk in Figure 4-21.

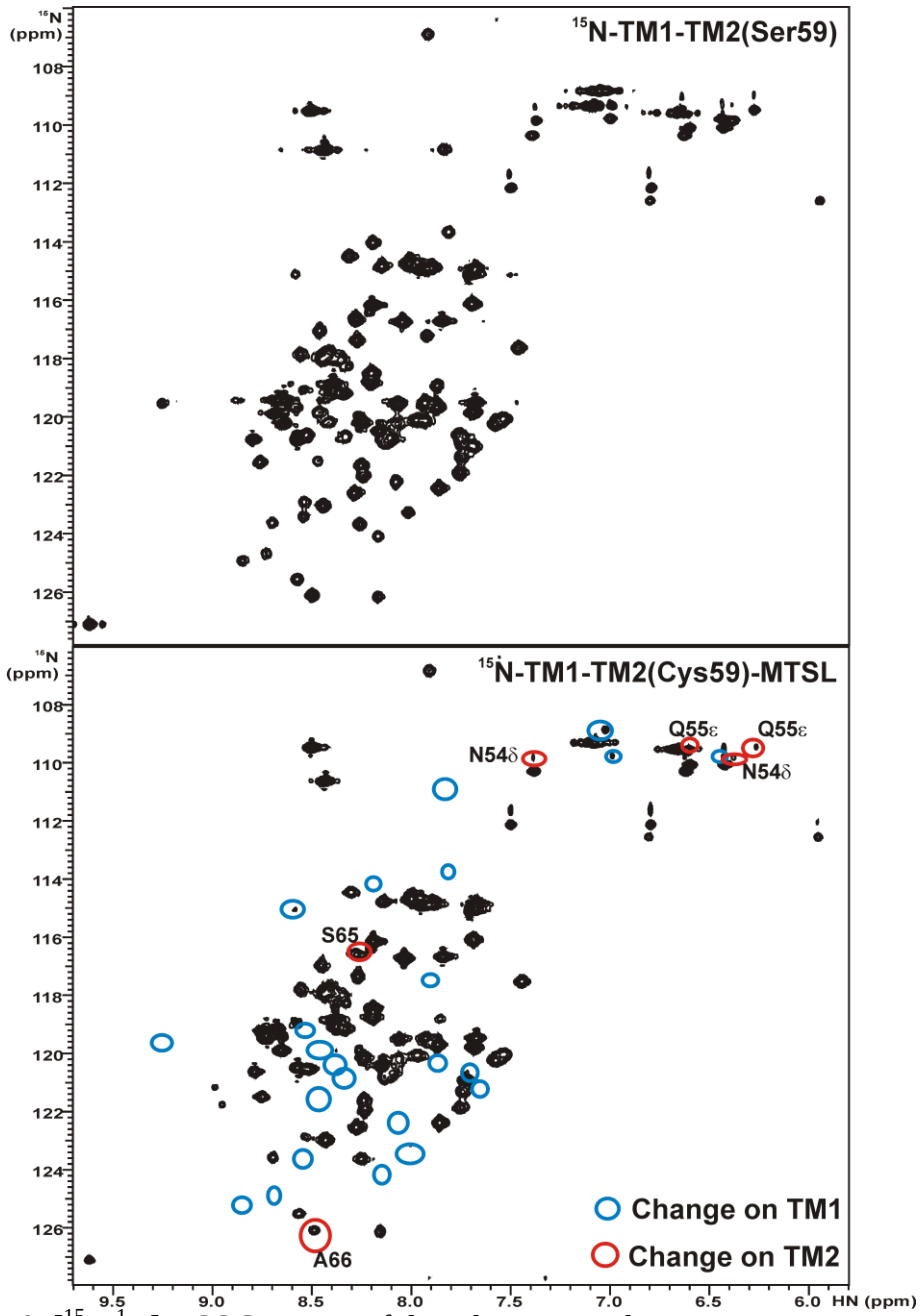


Figure 4-20. [^{15}N,^{1}H]-HSQC spectra of directly expressed TM1-TM2 in TFE:water(0.1% TFA) at 25ºC. The top panel is the Met-His$_6$-Ste2p(G31-T110; M54L,C59S,M69V,M71I) peptide that does not contain a spin label moiety. The bottom panel is the Met-His$_6$-Ste2p(G31-T110; M54L,M69V,M71I) peptide that has the MTSL covalently linked to C59. Blue circles: effects on peaks in TM1. Orange circles: effects on peaks in TM2.

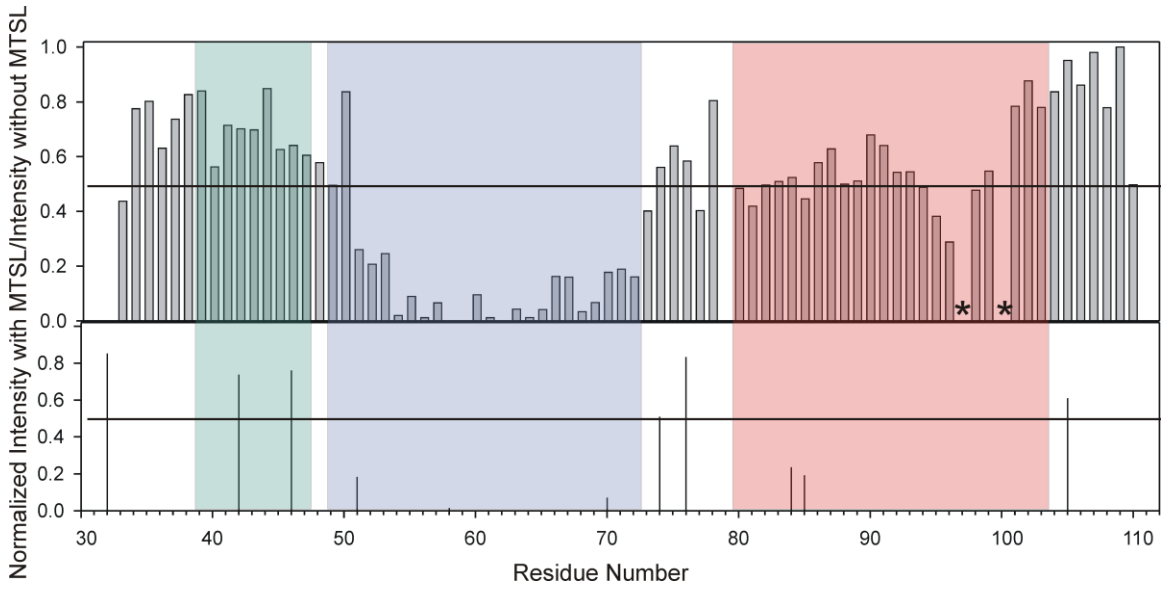


Figure 4-21. Peak intensity affects in the presence of paramagnetic spin label. The peak intensities of [^{15}N]-Met-His$_6$-TM1-TM2 and [^{15}N]-Met-His$_6$-TM1-TM2-C59-MTSL were determined by analysis of the [^{15}N,^{1}H]-HSQC spectra in Figure 4-20. Two residues could not be assigned as indicated by the * and therefore the affect of the MTSL could not be observed. The affect of the paramagnetic spin label was determined by comparing the intensity with MTSL by the intensity without it and then the data was normalized so that the highest ratio was 1. The top panel is the affect on backbone resonances and the bottom panel is the affect on side chain resonances. The line indicates where the ratio is 0.5 and the boxes indicate the predicted helices from the structure of TM1-TM2 determined in LPPG micelles (*213*):NT helix in green, TM1 in blue, TM2 in red.

Structure calculations in DYANA of TM1-TM2. Once the chemical shifts have been determined, and the connectivities have been assigned using NMRViewJ, files are generated for DYANA analysis. The upper and lower limit files are based on the NOE intensities that were extracted in NMRView5 by bin method calibrations. The angle constraints determined from TALOS were also included in the analysis. One hundred randomized structures were calculated for TM1-TM2 in TFE:water at 25°C and the 20 structures with the lowest energies were used for the structure analysis. The connectivities for TM1-TM2 are displayed in Figure 4-22. There are a number of short-range connectivities that are intraresidue connectivites. Furthermore, the residues are all connected by d_{NN} (*i*,*i*+2) connectivities. There is a strong N-terminal helix indicated by the *i*, *i*+3 and *i*,*i*+4 connectivities between residues 36 to 49. The *i*, *i*+3 and *i*,*i*+4 connectivites are not as strong in the putative TM1 and TM2 helices, but there are enough connectivities to predict that TM1 encompasses residues 50 to 72 and TM2 encompasses residues 80 to 101. The overall number of constraints are shown in Figure 4-23. The majority of the constraints are intra-residue and short-range inter-residue connectivities as indicated by the top graph (Figure 4-23A) and the white and light gray boxes in the bottom graph (Figure 4-23B). The presence of long-range connectivities (black boxes in the bottom graph) is indicative of inter-helical connectivities. The total number of constraints used to determine this structure were 1666 which included the long-range NOEs and the restraints determined by the spin label. The calculated structures had ~200 violations of ideal distances or angles. Most of the distance violations are short-range and may be due to flexibility in the peptide in the organic:aqueous environment (see below).

The 20 minimized structures are shown as bundles in Figure 4-24. They were generated by structure calculation in DYANA and then analyzed in Molmol. The top three panels show

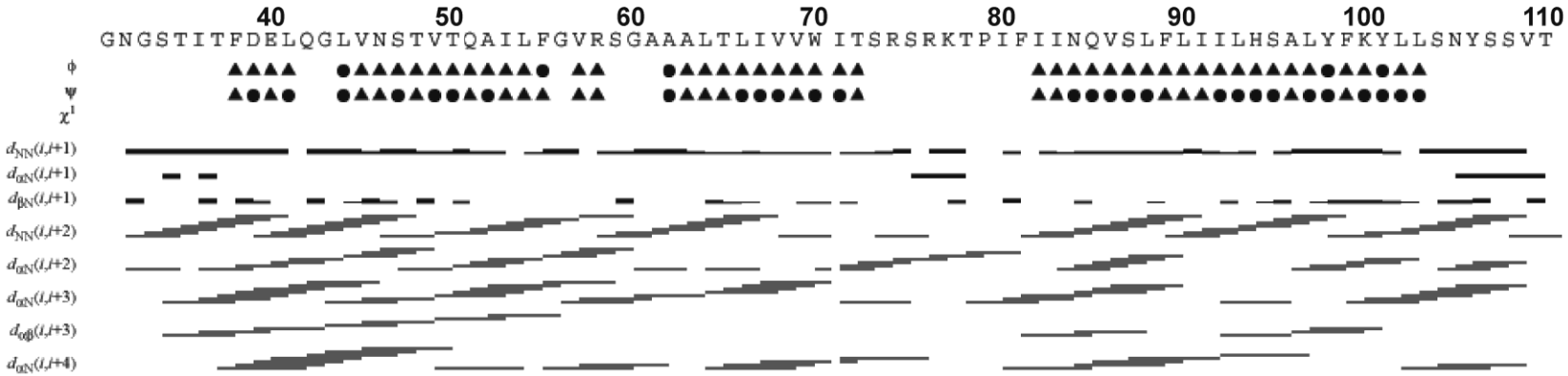


Figure 4-22. NMR connectivities for TM1-TM2 in TFE:water(0.1% TFA) (1:1,v:v). The connectivities were determined by a series of ^{15}N- and ^{13}C-editied NOESY experiments as described in Chapter 4. The TM1-TM2 was either unlabeled, uniformly [^{15}N]- or [^{15}N,^{13}C]-labeled or [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-labeled. The angle constraints used in the structure calculations as determined by TALOS (*98*) are indicated by triangles (torsion angles for ideal α-helix or 3_{10}-helix) and circles (excludes the torsion angles of regular secondary structure) (*109*).

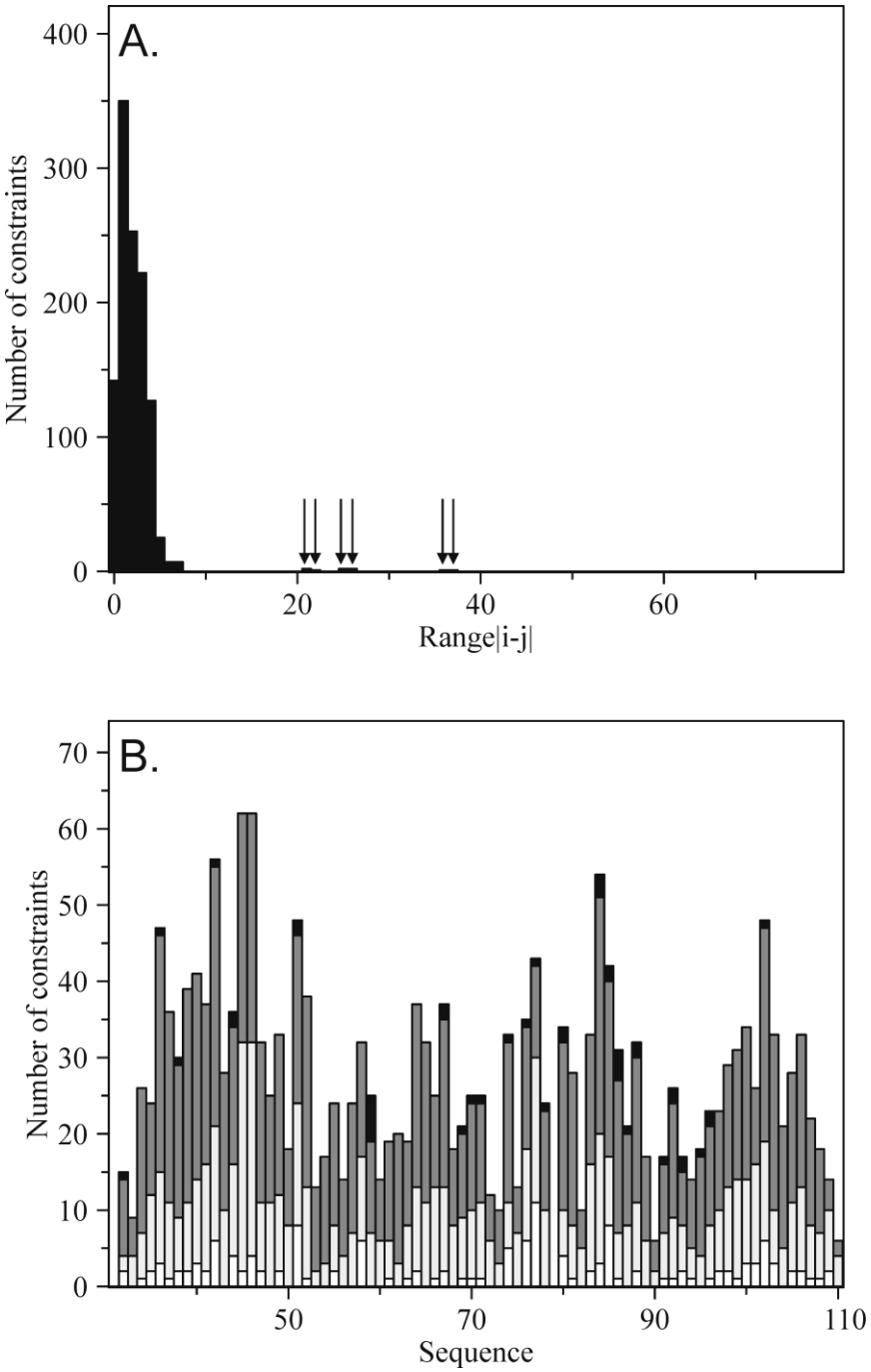


Figure 4-23. The number of constraints determined for TM1-TM2 in TFE:water+0.1%TFA (1:1,v:v). A) The number of constraints plotted versus distance between residues. The arrows indicate where the long-range constraints are on the graph. B) The number of constraints plotted for each residue. The type of constraints are indicated by the color of the bar: white (intra-residue), light gray (short-range), dark gray (medium-range) and black (long-range).

each of the helices as separate entities. As the connectivity data in Figure 4-22 suggested, the region with the tightest helix and the best bundle of structures is the N-terminal helix. This helix overlaps with a backbone RMSD of 1.13±0.52 Å (Table 4-4). The TM1 domain is missing a lot of connectivity data in the middle of the helix. This region contains a GxxxG motif which appears to be causing a kink in the structure. This motif is followed by three Ala residues which seem to lack inter-residue connectivities. Therefore, the backbone RMSD of this bundle is higher than the N-terminal helix at 3.61±0.99 Å (Table 4-4). TM2 has more connectivities throughout but the TALOS angle constraints (Figure 4-22) were often outside of the ideal range for the ψ angle of a helical residue. The backbone RMSD of this helix, 2.60±1.11Å, is between TM1 and the N-terminal helix. The overall structure when aligning residues 50-102 is a folded structure that is closely packed at the bottom and splays apart at the top (Figure 4-24D). The backbone RMSD of this bundle is 6.7±2.5Å. The use of an organic:aqueous solvent probably plays a role in the high RMSD exhibited by my calculated structure. It is possible that the polypeptide chain is rapidly interconverting between a population of closely related but different conformations and this is why the calculated structures do not converge. The presence of water that may solvate the peptide is a major difference between the micelle and organic:aqueous environment. The choice of TFE:water in a 1:1 ratio (v:v) was based on comparisons of the predicted helicity of the peptide region and the CD analysis in both TFE:water and micellar environments. These comparisons suggested that TFE:water (1:1) was the most biologically relevant condition. However, the presence of 50% water which is actually 80% on a mole basis may be playing a role in the lack of structure convergence of TM1-TM2 in this medium. Furthermore, the initial screenings by both CD and [^{15}N,^{1}H]-HSQC do not take helix-helix interactions into account.

Table 4-4. RMSD values for the overlay of 20 calculated structures of TM1-TM2 in TFE:water(0.1%TFA) (1:1,v:v).

	Residues aligned	RMSD bb (Å)	RMSD heavy (Å)
N-terminal helix	36-49	1.13±0.52	1.92±0.57
TM1	50-72	3.61±0.99	4.80±1.09
TM2	80-102	2.60±1.11	3.94±1.22
TM1-TM2	50-102	6.70±2.50	7.64±2.48

Table 4-5. Ramachandran analysis of 20 calculated structures of TM1-TM2 in TFE:water(0.1%TFA) (1:1,v:v).

Regions on Ramachandran Plot	Percent (%)
Favored	70
Additional Allowed	18
Generously Allowed	9
Disallowed	3

Table 4-6. Comparison of TM1-TM2 structures determined in TFE:water at 25°C and LPPG micelles at 45°C.

	Helix Definitions	Long-Range NOE Conectivities	Long-Range Spin-Label from Cys59 Connectivities
TFE:water			
N-terminal helix	36-49	V69-I91	Side chain of N84, Q85
TM1	50-72	I67-L88	Backbone of S95, A96
TM2	80-102		
LPPG			
N-terminal helix	38-47	L66-V86	N/D[a]
TM1	49-72	A63-V86	
TM2	80-103	A63-L90	

[a]N/D, Not Determined.

A ribbon structure was generated for one of the structures generated (Figure 4-25A). From this figure, it is clear that the structure is predominantly helical, but that the helices are distorted based on the absence of red and yellow ribbons as seen in the ribbon structure of the rhodopsin-templated model (*8*) (Figure 4-25B). This is in agreement with the data from the Ramachandran analysis in which 70% of the residues are in the "Favored" region of the plot, whereas 27% are in the "Additonal allowed" and "Generously allowed" regions (Table 4-5). The kink in TM1 appears as a twist in the helix between residues 59 and 65. The interhelical loop at the bottom of the structure appears to have a small turn which may make it smaller and therefore bring the bottom of the structure closer together. TM2 appears to be slightly curved towards TM1. The two helices in the rhodopsin-templated model appear to be close together at the top which may be due to aromatic residue stacking. Due to the fact that I was unable to assign the majority of the aromatic residues, I was unable to include these constraints into my structure calculations.

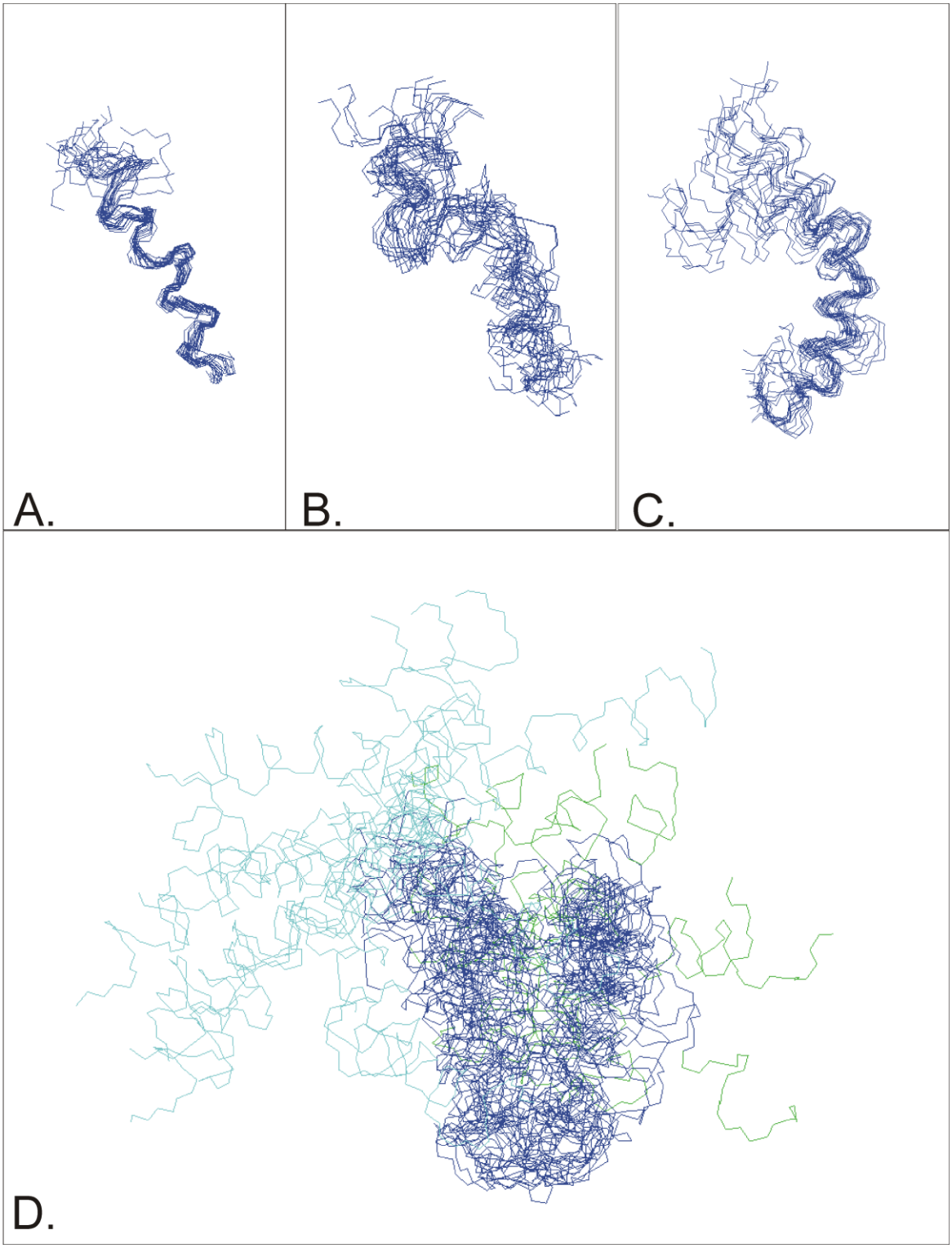


Figure 4-24. Calculated structures of TM1-TM2 in TFE:water+0.1%TFA (1:1,v:v). A) The N-terminal helix is shown from residues 31-49 and was aligned from residues 36-49. B) TM1 is shown from residues 50-72, which were used for the alignment. C) TM2 is shown from residues 80-102, which were used for the alignment. D) The whole TM1-TM2 fragment is shown and it was aligned from residues 50-102. The RMSD values can be found in Table 4-5.

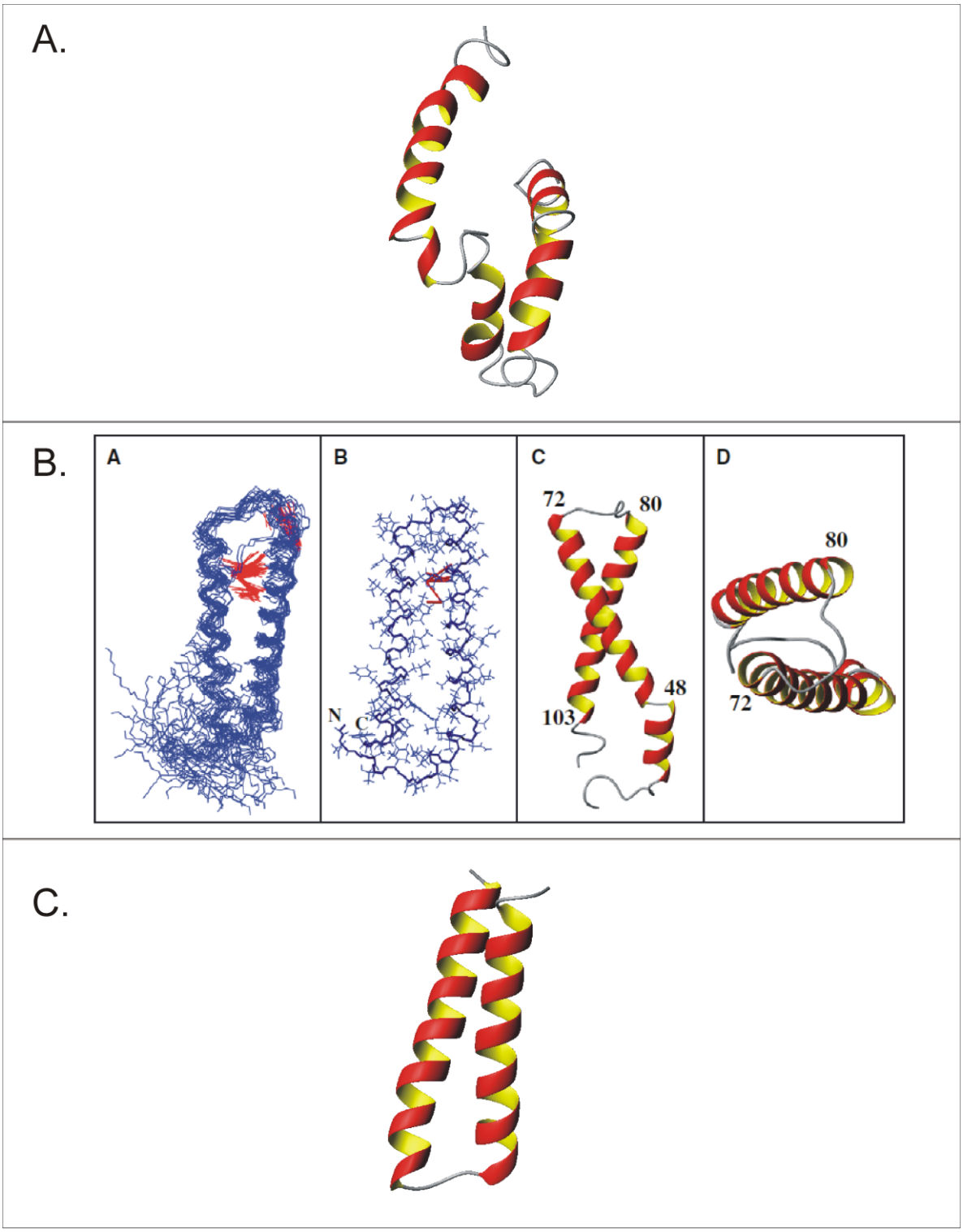


Figure 4-25. Ribbon structure of TM1-TM2. A) The ribbon structure was determined for the 20 calculated structures of TM1-TM2 in TFE:water(0.1%TFA) (1:1,v:v). The lowest energy structure is shown. B) The calculated structure of TM1-TM2 in LPPG micelles taken from Neumoin et al 2009 Figure 5 (*213*). C) The ribbon structure of TM1 and TM2 as determined in the context of the full receptor in the rhodopsin-templated model (*8*).

Comparison of the structure of TM1-TM2 in TFE:water and LPPG micelles. The use of organic:aqueous media to determine the structure of IMPs may be helpful to determine long-range connectivities between helices, as described above. We hypothesized that one could use the backbone structure determined in micelles and or bicelles and overlay the chemical shifts determined in organic:aqueous media to determine a high-resolution structure. To test this hypothesis, we worked with a collaborator at the University of Zurich, Dr. Oliver Zerbe, and his graduate student, Alexey Neumoin, to determine the structure in LPPG micelles (*213*). The structure in LPPG micelles was determined to an overall RMSD of 2.36±0.97 Å for residues 39-103 which encompass the two TM domains (Figure 4-25B). The secondary structure and dynamics of the peptide in the two membrane mimetics appear to be similar as described above. However, the interhelical connectivities are very different. Using an idealized helical wheel representation the connectivities begin on the face of TM1 containing Ser59. In the LPPG structure the interhelical connectivities are close in space (Figure 4-26B). In comparison the interhelical connectivities in the TFE:water structure are more dispersed and involve residues that are now on the opposite face of an idealized helix (Figure 4-26A) This suggests that in these structures there is a twist in the position of TM2 relative to TM1. The acyl chains in the micelle likely provide an environment which stabilizes the relative positions of the helices and reduces the flexibility in the GxxxG motif. Conversely, the organic:aqueous mixture may allow water to interact with regions in the TM domains and, therefore, reduce the stability of the helices and increase the flexibility surrounding the GxxxG motif.

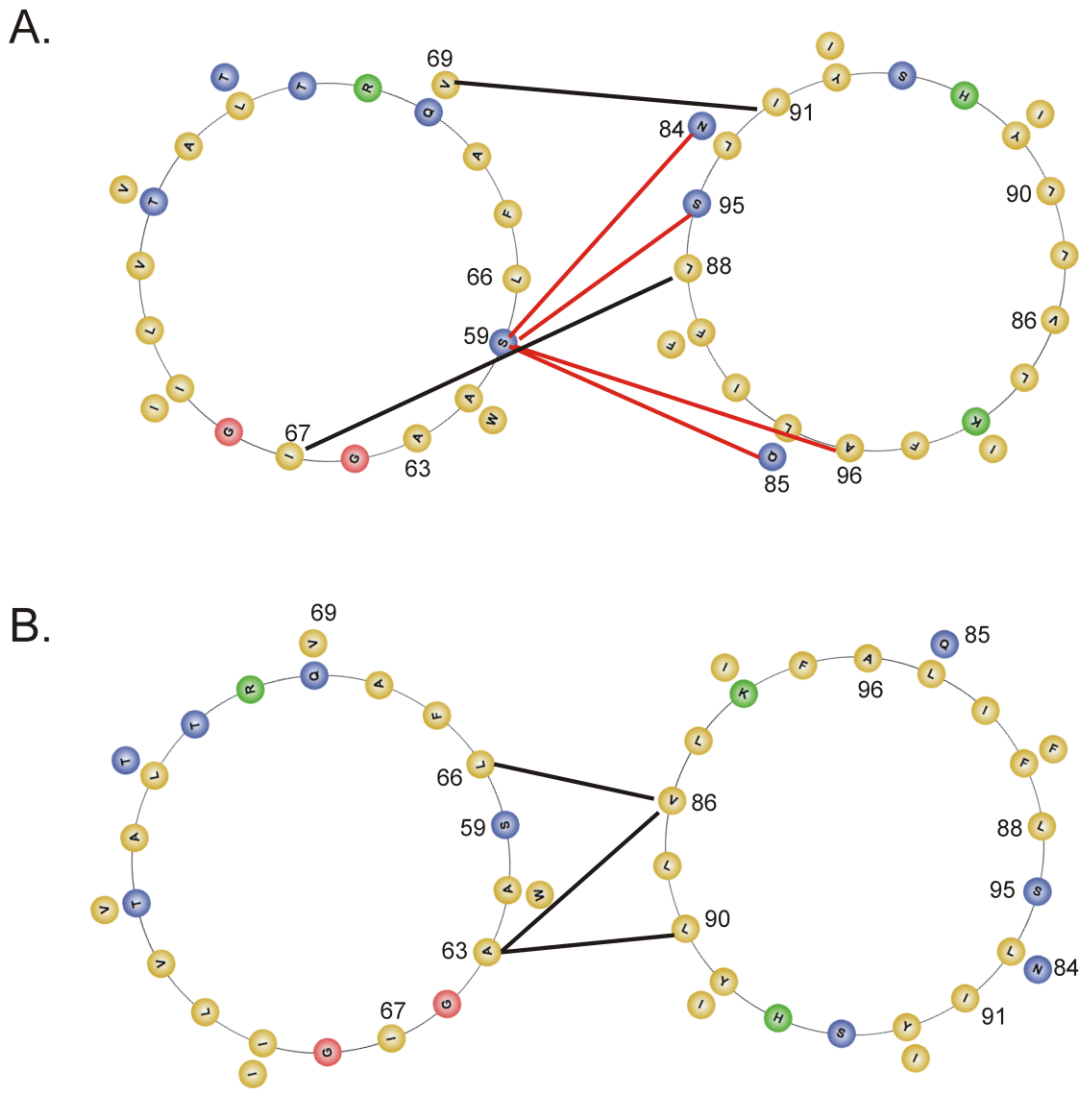


Figure 4-26. Comparison of interhelical connectivities of TM1-TM2 in TFE:water (A) and LPPG micelles (B). Helical wheels were generated at http://www-nmr.cabm.rutgers.edu/bioinformatics/Proteomic_tools/Helical_wheel/ by inputting the amino acid sequence for each TM domain. The sequence for TM2 was put in backwards due to the antiparallel nature of the helices. The NOE connectivities are indicated by black lines and the connectivites determined by spin label broadening are indicated by red lines.

4d. Summary

A structure of TM1-TM2 in TFE:water was elucidated from data acquired by solution NMR at 25°C. The use of a variety of different labeling techniques as described in Chapter 2 and the use of specialized NMR experiments created by the Kay laboratory for selectively protonated methyl groups allowed for assignments to be made for 84% of the observable resonances and over 1500 connectivites were determined including intra-residue, short-, medium, and long-range NOEs. Additional angle constraints were obtained using TALOS and were input into the structure calculation of TM1-TM2. The major feature of the structure is that the polypeptide is twisted into a helical hairpin-like conformation even in organic:aqueous medium. The overall convergence of the structures remains rather poor and there were a large number of violations that resulted during the structure calculations in DYANA. This TM1-TM2 structure is the first example of a GPCR multi-transmembrane domain peptide fragment structure solved in organic:aqueous media. By comparing the structures determined in TFE:water and LPPG, I determined that although both form helical hairpins the peptides fold into different structures and the structure in the micelle is better converged.

Chapter 5

Expression Analysis of Ste2p(G31-T110) Containing Cysteine Residues

5a. Introduction

Secondary structure is an important part of protein folding. The formation of a tertiary fold is dependent on the correct formation of an α-helix or a β-sheet. The information presented in Chapters 3 and 4 show that the helical formation of the two TM domains appear to match the predicted helical regions as determined by secondary structure prediction analysis and hydropathy plots. Unfortunately, the tertiary fold in LPPG micelles and TFE:water is different.

The introduction of disulfide bonds as a means to constrain a tertiary fold has been used with many protein fragments as described in Chapter 1 (*126-131*). For example, our lab is working on generating peptides for use in vaccine generation against HIV-1. Experiments done by our collaborators in Israel indicate that one of the more important determinants for the generation of an active anti V3 vaccine is that the sequence remain in a loop configuration (*256, 257*). To accomplish this two cysteine residues were introduced into the putative N- and C-strands of a V3-like peptide and the peptide was cyclized and found to assume the predicted β-hairpin-like structure. In the case of TM1-TM2, the additional constraint introduced by the two Cys residues might enable the formation of similar structures in both membrane mimetic environments.

Mutation of the plasmid to incorporate two Cys residues included reintroduction of the native Cys at Ser59 on TM1 and the introduction of a second Cys residue in TM2 based on distance modeling using the rhodopsin-templated model by the Konopka group (*8*), Expression

and purification of the peptides were attempted for use in biophysical analyses. In addition, the peptide was expressed with a single Cys at residue 59 for use in paramagnetic label incorporation to aid NMR analysis by examining long-range perturbations as described in Chapter 4. The nitroxide spin label was attached under denaturing conditions which is important when working with a GPCR fragment.

5b. Materials and Methods

Generation of Cys-containing 2TM constructs. The Ste2p gene in the pBEC1 plasmid (*27*) was mutated by a double stranded mutagenesis protocol as described in Chapter 2 by Hee Jung Kim at the University of Tennessee. Biological activity of these proteins was determined by performing Halo assays (*24*, *25*) at the University of Tennessee. Using the pBEC1 based plasmids as the template plasmid, DNA was amplified by PCR using the primers described in Chapter 2. The PCR products were cloned into pSW02 and mutated to remove the Met residues as described in Chapter 2.

Cloning of directly expressed (DE) TM1-TM2 peptides. A primer had been generated that contained the codons for a Met residue followed by a His_6 tag then the beginning of the TM1 sequence (5′-GGAATTCCATATG**CACCATCACCATCACCAT**GGGAATGGATCTACC-3′) (Racha Estephan, unpublished results). The underlined region is the *Neisseria denitrificans* restriction enzyme I (*Nde*I) restriction enzyme site and the bold region is the His_6 sequence. This primer along with the M2 reverse primer from Chapter 2 was used in PCR reactions to amplify the TM1-TM2 region and then it was cloned into the pSW02 plasmid as described. The final DNA sequences were verified by restriction analysis and DNA sequencing.

Expression of directly expressed TM1-TM2 peptides. The plasmids containing the DE-TM1-TM2 sequences were transformed into BL21-AI and expressed in a similar manner as the TrpΔLE containing plasmids. Some optimization was performed as described in the Results section below.

Purification of TM1-TM2 fusion proteins and directly expressed peptides. The DE peptides were purified by RP-HPLC using a Zorbax 300SB-C3 column as described in Chapter 2. Acetonitrile:water gradients with 0.1% TFA and either 10% or 20% isopropanol were used to purify the DE peptides at 60°C. Details are discussed in the Results section.

Analysis of fractions from the IB preparation by analytical HPLC. A small amount of TFA was added to each fraction because the column works better under acidic conditions. When the TFA was added to the lysis mix, the lysis buffer and the mild detergent fractions, a white precipitate formed. These samples were centrifuged at high speed for 15 minutes. The supernatant was loaded straight onto the column and the pellet was solubilized in TFE:water (1:1) to a volume equivalent to the original volume prior to TFA addition.

Attachment of (1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) Methanethio-sulfonate (MTSL) to ^{15}N-DE-TM1-TM2-Cys59. This method is based on the method described in (*253*) with some modifications. The ^{15}N-DE-TM1-TM2-Cys59 peptide was solubilized in 100 μL 6M GnHCl, 0.1 M NH_4OH buffer, pH 8.5 by sonication, but it was cloudy and clumpy so 100 μL TFE was added and the mixture sonicated until clear. The MTSL label (10 μL) that had been resuspended in acetonitrile to 200 mM final concentration was added. The reaction incubated at room temperature and was followed over time by analytical HPLC and the MW of the product analyzed by ESI-MS.

5c. Results

Determination of residues that would be useful in disulfide formation screening. Using the rhodopsin-templated model of Ste2p (*8*) the distance was measured from the -SH group of Cys59 to the nearest side chain on TM2. The smallest distance, 2.44 Å was between Cys59 and Ile92 (Figure 5-1B). The distance between Ile91 and Cys59 is 4.66 Å (Figure 5-1A). The side chains of the other residues in this region do not face the Cys59 side chain and so their distances were not determined. The distance between the residues would be increased when the native residues are replaced by the Cys residues (Cys91 9.08 Å, Cys92 4.41 Å). The mutation Ile92Cys should not interfere with disulfide bond formation, but the mutation Ile91Cys may be too far apart to form a disulfide bond. All of the distances were determined in Rasmol and the amino acid modifications were made in Molmol.

Full-length Ste2p constructs were generated with the native Cys at residue 59, a mutated Cys252Ser, and a single Cys mutation in TM2 (I91C, I92C, L93C, or H94C) (Kim, Hee Jung, Unpublished Results). Halo assays were performed by Hee Jung Kim at the University of Tennessee and all of the constructs appeared to have similar activity when compared to WT Ste2p, except for I91C (Figure 5-2). This mutant protein exhibited similar activity to WT at high concentrations of α-factor, but at low pheromone concentrations did not have as large a halo. The mutant I92C was slightly more active than the WT protein and therefore may be a good candidate for the analysis of disulfide bond formation.

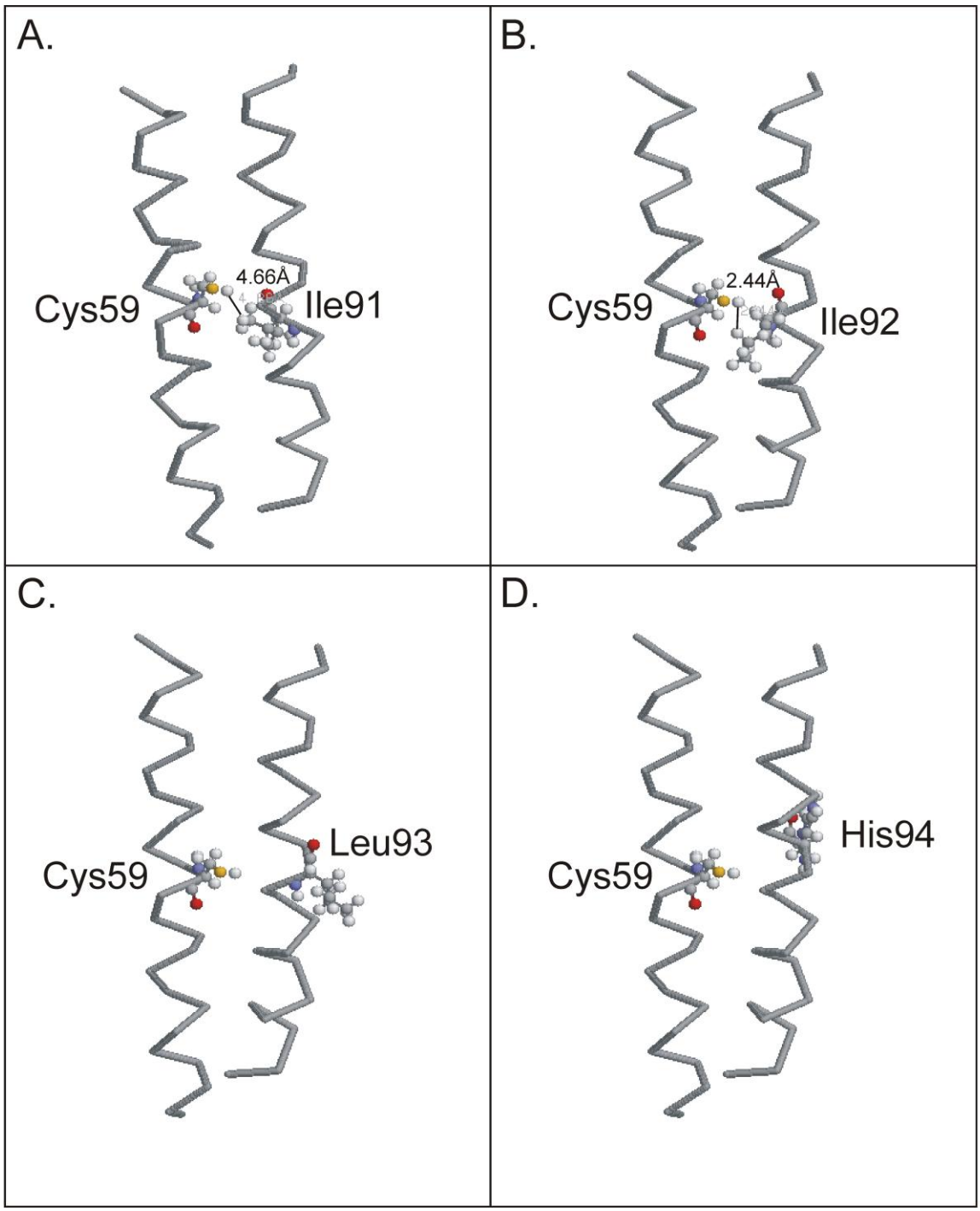


Figure 5-1. Backbone structures of the TM1-TM2 helices. Rasmol was used to determine the distances between Cys59 and Ile91 (A) and Ile92 (B). The positions of the side chain residues Leu93 (C) and His94 (D) were on the opposite face of the TM2 helix relative to TM1 so the distances were not measured.

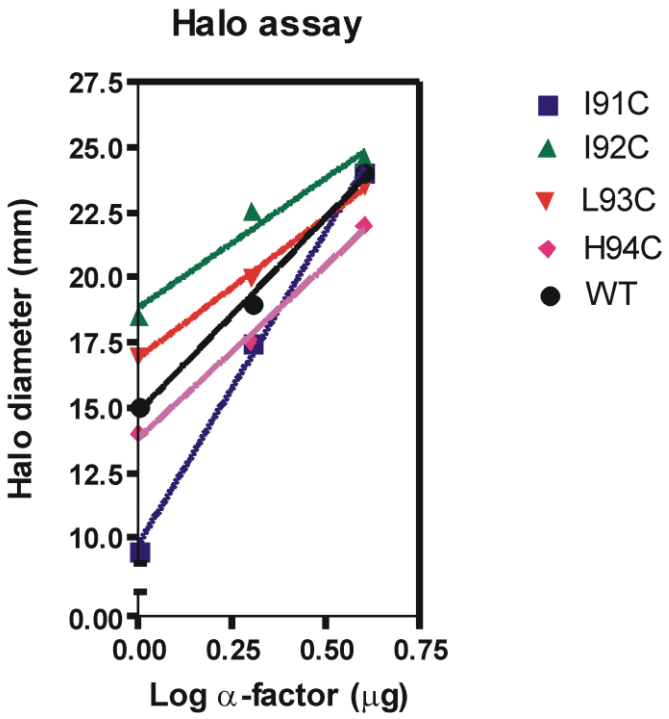


Figure 5-2. Plot of halo diameter vs. the amount of α-factor added to the disk. All constructs contained a mutation in the Ste2p gene coding for Cys252Ser. Ile91Cys, Ile92Cys, Leu93Cys and His94Cys mutations were incorporated into the Cys252Ser background. The WT contains Cys at position 59 and 252. The growth arrest activity as a diameter of growth inhibition of each construct was assayed and compared to WT.

Table 5-1. Sequence and molecular weights of TM1-TM2 Cys containing fusion proteins and directly expressed peptides.

					Molecular Weight (Da)[c]			
Protein Name	Cys Residue	Plasmid[a]	Amino acid sequence[b]	Label Incorporated	Calculated	Observed	% Incorporation[d,e]	Yield (mg/L)
TM1-TM2-FP	59	pLC13	MHHHHHHHHH-(TrpΔLE)-M-GN-GSTITFDELQGLVNSTVTQAILFGVR**C**GAAALTLIVVWITSRSRKTPIFIINQVSLFL**IILH**SALYFKYLLSNYSSVT	^{15}N	22619	22612	97%	N/D
	59/91	pLC14		^{15}N	22608	22596	96%	N/D
	59/92	pLC15		No label	22319	N/D	N/A	N/D
	59/93	pLC16		No label	22319	N/D	N/A	N/D
	59/94	pLC17		No label	22293	N/D	N/A	N/D
DE- TM1-TM2	NoCys	pLC24	MHHHHHH-GNGSTITFDELQGLVN-STVTQAILFGVR**C**GAAALTLIVVWITSRSRKTPIFIINQVSLFL**IILH**SALYFKYLLSNYSSVT	No label	9705.34	9704.5	N/A	N/D
				^{15}N	9825.34	9819.1	95%	4.2
	59	pLC25		No label	9721.41	9719.5	N/A	N/D
				^{15}N	9841.41	9834.9	95%	1.4
	59/91	pLC26		No label	9711.39	9710.1	N/A	N/D
				^{15}N	9831.39	9826.9	96%	N/D
	59/92	pLC27		No label	9711.39	9710.56	N/A	N/D
	59/93	pLC28		No label	9711.39	9710.78	N/A	N/D
				^{15}N	9831.39	9826.77	96%	2.6
	59/94	pLC29		No label	9685.39	N/D	N/A	N/D

[a] Plasmids pLC15, pLC16 and pLC17 were generated but were not analyzed for expression when the CNBr cleavges failed.
[b]The bolded and underlined residues are the ones that were mutated to Cys.
[c]Calculated MWs are given for reduced Cys residues.
[d]Percent Incorporation of ^{15}N-labeled amino acids was determined using the calculated MW.
[e]N/A is not applicable because no isotopes were incorporated, N/D was not determined for the samples indicated.

Cloning and expression of the Cys containing TrpΔLE-TM1-TM2 fusion proteins in BL21-AI. The plasmids generated for the activity assays were used as templates for subcloning of TM1-TM2 Cys containing peptides in the TrpΔLE plasmid for expression. These plasmids were transformed into BL21-AI and the proteins were expressed in ^{15}N-labeling minimal media (Figure 5-3). Based on the SDS-PAGE analysis, TM1-TM2-Cys59/Cys92-FP, TM1-TM2-Cys59/Cys91-FP and TM1-TM2-Cys59-FP were expressed at levels similar to that of Cys-less TM1-TM2-FP. Expression of TM1-TM2 Cys59/Cys93-FP was not as high. Nevertheless, in all cases sufficient fusion protein was expressed for cleavage and isolation. The ^{15}N-TM1-TM2-Cys59/Cys94-FP sample was lost so its expression could not be analyzed on this gel, but previous analysis indicated that the expression levels would be the same as the Cys92 and Cys91 containing constructs (data not shown).

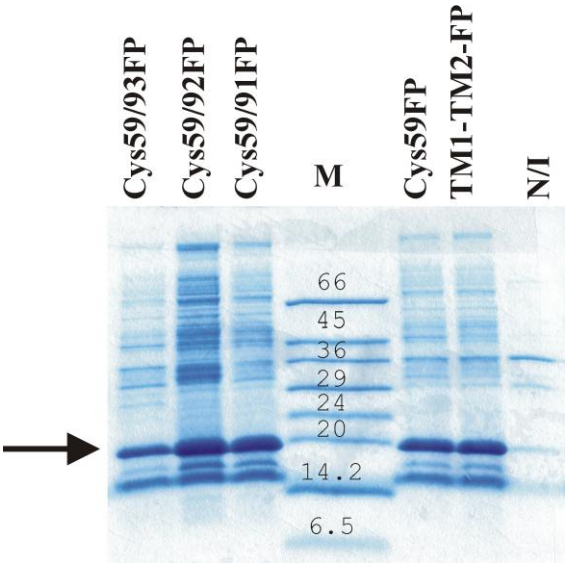


Figure 5-3. Large scale inductions of Cys containing TM1-TM2 fusion proteins. Plasmids containing the Cys containing DNA sequences were transformed into BL21-AI and expression was induced in ^{15}N-labeling media as described in Chapter 2. Lane 6 contains ^{15}N-TM1-TM2-FP as a control to determine growth conditions. The arrow indicates the protein at the expected MW.

Cleavage of TrpΔLE-TM1-TM2-Cys59 by CNBr in the inclusion bodies. Inclusion bodies were generated for the fusion protein containing a single Cys residue at position 59. As described in Chapter 2, the inclusion bodies were solubilized in TFA and cleaved at room temperature with CNBr. The cleavage reaction was then injected onto a Zorbax 300SB-C3 column and the peaks were collected for MS analysis (Figure 5-4). The fusion protein (Retention time=35.5 min) slowly decreased in peak height with time (data not shown). However, the desired TM1-TM2-Cys59 peak could not be identified.

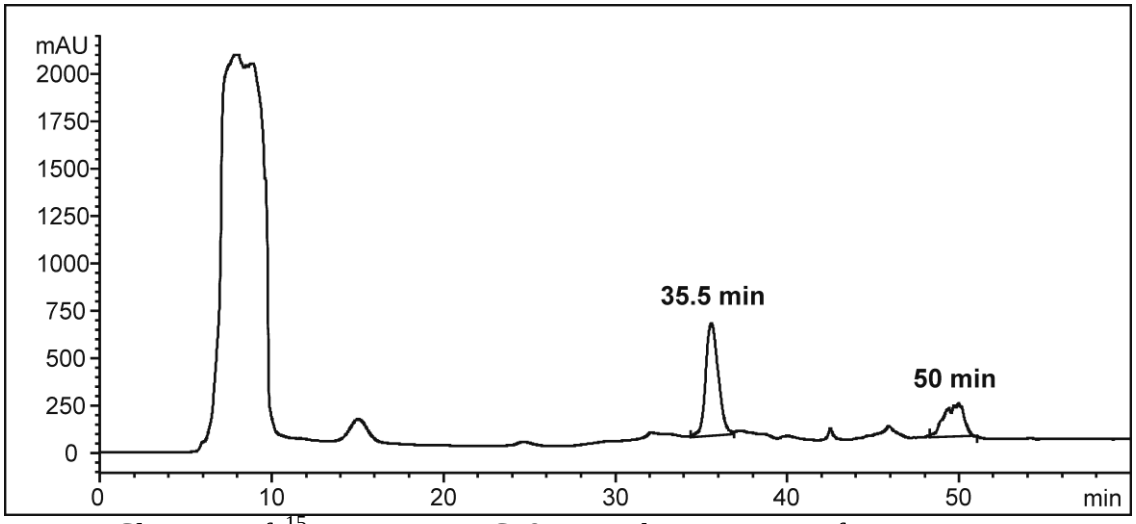


Figure 5-4. Cleavage of ^{15}N-TM1-TM2-C59-FP. The TM1-TM2 fusion protein containing a single Cys at position 59 was cleaved in the inclusion bodies. The reaction was loaded onto a Zorbax 300-SB C3 Prep-HT column and separated by a 36-72% acetonitrile:water gradient with 10% isopropanol and 0.1% TFA at 60°C with a flow rate of 5 ml/min. Peaks were collected and analyzed for the presence of TM1-TM2-C59.

Purification and cleavage of TrpΔLE-TM1-TM2-Cys59 and TrpΔLE-TM1-TM2-Cys59/91 by RP-HPLC. Our lab has had success in the past with performing CNBr cleavage on a fusion protein of TrpΔLE and a Cys containing 73-residue peptide, though at a significantly reduced yield compared with the Ser containing analog (*79*). To purify the TM1-TM2-Cys59-FP I used a 40-72% acetonitrile:water gradient with 20% isopropanol and 0.1% TFA. The fusion protein

eluted at 37 minutes (Figure 5-5A) and the MW was verified by ESI-MS (Figure 5-5C and Table 5-1, $MW_{expected\ for\ 100\%\ incorporation}$=22619, $MW_{observed}$=22612) which would indicate that the fusion protein contained 97% ^{15}N atoms. An injection of 70% TFA just after the purification indicated that there was no fusion protein left on the column (data not shown).

After purification of the ^{15}N-TM1-TM2-Cys59-FP, I attempted to purify one of the ^{15}N-TM1-TM2-Cys59/91-FP (Figure 5-5B). The fusion protein eluted at 36.5 minutes and the incorporation on ^{15}N atoms, based on ESI-MS, was 96% (Figure 5-5D, $MW_{expected\ for\ 100\%\ incorporation}$=22608, $MW_{observed}$=22596). The samples that were purified above were lyophilized and then different cleavage protocols were attempted.

Three different methods were used to attempt to cleave the fusion protein from the Cys containing peptide (data not shown). The first method was the method described in Chapter 2 except that we used purified fusion protein. A second method was performed in formic acid over a longer period of time (*258*). The sample was solubilized in formic acid to a final concentration of 70% formic acid. CNBr was added to the reaction mixture to a final concentration of 20 mg/mL and the reaction mixture was then incubated at room temperature for 16 hours. The third reaction was a modified method that was originally presented by Staley and Kim (*133*). The two cysteine containing fusion protein was allowed to oxidize by air prior to cleavage. In all three methods, the reaction products were treated with dithiothreitol prior to purification. I was able to verify the presence of the peptide by ESI-MS in all cases, but the amount of peptide purified was very low and would not be useful for further analysis (data not shown). In the formic acid cleavage reaction, the peptide was formylated and multiple peaks appeared on the MS data (data not shown). I concluded that CNBr cleavage of the Cys containing peptides would not provide enough peptide to perform characterization studies so I switched to a direct expression strategy

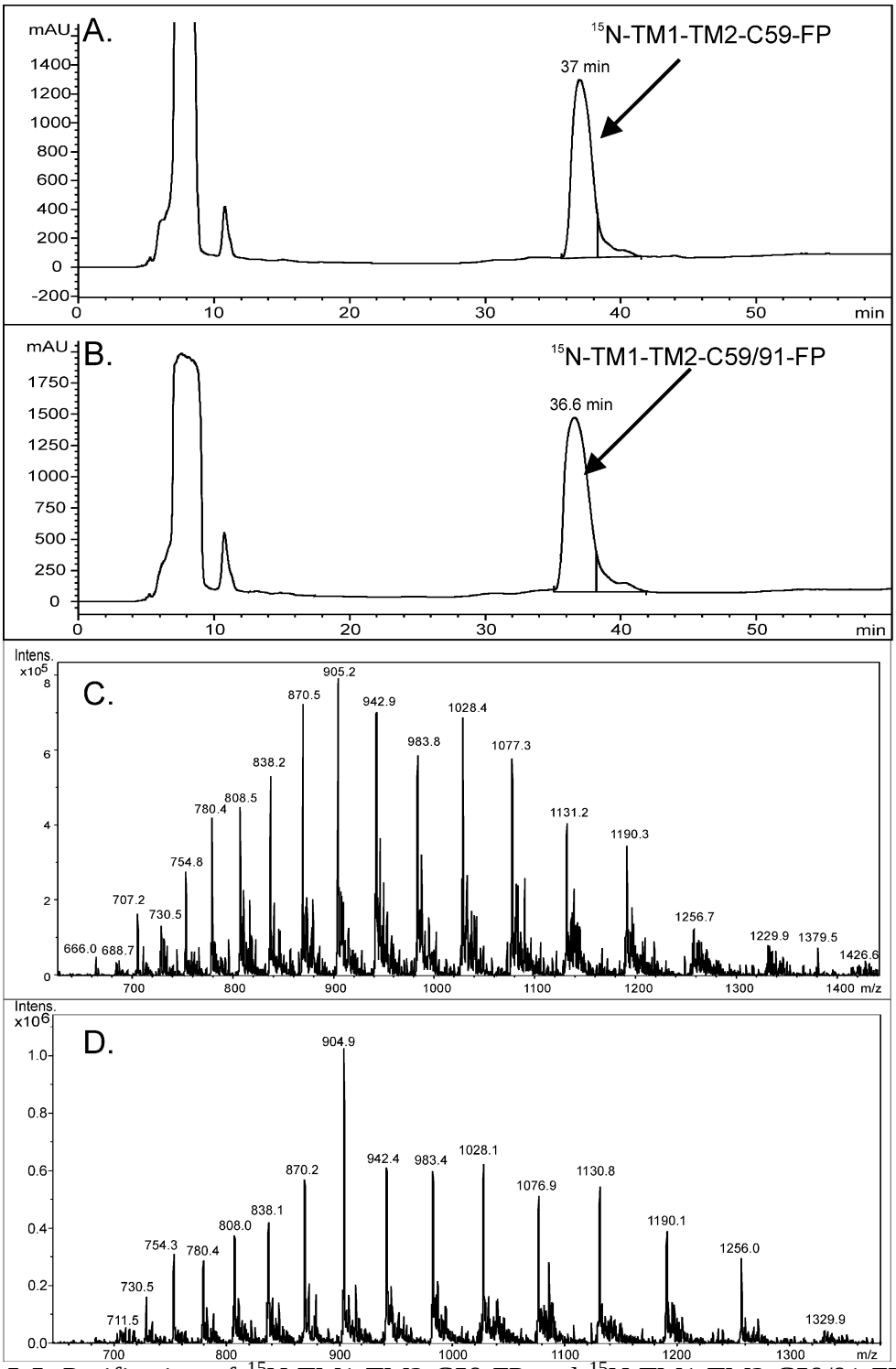


Figure 5-5. Purification of ^{15}N-TM1-TM2-C59-FP and ^{15}N-TM1-TM2-C59/91-FP. The fusion proteins containing a single Cys residue at 59 (A, C) and two Cys residues at positions 59 and 91 (B, D). The proteins were purified on a Zorbax 300-SB C3 Prep-HT column and separated by a 40-72% acetonitrile:water gradient with 20% isopropanol and 0.1% TFA at 60°C with a flow rate of 5 ml/min. The collected fractions were analyzed by ESI-MS. The observed MWs were 22612 for C59 (C) and 22596 for C59/91 (D).

Cloning and expression of the directly expressed TM1-TM2 proteins in BL21-AI. A way to circumvent the need to release the target peptide from the fusion protein is to express it directly (*167, 168*). This may lead to low levels of expression due to cell toxicity as originally observed for several Ste2p TM peptides (*6*). A small scale growth was performed using BL21-AI cells containing the plasmid that expresses directly expressed (DE) TM1-TM2-Cys59/91. The expression was induced in minimal media as above and the IBs from the 1 ml fractions were examined by SDS-PAGE (Figure 5-6A). The directly expressed protein appears as a slightly smeared band which runs just below the 14 kDa marker protein. The Lysis Mix and Mild Detergent fractions of the 1 mL IB preparations were also analyzed to verify that the directly expressed protein was in the inclusion bodies fraction (Figure 5-6B). In fact, the majority of the protein was observed in the inclusion body fraction as indicated by the band next to the arrow in the last lane. When the IBs were generated for the remaining 50 mL culture pellet, all of the supernatant fractions were collected and analyzed for the presence of DE-TM1-TM2-Cys59/Cys91. There appears to be peptide in all of the lanes (Figure 5-6C). Analytical HPLC in combination with MS analysis allowed me to identify the retention time and MW of the peptide (Figure 5-7). Further analysis to determine which fractions contained the peptide was performed by analytical HPLC of the ^{15}N-DE-TM1-TM2 IB isolation (see below).

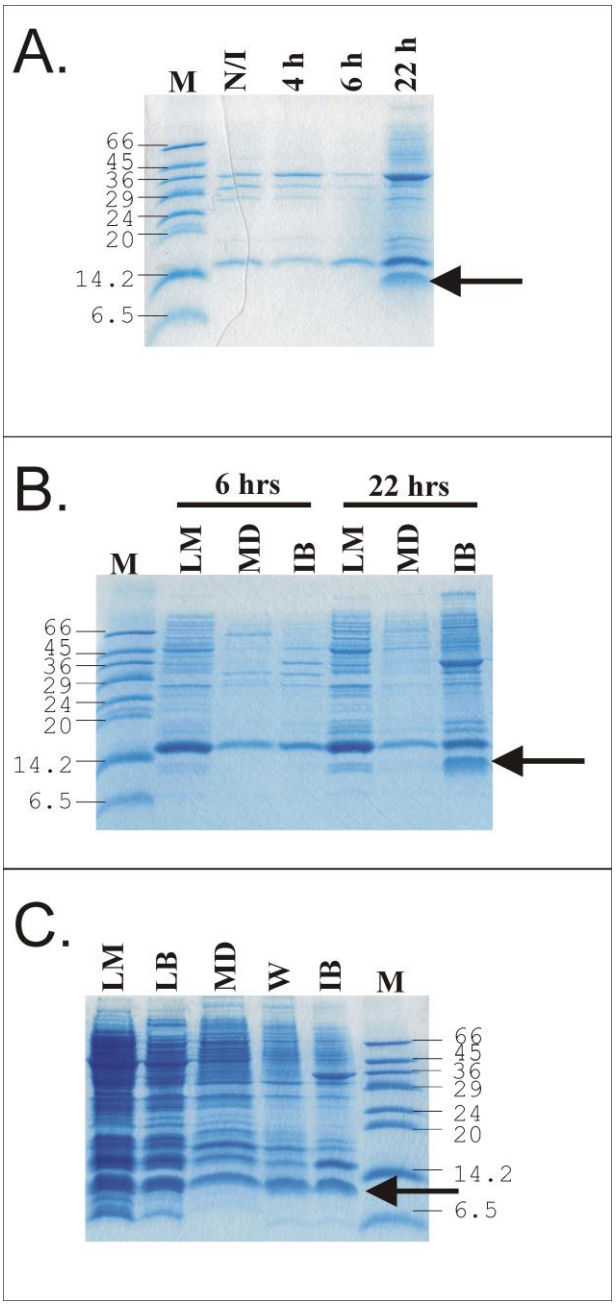


Figure 5-6. Small-scale induction of DE-TM1-TM2-C59/91. A) A plasmid directly expressing TM1-TM2 with two Cys residues at positions 59 and 91 was transformed into BL21-AI and expression was induced in minimal media as described in Chapter 2. Expression was followed over time. The time points are indicated at the top of the gel. The arrow indicates the protein at the expected MW. B) Analysis of all fractions of the small-scale IB preparation at 6 hours and 22 hours induction time. C) Analysis of fractions from the large-scale IB preparation at 22 hours induction time. LM Lysis Mix; LB Lysis Buffer; MD Mild Detergent; W Water; IB Inclusion Bodies.

Large and small scale expressions of DE-TM1-TM2 and DE-TM1-TM2-Cys59 were performed in 500 mL of ^{15}N-labeling media and 50 ml LB and expression was induced as described above (Figure 5-8A). There were high levels of expression of both unlabeled and ^{15}N-labeled DE-TM1-TM2noCys as seen in lanes 3 and 7 of Figure 5-8A). Compared to the noCys peptide, the expression decreases when a Cys is present at residue 59 (compare lane 5 to lane 3 and lane 9 to lane 7, Figure 5-8A) and the expression of the Cys peptide is increased in minimal media relative to expression of the Cys peptide in LB. All of the fractions collected during the IB preparation were also evaluated. A protein band that corresponded to the MW of the DE-TM1-TM2 was found in the mild detergent and the water fractions as well as the IBs (Figure 5-8B).

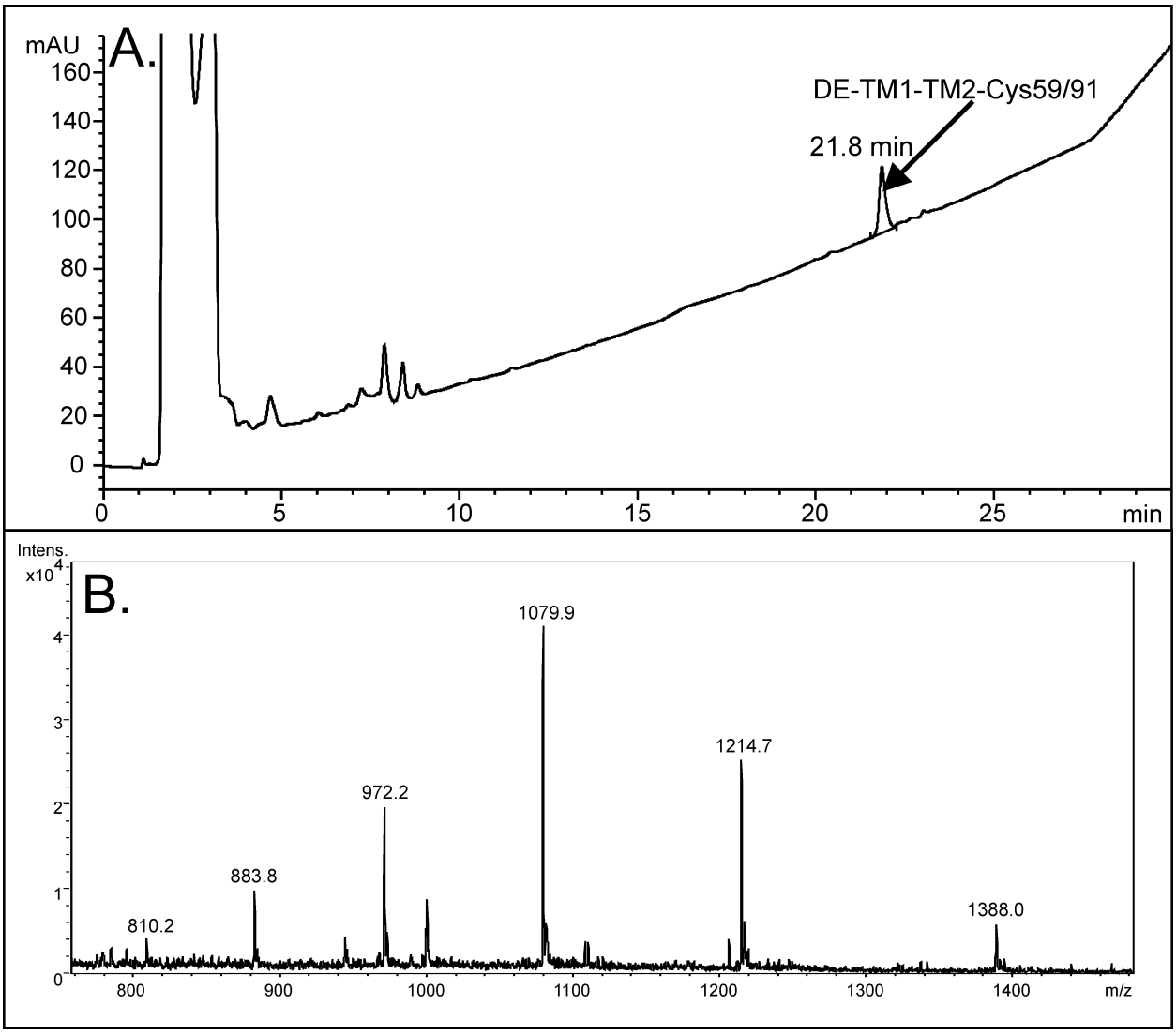


Figure 5-7. Analytical purification of DE-TM1-TM2-C59/91. A) The IBs from a small-scale induction were analyzed on an analytical Zorbax 300SB-C3 column with a 36-72% acetonitrile:water gradient with 10% isopropanol and 0.1% TFA at 60°C. B) The peak at 21.8 minutes was collected and analyzed by ESI-MS. The calculated MW was 9711.39 Da and the observed MW was 9710.1 Da.

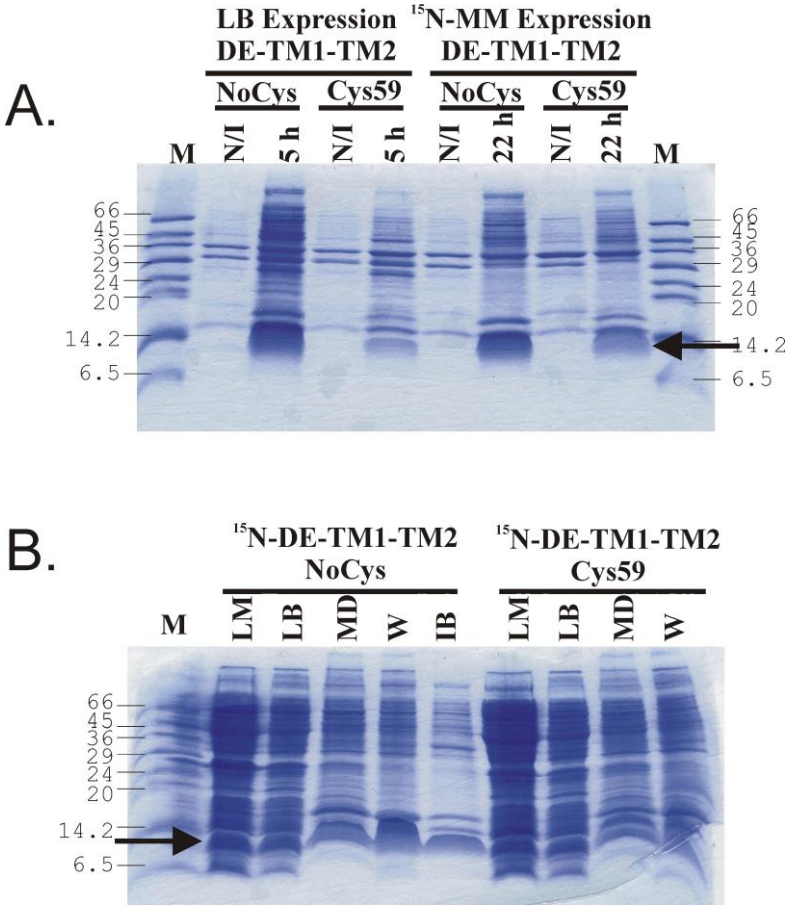


Figure 5-8. Induction of DE-TM1-TM2 and DE-TM1-TM2-C59. A) Expression of the directly expressed TM1-TM2 peptide without a Cys (NoCys) and with a single Cys residue at 59 (Cys59) was performed in a small scale growth in LB and a large scale growth in ^{15}N-labeling media. Expression was induced for 5 hours at 37°C in LB and 22 hours at 22°C in minimal media (MM). The arrow indicates the protein at the expected MW. B) Analysis of the large-scale IB preparation for the ^{15}N-labeling growths to determine which fraction the peptide was in. LM Lysis Mix; LB Lysis Buffer; MD Mild Detergent; W Water; IB Inclusion Bodies.

To determine whether this protein band was the target product, the fractions that were collected during the IB isolation of the DE-TM1-TM2 peptide (see Methods in Chapter 2 for details) were analyzed by analytical HPLC (Figure 5-9). As judged by the absence of a peak at about 20 minutes the lysis mix did not contain any DE-TM1-TM2 in either the TFA supernatant or pellet. A small amount of DE-TM1-TM2 was present in both the supernatant and pellet from the TFA precipitation of the lysis buffer. The mild detergent wash did not have any DE-TM1-TM2 peptide in it, but surprisingly, the water fraction contained a considerable amount of

peptide mostly in the TFA pellet. The large majority of the desired product was in the IB pellet. ESI-MS verified the correct MW for a peptide that has 95% incorporation of the ^{15}N atom ($MW_{expected\ for\ 100\%\ incorporation}$=9825.34, $MW_{observed}$=9818.85). Similar results were obtained for DE-TM1-TM2-Cys59 ($MW_{expected\ for\ 100\%\ incorporation}$=9841.41, $MW_{observed}$=9835.14). The water samples can be lyophilized and then the detergent removed by a modified acetone precipitation (*259*) prior to HPLC purification.

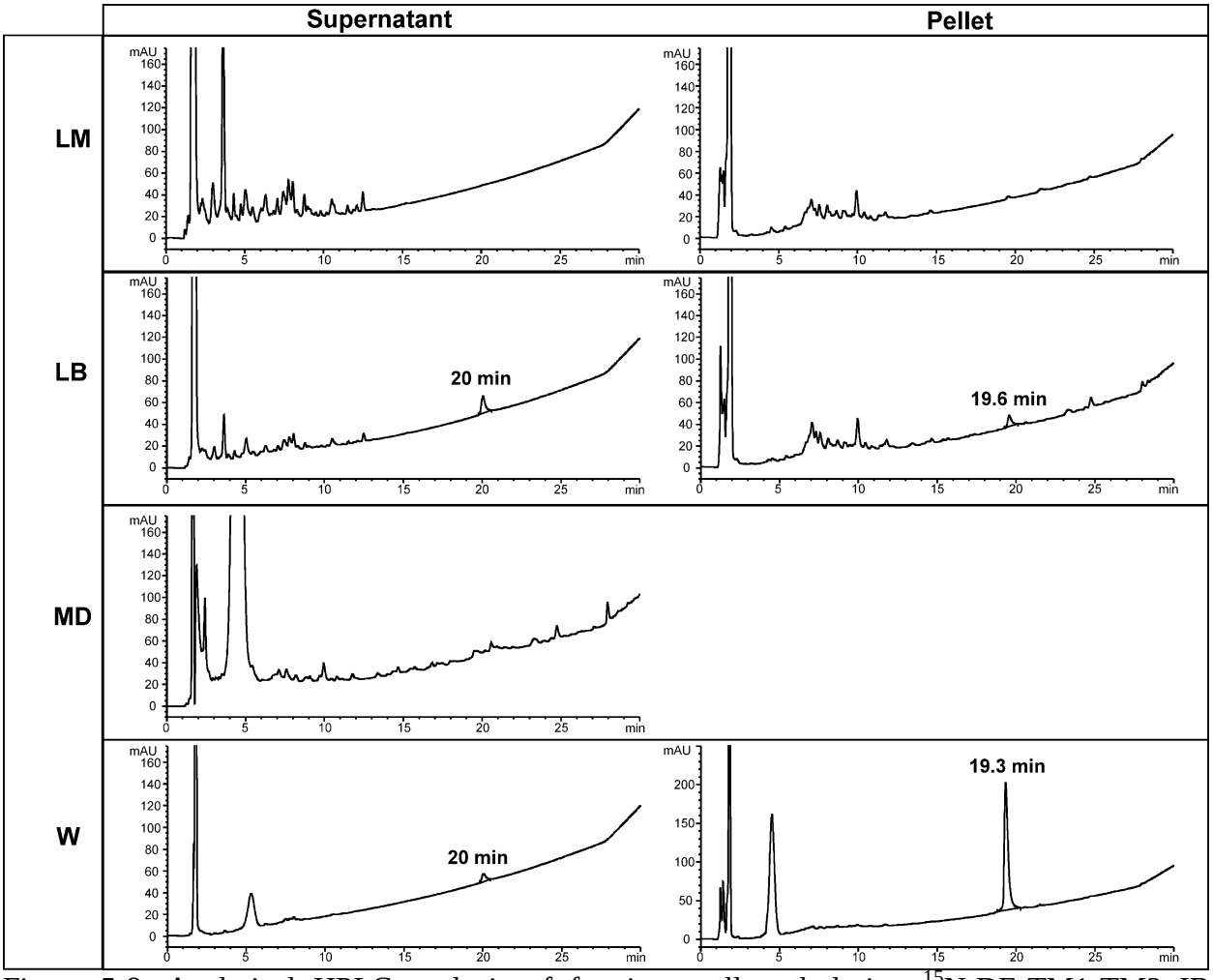


Figure 5-9. Analytical HPLC analysis of fractions collected during ^{15}N-DE-TM1-TM2 IB preparation. The IB preparation fractions were treated with TFA and the precipitate was pelleted. The supernatant and the pellet which was redissolved in TFE;water (1:1, v:v) were analyzed by analytical HPLC on a Zorbax 300SB-C3 column with a 36-72% acetonitirile water gradient with 10% isopropanol and 0.1% TFA. The ^{15}N-DE-TM1-TM2 peak eluted between 19 and 20 minutes as indicated by the retention time on each chromatogram.

Optimization of expression experiments for DE-TM1-TM2-Cys59 and DE-TM1-TM2-Cys59/Cys91 were performed in BL21-AI cells. These were induced in minimal media at two different temperatures (22°C and 37°C) and for two different times (5 h and 22 h). The gels in Figure 5-10A show that at both temperatures there is low expression with a 5 h incubation after induction with 0.5% L-arabinose. The level of expression for both peptides is optimal at 37°C with overnight incubation. These conditions were used to express all of the Cys containing peptides (Figure 5-10B). The expression appeared lower than expected for most of the peptides in minimal media (Figure 5-10B, left gel) and something was wrong with the Cys59/Cys91 sample. The anomaly on the gel occurred on two separate gels run with the same samples. There may be a lot of detergent in this IB sample thus affecting the gel pattern. The samples from the 50 mL LB cultures were better, but the Cys59/Cys91 and Cys59/Cys93 peptides were poorly expressed. These samples were used to verify the MW of each peptide which agreed within two Daltons of the expected values for the unlabelled polypeptides (Table 5-1).

Purification of DE-TM1-TM2. The DE-TM1-TM2 peptide was purified by prep RP-HPLC and the MW was verified by ESI-MS (Figure 5-11A). The peptide eluted at about 31 minutes using a 45-80% acetonitrile:water gradient with 10% isopropanol and 0.1 % TFA at 60°C. The DE-TM1-TM2-Cys59 peptide eluted at 34 minutes under the same conditions (Figure 5-11B). The peak height was considerable less, as expected from the SDS-PAGE results. The IBs from the large scale growths in ^{15}N-labeling media were purified under similar conditions and the fractions were collected, ESI-MS analysis was performed (Table 5-1) and the samples were lyophilized. During purification of the ^{15}N-DE-TM1-TM2 peptide, I noticed that there was peptide remaining on the column so I ran the gradient 32-72% acetonitrile:water with 20%

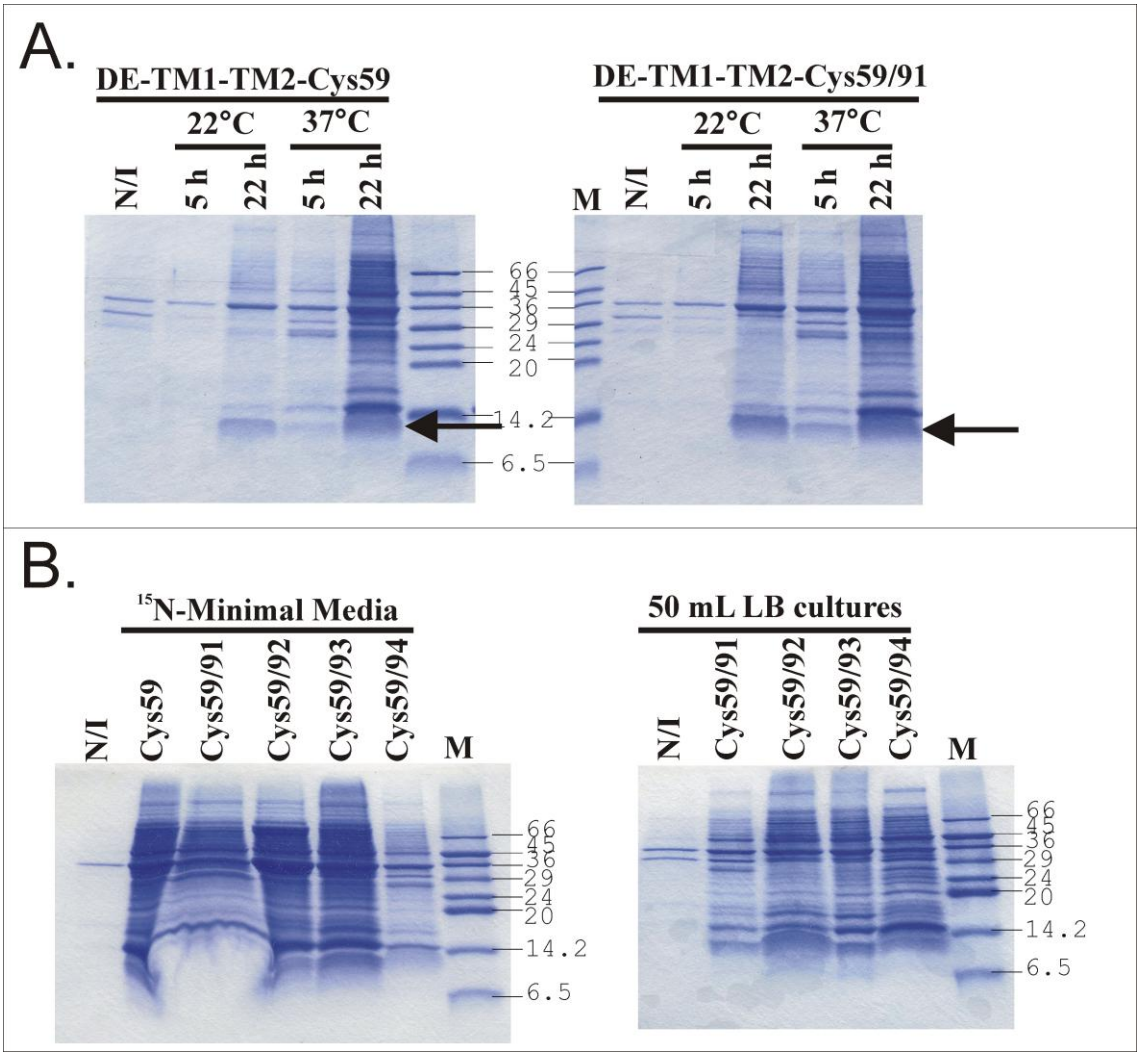


Figure 5-10. Expression optimization in BL21-AI of DE-TM1-TM2 Cys containing peptides. A) Expression of the directly expressed TM1-TM2 peptide with a single Cys residue at 59 (Cys59) and double Cys residues at 59 and 91 (Cys59/91) was performed in a small scale growth in minimal media at 22°C and 37°C. Expression was checked at 5 hours and 22 hours after induction. The arrow indicates the protein at the expected MW. B) Analysis of the large-scale ^{15}N-labeling inductions of all of the Cys containing peptides after 22 hours at 37°C (left gel) and small scale inductions in LB after 5 hours at 37°C (right gel).

isopropanol and 0.1% TFA at 60°C. The water wash fractions were also purified for each of the peptides after precipitation by either acetone or TFA. In the end, I purified ~4 mg per liter of culture of ^{15}N-DE-TM1-TM2 and ~1.4 mg per liter of culture of ^{15}N-DE-TM1-TM2-Cys59 with high levels (95%) of ^{15}N incorporation.

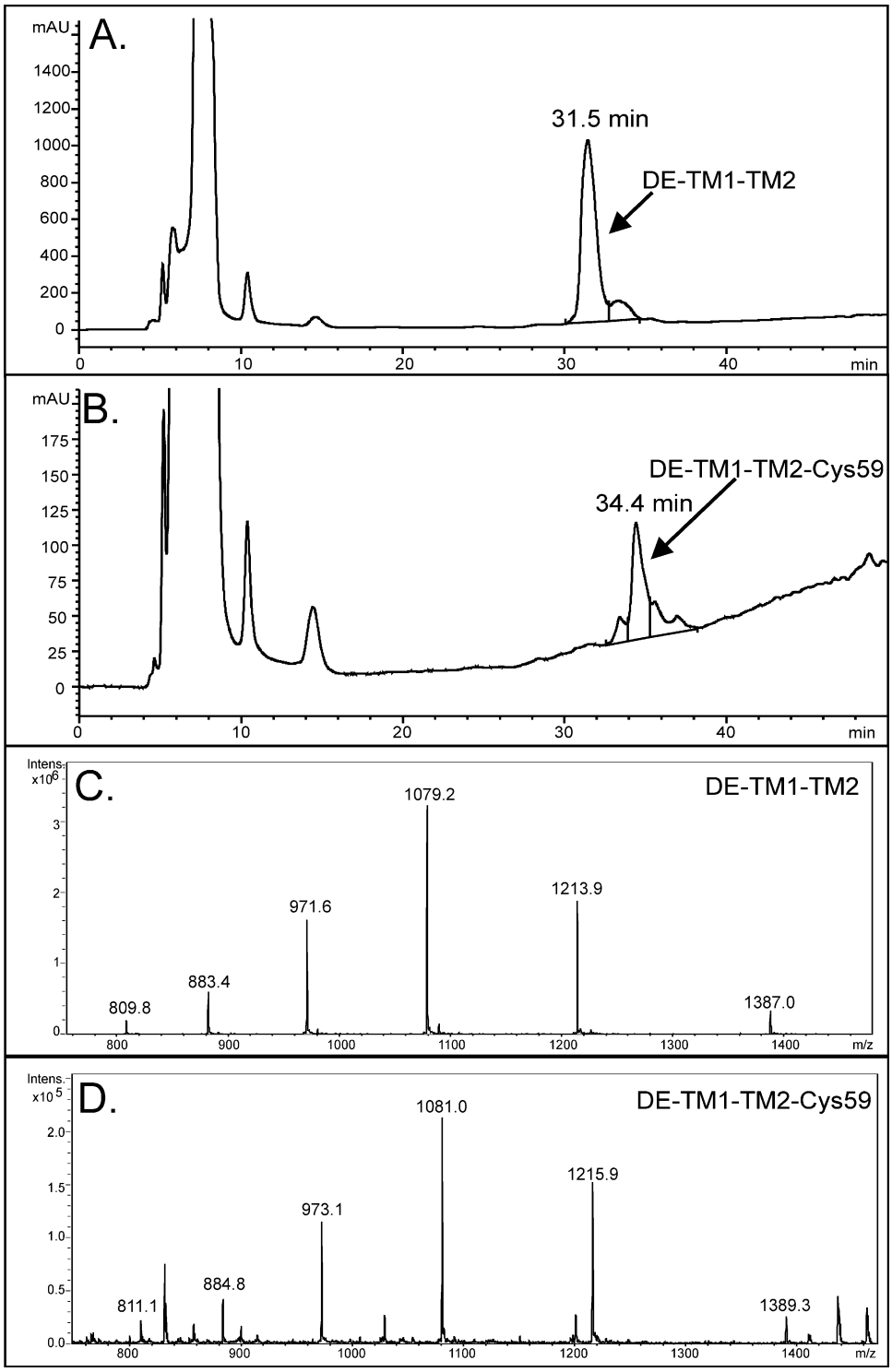


Figure 5-11. Preparative HPLC purification of DE-TM1-TM2 and DE-TM1-TM2-Cys59. Preparative HPLC was performed to purify DE-TM1-TM2 (A) and DE-TM1-TM2-Cys59 (B) from IBs generated from small-scale inductions in LB. The MW of the peptides was analyzed by RP-HPLC. C) The expected MW of DE-TM1-TM2 is 9705.34 Da and the observed MW is 9704.5 Da. D) The expected MW of DE-TM1-TM2-Cys59 is 9721.41 Da and the observed MW is 9719.5 Da.

Attachment of MTSL paramagnetic label to ^{15}N-DE-TM1-TM2-Cys59. The peptide was solubilized in 6M GnHCl buffer and TFE as described in the methods and the MTSL was added.

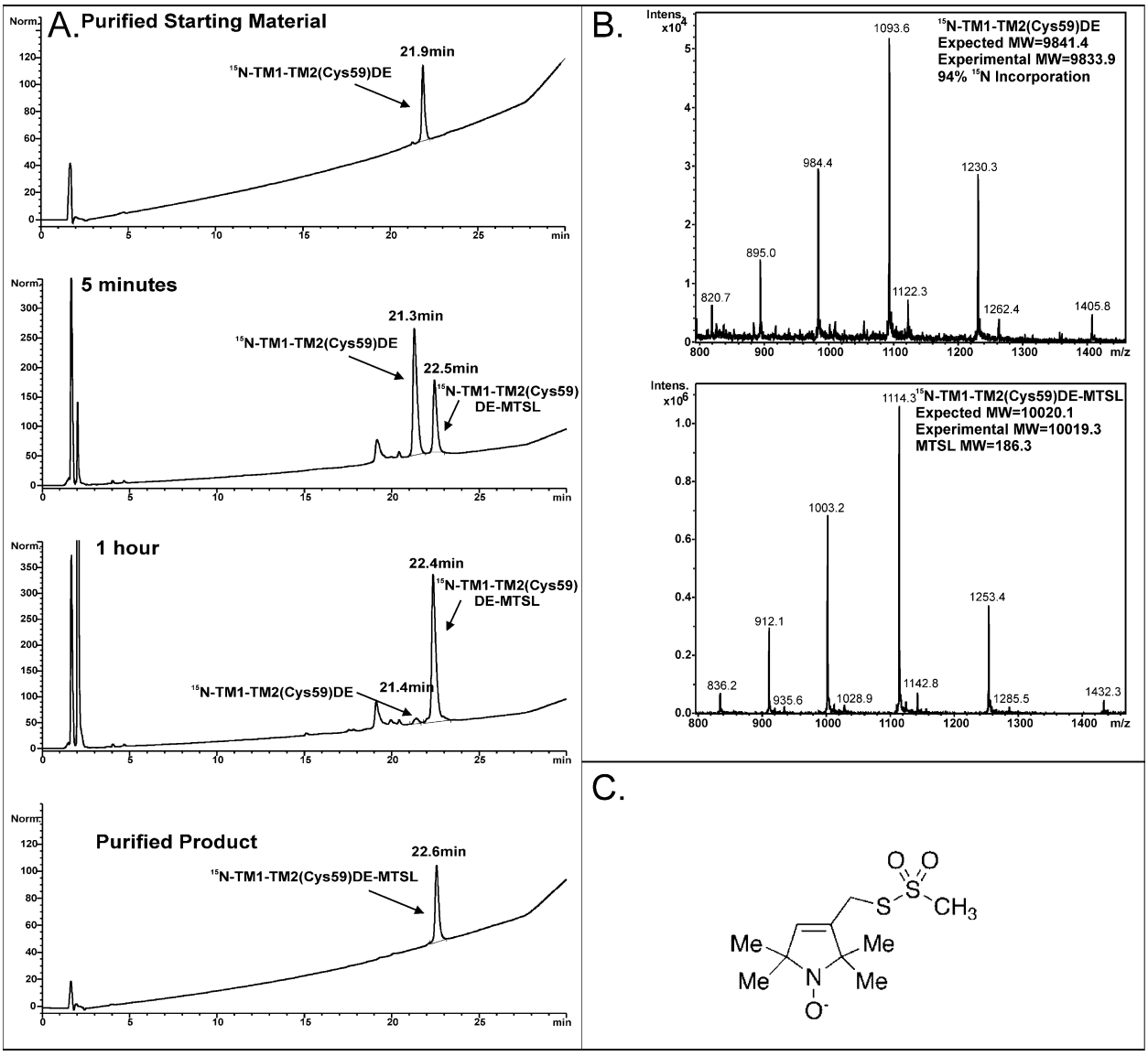


Figure 5-12. Addition of the S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothionate (MTSL) paramagentic label. A) The labeling reaction of ^{15}N-DE-TM1-TM2-C59 with the paramagnetic spin label was performed in 6M GnHCl:TFE (1:1) with 10 mM MTSL. The reaction was followed over time by analytical RP-HPLC as described and the peaks were collected and analyzed by ESI-MS. B) The ESI-MS analysis for the starting material (top panel) and product of the labeling reaction (bottom panel). The MWs are given in each panel. C) The MTSL paramagnetic label.

Attachment of the spin label began almost immediately as judged by the disappearance of the 21.3 min peak and the appearance of a peak at 22.6 min (Figure 5-12). The peaks were collected after 5 minutes and the MW was determined for each by ESI-MS. The MW of the 21.3 min peak was 9834.8 Da which is the ^{15}N-DE-TM1-TM2-Cys59 peptide and the MW of the 22.4 minute peak was 10019.32 Da. The difference in the MW of the two peptides was 185.9 Da and the expected MW of the spin label was 186.3 Da confirming the synthesis of the correct spin labeled product. After 1 h HPLC showed that the reaction was complete. The peptide was purified by multiple loadings on the analytical HPLC with a 36-80% acetonitrile:water gradient with 10% isopropanol and 0.1% TFA at 60°C. The final product (~0.24 mg) was pure as analyzed on RP-HPLC and had the expected molecular weight (10019.3 Da). This was used for NMR analysis to help increase the number of long-range constraints in the structure calculations (see Chapter 4).

Purification of Double Cys containing DE peptides. Inclusion bodies containing ^{15}N-DE-TM1-TM2-Cys59/93 were loaded onto a prepHT Zorbax 300SB-C3 column and the peptide was purified (Figure 5-13). The purified peptide was found by ESI-MS to have 96% incorporation of the ^{15}N isotope ($MW_{\text{expected for 100\% incorporation}}$=9831.39, MW_{observed}=9826.77) and was quantitated by UV to yield 2.6 mg per liter of culture. The column was cleaned and then inclusion bodies containing ^{15}N-DE-TM1-TM2-Cys59/Cys92 were loaded onto the column. No peptide peak corresponding to the correct retention time or MW appeared on the chromatogram. This purification was attempted with all of the double Cys containing peptides.

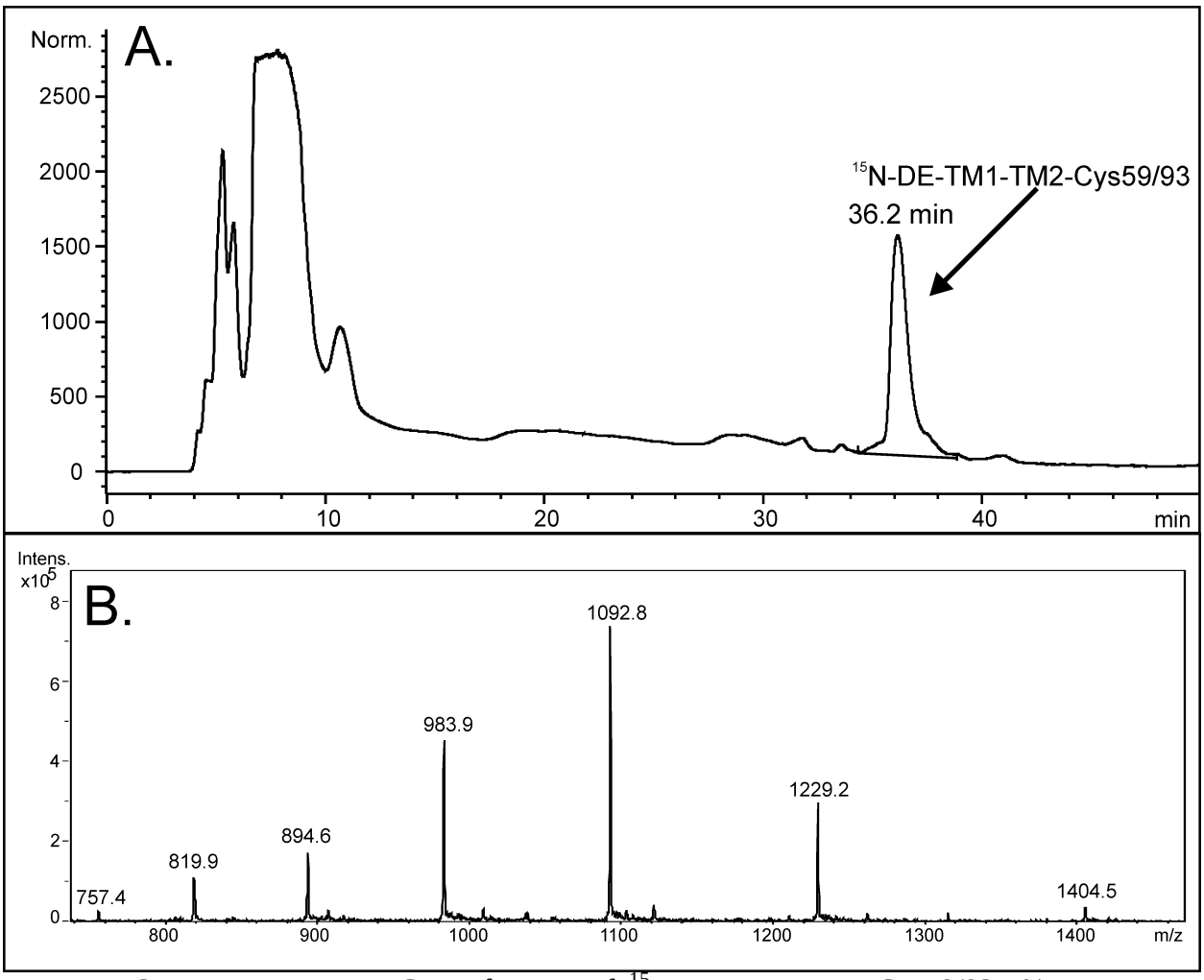


Figure 5-13. Preparative HPLC purification of ^{15}N-DE-TM1-TM2-Cys59/93. A) Preparative HPLC was performed to purify ^{15}N-DE-TM1-Cys59/93 from IBs generated from a large-scale induction in ^{15}N-labeling media. The peak at 36.2 minutes was collected and characterized by ESI-MS. The expected MW of 15N-DE-TM1-TM2-Cys59/93 if there is 100% incorporation is 9831.39 Da and the observed MW is 9826.77 Da which resulted in 96% incorporation of the ^{15}N moiety.

The unlabeled IBs from the 50 mL LB cultures were purified on the analytical column to verify the MW of each peptide (Table 5-1). When purifying on the analytical, I noticed that there was a lot of other protein coming off the column in the beginning. This mixture of proteins was very difficult to remove from the column even after washes with TFA, DMSO and high to low acetonitrile gradients. I hypothesized that I could use the His_6 tag to purify the peptide on a Ni^{2+}-NTA column with an imidazole gradient. The ^{15}N-TM1-TM2-Cys 59 IBs were solubilized

in TFE:GnHCl and incubated with the Ni^{2+}-NTA resin for 2 hours. The fractions were analyzed by SDS-PAGE and the peptide was separated from the other proteins in the IBs, but it was very dilute and spread throughout the elutions (Figure 5-14). This is not a good way to purify the Cys containing peptide but it may be useful as a first step in the purification process. The removal of a lot of the higher MW proteins in the flow through and washes indicates that the peptide will be enriched in the eluents when combined and concentrated. The buffer can be exchanged using a spin column and then the peptide can be injected onto the RP-HPLC column for final purification.

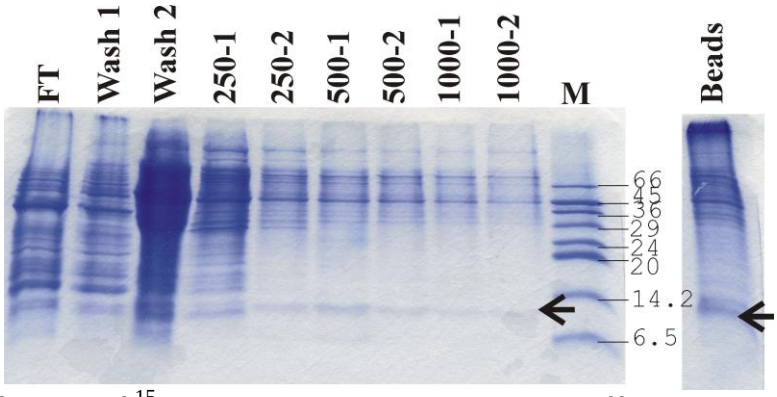


Figure 5-14. Purification of ^{15}N-DE-TM1-TM2-C59 by Nickel affinity column chromatography. Purification of ^{15}N-DE-TM1-TM2-C59 under denaturing conditions by an imidazole gradient. The arrow indicates where the protein of interest would run. FT=flow through, W= 8M urea buffer with 50 mM imidazole,, E250=8M urea buffer with 250 mM imidazole, E500=8M urea buffer with 500 mM imidazole, E1M=8M urea buffer with 1M imidazole. Beads= Ni^{2+}-NTA resin after the E1M elution were taken from the column and resuspended in SDS loading buffer.

5d. Summary

Generation of plasmids coding for Cys containing fusion proteins was done in collaboration with the University of Tennessee (Ms Hee Jung Kim). The expressed Ste2p mutant proteins were were all shown to respond to α-factor similarly to WT Ste2p. Although the cloned TrpΔLETM1-TM2 Cys containing fusion proteins expressed well, CNBr cleavage and purification of these proved difficult. Directly expressed TM1-TM2 peptides expressed at reasonable levels in BL21-AI and I was able to purify more than 4 mg per liter of culture of ^{15}N-DE-TM1-TM2 and 1.4 mg per liter of culture of ^{15}N-DE-TM1-TM2-C59. The ^{15}N-DE-TM1-TM2-C59 peptide was used to incorporate the MTSL paramagnetic spin label and this was useful in determining additional long range constraints for calculation of the TFE:water derived structure of TM1-TM2 (see Chapter 4). Even though the purification of the double Cys peptides has proven to be problematic, the direct expression of TM1-TM2 with both single and double cysteine residues was accomplished. Yields are modest but can be optimized.

Chapter 6

Conclusions and Future Directions

The study of the structure of protein fragments by solution NMR is a practical experimental paradigm. At the time I began this research project the unanswered questions relating to their use were: 1) Whether 2TM peptide fragments of a GPCR could be biosynthesized and purified in the mg quantities necessary for use in NMR analysis; 2) Whether small fragments of a GPCR will fold into a tertiary structure; 3) Whether organic:aqueous media can be used as surrogates for the membrane environment; and 4) Whether the tertiary structure of the GPCR fragment is the same as when this part of the protein in the context of the full receptor.

1) Biosynthesis of 2TM peptides for use in solution NMR analysis.

An NIH review panel was highly skeptical that fragments of Ste2p could be expressed and purified in quantities necessary for the high resolution NMR studies that we proposed. By optimizing expression conditions I was able to successfully meet this challenge. Best yield of TM1-TM2-FP were achieved in *E. coli* BL21-AI and high expression was maintained for this strain grown in minimal medium allowing efficient production of isotopically labeled peptide. Further increases in yield were achieved by the direct cleavage of the fusion protein in the inclusion bodies. Overall the yields of TM1-TM2 in multiple isotopically labeled backgrounds were from 6-20 mg of peptide per liter of culture. The yields of TM6-TM7-CT40 were reduced due to putative aggregation during the HPLC purification. The expression and purification of directly expressed TM1-TM2 resulted in decreased yields, but by eliminating the need for CNBr cleavage I was able to prepare Cys containing peptides for spin labeling studies. This greatly

benefited this project. Our lab has continued to generate multi-mg quantities of larger peptide fragments. Katrina Caroccia has been able to express, cleave and purify a 3TM construct in a uniformly ^{15}N-labeled background to 15-20 mg per liter of culture. Dr. Zhanna Potetinova has been able to directly express and purify a 5TM peptide to similar yields.

Multiple labeling schemes were needed for the NMR evaluation of TM1-TM2 in both TFE:water and LPPG micelles. I was able to label uniformly with ^{15}N and ^{13}C with no reduction in yields. Furthermore, by extensive optimization, I was able to partially and fully deuterate the TM1-TM2 fusion protein with only a small decrease in yields. The perdeuterated protein was selectively labeled at the methyl groups of Leu, Ile and Val. Due to the high cost of the deuterated components of the media, optimization is very important to assure good protein yields. The work that I performed by optimizing the expression, cleavage and purification of these peptides is a good basis for the structural analysis of this peptide and allowed us to generate structures in two membrane mimetic environments.

2) Tertiary fold of a GPCR peptide fragment.

Peptide fragments of membrane proteins have been shown to fold in organic:aqueous membrane mimetics (*89, 212*). Prior to commencing this study, however, no information concerning tertiary structures of a multi-transmembrane domain fragment of a heptahelical receptor was available. Recently, working with our collaborator in Zurich, the structure of TM1-TM2 in LPPG micelles was solved to 2.36±0.97 Å (*213*). Using heteronuclear multidimensional NMR and selective methyl group labeling we have shown that this GPCR peptide fragment folds into a helical hairpin in detergent. Furthermore, this hairpin structure can be superposed on the structure of this region of Ste2p determined using the rhodopsin-templated model of this GPCR

with an RMSD of about 1.8 Å (*8*, *213*). In this work, I was able, using the selective methyl group labeled peptide and a paramagnetic spin labeled peptide, to show that TM1-TM2 also folds into a hairpin structure in TFE:water. Though the tertiary structure in TFE:water does not converge as well as that of this peptide in LPPG micelles, both membrane mimetic environments support the formation of a tertiary fold. In both membrane mimetics there appears to be a flexible region from G56 to G60 which is biologically relevant because the α-factor peptide ligand has been shown to crosslink to residues in this region of Ste2p(*19-21*). The Naider and Becker labs have hypothesized that the flexibility of the GxxxG domain, which is towards the middle of TM1, is important to provide access for the α-factor to its binding partners in this region of the receptor. Thus I conclude that peptide fragments representing a sufficiently large domain of a GPCR can assume a structure that resembles that region in the intact receptor.

3) Trifluoroethanol:water as a surrogate for the membrane environment.

Membrane proteins require a hydrophobic environment. Detergent micelles and lipid bicelles and/or bilayers have been used to mimic the membrane for biophysical investigations of IMPs. However, the latter two environments are slightly prohibitive, due to an increase in the size of the macromolecular complex, for the analysis of IMP structures by solution NMR. Biophysical studies of GPCR fragments have been performed in both detergent/lipid and organic:aqueous environments. In this thesis, I have presented a comprehensive comparison of the effects of micelles and TFE:water on the secondary structure of TM1-TM2 and TM6-TM7-CT40. Both of the environments supported highly helical conformations. The TM1-TM2 peptides were helical in all of the TFE conditions tested and the helicity in micelles was dependent on acyl chain length and head group size. Similar results were observed with the

TM6-TM7-CT40 peptide, although the helicity was much lower due to the presence of the disordered carboxy-tail. Based on my analysis I conclude that organic:aqueous media, specifically TFE:water, is as good a surrogate as detergent micelles for determining whether the TM regions of a GPCR assume helical structures. As expected, the percent helicity in TFE:water was almost always higher than that in detergent. Due to the helix forming nature of TFE it was assumed that the organic:aqueous medium will likely cause some nontransmembrane regions of the polypeptide chain to assume non-native helical structures.

At the tertiary structure level important differences between LPPG and TFE:water were observed. As described above, the structure determined for TM1-TM2 in TFE:water was not very well defined. Many violations of NOE constraints remain in my final structure, only 88% of the residues in the helical regions are in favored or additionally favored regions of the Ramachandran map, and the RMSD of the overall helix hairpin was more than 6.5 Å. Conversely, in the structure determined in LPPG micelles, the residues in the TM regions assumed ideal α-helical angles and the RMSD was 2.36±0.97 Å. Thus, using solution NMR structure determination of a 2TM GPCR fragment as the criterion, I conclude that the use of organic:aqueous media at a ratio of 1:1 (v:v) would not result in a useful model of the native structure of the TM1-TM2 region of Ste2p. It is possible that more relevant structures will occur using higher percentages of TFE (i.e. 2:1 ratio) or from the investigation of larger GPCR fragments. The increase in TFE may result in more stable helices by excluding the water from the TM domains, thus making the medium more similar to the acyl chain environment of the membrane. As more helices are added to the GPCR fragment, more interhelical contacts would help to stabilize the tertiary fold of the peptide. In either case, structures that are more comparable to those in detergent and/or lipid environments may result. If this occurs, the

chemical shift assignment in TFE:water should correlate with the assignments in the larger detergent/peptide or lipid/peptide macromolecular complexes and can be used together to generate higher resolution structures. In our lab, an NMR analysis of a 3TM peptide is being done by Katrina Caroccia on the TM1-TM3 peptide in TFE:water, while our Zurich collaborators are determining the structure of the same peptide in a micelle. Comparison of these structures will test the above hypothesis.

4) Biological relevance of the 3D structure of TM1-TM2 peptide fragment and investigations in the context of a full receptor.

Due to the vast amount of biological and biochemical data for Ste2p in *Sacchromyces cerevisiae*, we are able to consider the biological relevance of the calculated structure. For example, it has been shown that the α-factor interacts with residues R58 and C59 which are located towards the middle of TM1. In both TM1-TM2 structures, we have observed an increased region of flexibility surrounding these residues which would indicate an opening for the peptide to interact with the protein. We can also compare the NMR structures of TM1-TM2 to the rhodopsin-templated model (*8*). Analysis of the TM1 and TM2 helices indicates that the residues on TM2 that are closest in space to the C59 in TM1 would be Ile91 and Ile92. Determining the distance between residues in the structure bundle (PDB 2K9P) of TM1-TM2 in LPPG micelles, the S59 is separated from Ile93 by about 2 Å and Ile 92 is twisted away from the Ser. Further analysis of the LPPG structure showed that Ser59 and Phe89 are actually the closest in space, approximately 1 Å. It is difficult to determine the exact distance because the pdb file is a bundle of conformers and therefore there are 20 different Phe residues and 20 different Ser residues. However, my comparison indicates that although the structure is well formed in LPPG

micelles, it varies from the rhodopsin-templated model. The structure in TFE:water was not well-defined and so it was difficult to ascertain which residues were closest in space to Ser59. However, based on the helical wheel (Figure 4-26), Ile92 is on the side of TM2 that faces TM1 and might be in position to interact with Ser59 which is in better agreement of the predicted TM positions of TM1-TM2 in the templated receptor model.

Future Goals

Refine structure in TFE:water. The structures in the bundle of TM1-TM2 in TFE:water(0.1% TFA) (1:1,v:v) had a high RMSD and there were many NOE and angle violations. This may be due to the presence of a distribution of conformations of this peptide in this medium. The structure determined in LPPG micelles is more constrained. It is possible that this reflects the interactions with the hydrophobic acyl chains of the lipid whose positions are, in general, more highly defined than those of the solvent TFE:water molecules as they interact with other acyl chains in the micelle. The result is that flexible regions in the TM domains are not as constrained. Therefore, in order to determine a high-resolution structure in TFE:water it will be necessary to add more constraints. One way to constrain TM1-TM2 is the addition of a non-native disulfide bond as described above. In terms of teasing out the relationship of TM1-TM2 in TFE:water I also propose the use of residual dipolar coupling (RDC) analysis.

As described in Chapter 5, difficulties were encountered during the purification of the double Cys mutants of TM1-TM2. I believe that I can overcome this problem by using a two-step purification method, Ni^{2+}-NTA followed by RP-HPLC. This strategy should result in the optimization of the purification and increase peptide yields. It remains to be seen whether these disulfide bonds will form and whether a unique structure will result. If numerous conformations

can form then several different Cys pairs may result in a disulfide constrained helical hairpin and it will not be possible to choose which one is the "native" structure. However, if only one of the combinations results in disulfide formation this may suggest that a relevant structure has been found. Of course this can be tested by comparison with the rhodopsin-templated model. Furthermore, disulfide constraints may be useful if added towards the top of the TM domains where the helices splay apart.

The use of RDC analysis will help to determine the position of the two TM domains relative to each other. I would like to generate compressed gels containing TM1-TM2 in TFE:water and determine the RDC values when compared to the non-compressed gel. Work has been done by Horst Kessler with RDCs in organic environments (*260-264*) and I would like to apply this to the TM1-TM2 system. Integration of this methodology into the Naider laboratory will also be useful as our lab moves to larger fragments and ultimately intact Ste2p.

Biophysical analysis of the disulfide TM1-TM2 constructs. The analysis of the double Cys mutants of TM1-TM2 is not limited to NMR. These peptides can be analyzed by Far UV CD to determine effects, if any, on secondary structure and Near UV CD to determine changes in the local environment of the Trp residue. The use of temperature to affect the secondary structure and/or folding of these peptides can be applied to both regions of the CD spectra, 222 and 280 nm. If differences between the breakdown of the tertiary folds and the secondary structure can be determined, evidence will have been found for the presence of a folded conformation, even at a lower resolution.

The folding of these peptides can also be followed by SDS-PAGE analysis as described by the Deber group (*173-176*). The interactions between the two TM domains can be analyzed in

the presence or absence of β-mercaptoethanol which would alter the oxidation state of the disulfide bond if present. The peptide would run at a faster rate if it were in a constrained conformation as opposed to a more open non-disulfide linked state.

Interactions of 2+5 constructs. One of the main arguments against the use of fragments is that the peptide region is not in the context of the full receptor. As discussed above, it has been shown with many GPCRs that co-expressed fragments of these proteins can find each other in the plasma membrane and reform functional units. To that end, work is being done by Dr. Subramayan Tantry and Dr. Zhanna Potetinova to express the remaining 5TMs of the peptide (TM3-TM7) in the hope that the two fragments will form a complex in a micellar environment. The 5TM peptide has been expressed and purified and analysis by SDS-PAGE is underway to determine if these peptides can interact and therefore run as heterodimers. Further analysis by both far and near UV CD can be useful in determining changes in secondary structure or the Trp residue environments. To determine if this complex is biologically active in a micellar environment, α-factor binding assays can be performed. Our collaborator at the University of Tennessee, Dr. B.K. Lee, has shown that full-length Ste2p can bind α-factor when solubilized in DDM micelles (*265*). After this complex has been characterized, isotopically labeled TM1-TM2 will be mixed with unlabeled 5TM and NMR characterization will be performed. Initial experiments will include [^{15}N,^{1}H]-HSQC to determine chemical shift perturbations due to the presence of the remaining TM domains and spin-label characterization to determine the relative positions of the helices. Future experiments may include the use of the [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-Ile, Leu, Val)]-labeled TM1-TM2 peptide to determine the assignments of the methyl groups and any long-range connectivities between the helices. There is also the potential to perform the

above experiments in the presence of α-factor which would be useful in determining the difference in conformation when in the active state.

List of publications as a result of this thesis work

Peer-Reviewed Journal Articles:

Cohen, LS, Neumoin A, Arshava B, Becker, JM, Zerbe O and F Naider. Structure Determination of Ste2p(G31-T110) in Trifluoroethanol:Water an Organic:Aqueous Membrane Mimetic. *Manuscript in preparation*.

Cohen, LS, Becker, JM and F Naider. Biosynthesis of peptide fragments of eukaryotic GPCRs in Escherichia coli by directing expression into inclusion bodies. *J Pept Sci*. 2010. **16(5)**:213-218.

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

Cohen LS, Arshava B, Estephan R, Englander J, Kim H, Hauser M, Zerbe O, Ceruso M, Becker JM, and F Naider. Expression and biophysical analysis of two double-tranmembrane domain-containing fragments from a yeast G protein-coupled receptor. *Biopolymers*. 2008; **90**:117-30.

Proceedings and Presentations:

Cohen LS, Neumoin A, Arshava B, Hauser M, Becker JM, Zerbe O and F Naider. NMR studies on a double transmembrane-containing fragment of a G protein-coupled receptor. Proceedings of the 21st American Peptide Society Symposium. 2009. Bloomington, Indiana. *Winner Young Investigator Poster Competition.*

Cohen LS, Arshava B, Estephan R, Englander J, Hauser M, Becker JM and F Naider. Expression of double transmembrane domain GPCR fragments for biophysical analysis. Proceedings of the 20th Annual American Peptide Society Symposium. 2007. Montreal, Canada. *Peptide Idol Poster Competition Participant.*

Presentations:

Cohen LS, Arshava B, Neumoin A, Hauser M, Becker JM, Zerbe O, and F Naider. Structural studies of Ste2p(G31-T110), a double transmembrane G protein-coupled receptor, in trifluoroethanol:water. Gordon Research Conference 2010, Peptides, Biology and Chemistry of. Ventura Beach, CA.

Cohen LS, Arshava B, Neumoin A, Hauser M, Zerbe O, Becker JM and F Naider. Conformational comparisons of a double transmembrane domain GPCR fragment in membrane mimetic media. 2008 Macromolecular Assemblies Peptide Engineering Conference at City College, CUNY, New York, NY.

Appendix

Supplementary Table S1. ^{15}N and ^{1}H chemical shifts assigned at 45°C

amino acid in Ste2p	N	HN	HA	HB	Others
G31	-	-	3.865	n/a	n/a
N32	118.77	8.548	4.811	2.85	Nδ2-; Hδ21-; Hδ22-
G33	109.21	8.417	3.982	n/a	n/a
S34	115.64	8.059	4.49	3.897/3.961	Hγ-
T35	116.03	7.947	4.283	4.283	Hγ1 1.254; Hγ21-
I36	121.84	7.759	4.151	1.88	Hγ12 0.912; Hγ11-; Hδ11-; Hγ21-
T37	115.64	7.612	4.206	4.526	Hγ1 1.219; Hγ21-
F38	121.10	8.065	4.303	3.162	Hδ1-; Hε1-; Hζ-; Hε2-; Hδ2-
D39	117.41	8.253	4.401	2.86/3.01	n/a
E40	120.21	8.076	4.083	2.092/2.311	Hγ2 2.42; Hγ1 2.556
L41	121.20	8.1	4.096	1.688	Hγ-; Hδ11-; Hδ21-
Q42	116.97	8.125	3.836	1.959/1.999	Hγ2 2.23; Hγ1 2.125; Nε2-; Hε21-; Hε22-
G43	106.52	7.886	3.87	n/a	n/a
L44	125.32	8.061	4.172	1.8/1.86	Hγ 1.69; Hδ11-; Hδ21-
V45	122.46	8.531	3.586	2.105	Hγ11 0.918; Hγ21-
N46	117.56	8.436	4.439	2.802/2.916	Nδ2-; Hδ21-; Hδ22-
S47	116.43	8.36	4.235	4.014/4.092	Hγ-
T48	119.32	8	3.983	4.364	Hγ1 1.24; Hγ21-
V49	122.22	8.384	3.691	2.128	Hγ11 0.962; Hγ21 1.046
T50	114.11	7.904	3.847	4.269	Hγ1 1.238; Hγ21-
Q51	118.66	7.797	3.961	2.21/2.321	Hγ2-; Hγ1-; Nε2-; Hε21-; Hε22-
A52	122.55	7.948	4.159	1.563	n/a
I53	120.48	8.322	3.784	2.025	Hγ12 1.144; Hγ11 1.569; Hδ11-; Hγ21-
L54	119.77	8.283	4.062	1.85/1.444	Hγ 0.868; Hδ11-; Hδ21-
F55	118.96	8.461	4.304	3.245	Hδ1-; Hε1-; Hζ-; Hε2-; Hδ2-
G56	109.98	8.343	3.755	n/a	n/a
V57	122.51	8.424	3.812	2.201	Hγ11 0.949; Hγ21 1.048
R58	119.42	8.293	4.062	1.875	Hγ2-; Hγ1 1.798; Hδ2-; Hδ1-; Nε-; Hε-; Nη2-; Hη21-; Hη22-; Nη1-; Hη11-; Hη12-
S59	113.53	8.071	4.178	3.609/3.824	Hγ-
G60	110.23	7.805	3.901	n/a	n/a
A61	123.48	8.07	4.139	1.43	n/a
A62	121.43	7.975	4.08	1.424	n/a

A63	120.04	7.685	4.059	1.447	n/a
L64	116.85	7.803	4.109	1.699/1.449	Hγ 0.87; Hδ11-; Hδ21-
T65	112.97	7.732	3.8	4.3	Hγ1 1.206; Hγ21-
L66	120.60	7.595	3.981	1.805/1.54	Hγ 1.654; Hδ11-; Hδ21-
I67	119.45	7.822	3.836	2.097	Hγ12 1.005; Hγ11 1.87; Hδ11-; Hγ21-
V68	120.06	8.184	3.69	2.281	Hγ11 1.08; Hγ21 0.945
V69	123.88	8.561	3.55	2.307	Hγ11 1.055; Hγ21 0.945
W70	124.23	8.72	4.265	3.584	Hδ1-; Hζ2-; Hη2-; Hζ3-; Hε3-; Nε1-; Hε1-;
I71	119.01	9.126	3.57	2.01	Hγ12 1.445; Hγ11 0.937; Hδ11-; Hγ21-
T72	114.33	8.472	3.945	4.265	Hγ1 1.285; Hγ21-
S73	117.59	8.247	4.214	3.98	Hγ-
R74	120.81	7.672	4.07	1.72	Hγ2-; Hγ1 1.37; Hδ2-; Hδ1-; Nε-; Hε-; Nη2-; Hη21-; Hη22-; Nη1-; Hη11-; Hη12-
S75	114.26	7.8	4.298	3.945/3.886	Hγ-
R76	120.63	7.659	4.215	1.88	Hγ2 1.59; Hγ1 1.69; Hδ2-; Hδ1-; Nε-; Hε-; Nη2-; Hη21-; Hη22-; Nη1-; Hη11-; Hη12-
K77	119.53	7.793	4.357	1.878	Hγ2 1.435; Hγ1 1.357; Hδ2-; Hδ1-; Hε2-; Hε1-; Nζ-; Hζ1-
T78	114.66	7.613	4.57	4.225	Hγ1 1.241; Hγ21-
P79					Hγ2 1.229; Hγ1 0.846; Hδ2-; Hδ1-
I80	117.54	7.417	3.874	1.813	Hγ12-; Hγ11-; Hδ11-; Hγ21-
F81	119.72	7.486	4.382	3.186	Hδ1-; Hε1-; Hζ-; Hε2-; Hδ2-
I82	119.61	7.439	3.828	1.937	Hγ12 1.206; Hγ11 0.887; Hδ11-; Hγ21-
I83	119.84	7.807	3.746	1.877	Hγ12 1.179; Hγ11 0.888; Hδ11-; Hγ21-
N84	119.65	8.125	4.448	2.868/2.763	Nδ2-; Hδ21-; Hδ22-
Q85	119.07	7.89	4.025	2.122/2.214	Hγ2-; Hγ1-; Nε2-; Hε21-; Hε22-
V86	119.77	8.349	3.746	2.148	Hγ11 1.007; Hγ21-
S87	114.45	8.042	4.147	4.073	Hγ-
L88	121.53	7.673	4.155	1.735	Hγ-; Hδ11-; Hδ21-
F89	119.01	7.98	4.256	3.307	Hδ1-; Hε1-; Hζ-; Hε2-; Hδ2-
L90	117.60	8.316	3.975	2.109	Hγ 1.387; Hδ11-; Hδ21-
I91	122.78	8.157	3.75	2.15	Hγ12 1.384; Hγ11 1.092; Hδ11-; Hγ21-
I92	124.76	8.46	3.638	2	Hγ12 1.368; Hγ11 0.871; Hδ11-; Hγ21-
L93	121.05	8.64	3.956	1.47/1.388	Hγ-; Hδ11-; Hδ21-
H94	114.14	8.225	4.17	3.328	Hε1-; Hδ2-; Nδ1-; Hδ1-; Nε2-; Hε2-
S95	116.06	8.218	4.279	4.167/3.974	Hγ-
A96	125.37	8.386	4.208	1.524	
L97	117.89	8.07	4.088	1.811	Hγ 1.514; Hδ11-; Hδ21-

Y98	119.91	7.97	4.28	3.023/3.087	Hδ1-; Hε1-; Hη-; Hε2-; Hδ2-
F99	118.63	8.217	4.238	3.202	Hδ1-; Hε1-; Hζ-; Hε2-; Hδ2-
K100	118.35	8.089	3.901	1.969	Hγ2-; Hγ1-; Hδ2-; Hδ1-; Hε2-; Hε1-, Nζ-; Hζ1-
Y101	121.61	8.167	4.137	3.231/3.094	Hδ1-; Hε1-; Hη-; Hε2-; Hδ2-
L102	122.26	8.296	3.814	1.729/1.456	Hγ-; Hδ11-; Hδ21-
L103	118.46	8.227	4.037	1.61	Hγ-; Hδ11-; Hδ21-
S104	114.05	7.883	4.223	3.896	Hγ-
N105	119.38	7.659	4.54	2.437	Nδ2-; Hδ21-; Hδ22-
Y106	119.88	8.015	4.433	2.995/3.097	Hδ1-; Hε1-; Hη-; Hε2-; Hδ2-
S107	114.61	7.833	4.328	3.899	Hγ-
S108	116.30	7.78	4.47	3.907	Hγ-
V109	119.20	7.621	4.235	2.176	Hγ11 0.941; Hγ21-
T110	114.64	7.634	4.454	4.319	Hγ1 1.311; Hγ21-

Supplementary Table S2. ^{13}C chemical shifts assigned at 45°C.

amino acid in Ste2p	N	CO	Cα	Cβ	Others
G31	-	169.396	42.67	-	
N32	118.77	175.109	52.73	37.98	Cγ-
G33	109.21	174.064	44.98	-	
S34	115.64	174.659	58.29	62.68	
T35	116.03	174.671	62.52	68.65	Cγ2-
I36	121.84	176.362	62.01	37.25	Cγ1-; Cδ1-; Cδ2-
T37	115.64	175.258	63.17	68.75	Cγ2
F38	121.10	176.795	60.27	37.73	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
D39	117.41	177.399	55.16	36.17	Cγ-
E40	120.21	177.776	57.86	26.91	Cδ-
L41	121.20	178.102	57.31	40.4	Cγ-; Cδ1-; Cδ2-
Q42	116.97	178.32	58.24	27.21	Cγ-; Cδ-
G43	106.52	175.993	46.14	-	
L44	125.32	179.905	57.48	40.62	Cγ-; Cδ1-; Cδ2-
V45	122.46	177.589	66.09	30.72	Cγ1-; Cγ2-
N46	117.56	178.33	56	37.6	Cγ-
S47	116.43	175.407	61.27	62.08	
T48	119.32	175.892	65.98	68.03	Cγ2-
V49	122.22	177.252	65.82	30.77	Cγ1-; Cγ2-
T50	114.11	175.686	66.52	67.85	Cγ2-
Q51	118.66	177.386	58.7	27.43	Cγ-; Cδ-
A52	122.55	180.475	54.75	16.35	
I53	120.48	178.198	64.15	36.67	Cγ1-; Cδ1-; Cγ2-
L54	119.77	179.156	57.35	40.33	Cγ-; Cδ1-; CD2-
F55	118.96	178.657	60.23	37.71	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
G56	109.98	175.801	46.55	-	
V57	122.51	178.141	65.11	30.82	Cγ1-; Cγ2-
R58	119.42	177.871	58.04	28.7	Cγ-; Cδ-; Cζ-
S59	113.53	175.537	59.77	62.47	
G60	110.23	174.509	45.42	-	
A61	123.48	179.532	53.91	16.61	
A62	121.43	178.53	54.21	16.55	
A63	120.04	178.648	54.17	16.53	
L64	116.85	178.066	57.06	40.51	Cγ-; Cδ1-; Cδ2-
T65	112.97	175.002	66.01	68.12	Cγ2-

L66	120.60	178.501	57.53	40.13	Cγ-; Cδ1-; Cδ2-
I67	119.45	177.565	64.34	36.65	Cγ1-; Cδ1-; Cγ2-
V68	120.06	180.043	66.31	30.53	Cγ1-; Cγ2-
V69	123.88	178.281	66.38	30.35	Cγ1-; Cγ2-
W70	124.23	178.656	60.8	27.79	Cγ-; Cδ1-; Cζ2-; Cη2-; Cζ3-; Cε3-; Cδ2-
I71	119.01	179.313	64.57	37.11	Cγ1-; Cδ1-; Cγ2-
T72	114.33	176.75	65.29	68.32	Cγ2-
S73	117.59	175.591	60.71	62.2	
R74	120.81	177.257	56.74	28.62	Cγ-; Cδ-; Cζ-
S75	114.26	174.441	59.03	62.66	
R76	120.63	175.646	55.95	28.86	Cγ-; Cδ-; Cζ-
K77	119.53	175.505	55.32	31.4	Cγ-; Cδ-; Cε-
T78	114.66	172.898	60.29	68.65	Cγ2-
P79		176.775	63.48	30.07	Cγ-; Cδ-
I80	117.54	176.223	62.36	36.9	Cγ1-; Cδ1-; Cγ2-
F81	119.72	176.875	59.26	37.86	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
I82	119.61	176.832	63	36.73	Cγ1-; Cδ1-; Cγ2-
I83	119.84	178.202	63.35	36.65	Cγ1-; Cδ1-; Cγ2-
N84	119.65	176.083	55.58	37.83	Cγ-
Q85	119.07	177.606	58.29	27.56	Cγ-; Cδ-
V86	119.77	177.267	65.52	30.71	Cγ1-; Cγ2-
S87	114.45	175.182	61.75	61.7	
L88	121.53	177.904	57.36	40.4	Cγ-; Cδ1-; Cδ2-
F89	119.01	176.824	60.37	37.48	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
L90	117.60	180.213	57.5	40.43	Cγ-; Cδ1-; Cδ2-
I91	122.78	179.138	64.76	36.77	Cγ1-; Cδ1-; Cγ2-
I92	124.76	178.274	64.78	36.28	Cγ1-; Cδ1-; Cγ2-
L93	121.05	178.594	57.39	40.55	Cγ-; Cδ1-; Cδ2-
H94	114.14	176.372	59.03	26.83	Cγ-; Cε1-; Cδ2-
S95	116.06	174.701	61.19	61.94	
A96	125.37	179.701	54.58	16.49	
L97	117.89	178.653	57.07	40.44	Cγ-; Cδ1-; Cδ2-
Y98	119.91	177.035	60.31	36.94	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
F99	118.63	177.036	60.62	37.41	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
K100	118.35	178.857	58.9	30.76	Cγ-; Cδ-; Cε-
Y101	121.61	177.784	60.45	36.59	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
L102	122.26	179.124	57.48	40.59	Cγ-; Cδ1-; Cδ2-
L103	118.46	179.29	56.55	40.43	Cγ-; Cδ1-; Cδ2-

S104	114.05	174.807	60.04	62.25	
N105	119.38	175.349	54.29	38.45	Cγ-
Y106	119.88	176.126	58.76	37.21	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
S107	114.61	174.133	58.54	62.74	
S108	116.30	173.891	58.19	62.85	
V109	119.20	175.546	61.74	31.48	Cγ1-; Cγ2-
T110	114.64	175.419	59.88	69.35	Cγ2-

Supplementary Table S3. ^{15}N and ^{1}H chemical shift assignments at 25°C.

amino acid in Ste2p	N	HN	Hα	Hβ	Others
G31	-	-	3.958/3.863	n/a	n/a
N32	118.86	8.674	4.799	2.85/2.85	Nδ2 126.9; Hδ21 6.793; Hδ22 7.492
G33	109.50	8.571	3.97/3.97	n/a	n/a
S34	116.18	8.169	4.45/4.454	3.887/3.946	Hγ-
T35	116.72	8.015	4.295	4.277	Hγ1-; Hγ21 1.259
I36	122.32	7.848	4.12	1.885	Hγ12 1.577; Hγ11 1.247; Hδ11 0.845; Hγ21 0.921
T37	116.08	7.685	4.29	4.29	Hγ1-; Hγ21 1.26
F38	121.53	8.227	4.262	3.149/3.21	Hδ1-; Hε1-; Hζ-; HE2-; Hδ2-
D39	117.84	8.409	4.369	3.06/2.856	n/a
E40	120.83	8.167	4.058	2.056/2.079	Hγ2 2.405; Hγ1 2.605
L41	121.94	8.255	4.073	1.817/1.873	Hγ 1.689; Hδ11 0.857; Hδ21-
Q42	117.27	8.267	3.807	1.91/1.939	Hγ2 2.085; Hγ1 2.192; Nε2 124.92; Hε21 6.422; Hε226.571
G43	106.82	7.915	3.861		
L44	126.12	8.168	4.166	1.805/1.871	Hγ 1.685; Hδ11 0.857; Hδ21-
V45	123.54	8.698	3.551	2.109	Hγ11 1.001; Hγ21 0.917
N46	117.76	8.551	4.405	2.766/2.924	Nδ2 125.19; Hδ21 6.617; Hδ22 7.388
S47	116.96	8.459	4.235	4.002/4.083	Hγ-
T48	120.14	8.066	3.941	4.394	Hγ1-; Hγ21 1.244
V49	122.83	8.529	3.668	2.127	Hγ11 1.049; Hγ21 0.974
T50	114.63	7.998	3.86	4.291	Hγ1-; Hγ21 1.257
Q51	118.84	7.867	3.961	2.21/2.21	Hγ2 2.33; Hγ1 2.54; Nε2 124.59; Hε21 6.43; Hε22 6.99
A52	123.20	8.011	4.179	1.581	
I53	121.42	8.465	-	-	Hγ12-; Hγ11-; Hδ11 0.796 ; Hγ21-
L54	120.07	8.41	-	-	Hγ-; Hδ11 0.836; Hδ21 0.846
F55	119.59	8.571	4.267	3.227/3.307	Hδ1-; Hε1-; Hζ-; Hε2-; Hδ2-
G56	110.72	8.425	3.74/3.999	n/a	n/a
V57	123.30	8.53	3.76	2.13	Hγ11 1.059; Hγ21 0.917
R58	119.73	8.452	-	1.88/1.81	Hγ2 1.66; Hγ1-; Hδ2 3.12; Hδ1 3.15; Nε 113.75; Hε 7.015; Nη2-; Hη21-; Hη22-; Nη1-; Hη11-; Hη12-

S59	113.93	8.182	4.179	3.857/3.922	Hγ-
G60	110.72	7.826	4.1/3.967	n/a	n/a
A61	123.97	8.156	4.058	1.44	n/a
A62	122.12	8.069	4.056	1.47	n/a
A63	120.53	7.755	4.052	1.443	n/a
L64	117.12	7.908	4.089	1.705/1.828	Hγ-; Hδ11 0.884; Hδ21 0.873
T65	113.55	7.804	3.807	4.297	Hγ1-; Hγ21 1.202
L66	120.92	7.678	3.954	1.307/1.484	Hγ 1.383; Hδ11 0.806; Hδ21 0.86
I67	120.06	7.915	3.859	2.092	Hγ12 1.141; Hγ11 1.113; Hδ11 0.815; Hγ21 1.009
V68	120.56	8.321	3.656	2.118	Hγ11 1.065; Hγ21 0.91
V69	124.60	8.72	3.548	2.3	Hγ11 1.077; Hγ21 0.946
W70	124.86	8.84	4.235	3.629/3.455	Hδ1 7.067; Hζ2 7.363; Hη2 7.12; Hζ3 7.022; Hε3 7.555; Nε1 127.124; Hε1 9.62
I71	119.45	9.249	3.543	1.997	Hγ12 1.209; Hγ11 0.934; Hδ11 0.877; Hγ21 0.888
T72	115.05	8.581	3.934	4.391	Hγ1-; Hγ21 1.25
S73	117.95	8.352	-	-	Hγ-
R74	121.30	7.742	-	1.73/1.61	Hγ2 1.38; Hγ1-; Hδ2 2.578; Hδ1-; Nε 114.484; Hε 6.646; Nη2-; Hη21-; Hη22-; Nη1-; Hη11-; Hη12-
S75	114.82	7.89	4.284	3.859/3.926	Hγ-
R76	120.86	7.715	4.208	1.843	Hγ2 1.536; Hγ1 1.575; Hδ2 3.095; Hδ1 3.156; Nε 114.256; Hε 7.042; Nη2-; Hη21-; Hη22-; Nη1-; Hη11-; Hη12-
K77	119.59	7.855	4.36	1.777/1.883	Hγ2 1.359; Hγ1 1.439; Hδ2 1.625; Hδ1 1.557; Hε2 3.069; Hε1 2.857; Nζ-; Hζ1-
T78	114.84	7.667	4.55	4.27	Hγ1-; Hγ21 1.253
P79	-	-	-	-	Hγ2-; Hγ1-; Hδ2-; Hδ1-
I80	117.48	7.437	3.866	1.84	Hγ12 1.237; Hγ11 1.491; Hδ11 0.845; Hγ21-
F81	120.17	7.573	4.336	3.205	Hδ1-; Hε1-; Hζ-; Hε2-; Hδ2-
I82	119.99	7.536	-	-	Hγ12-; Hγ11-; Hδ11 0.841; Hγ21-
I83	120.01	7.98	3.756	1.972	Hγ12 1.189; Hγ11 1.224; Hδ11 0.782; Hγ21 0.94
N84	120.11	8.244	4.405	2.876/2.713	Nδ2 124.64; Hδ21 6.375; Hδ22 7.361
Q85	119.42	7.924	4.015	2.1/2.1	Hγ2 2.17; Hγ1 2.17; Nε2 124.29; Hε21 6.27; Hε22 6.63

V86	120.53	8.525	3.728	2.129	Hγ11 1.02; Hγ21 0.974
S87	114.75	8.143	4.212	4.106	Hγ-
L88	121.82	7.752	4.157	1.814/1.817	Hγ 1.689; Hδ11 0.888; Hδ21 0.87
F89	119.45	8.058	4.331	3.209	Hδ1-; Hε1-; Hζ-, Hε2-; Hδ2-
L90	117.90	8.44			Hγ-; Hδ11 0.86; Hδ21 0.903
I91	123.61	8.256		2.14	Hγ12-; Hγ11-; Hδ11 0.796; Hγ21-
I92	125.47	8.567	3.616	1.998	Hγ12-; Hγ11-; Hδ11 0.845; Hγ21 0.889
L93	121.45	8.753		1.383	Hγ 1.291; Hδ11 0.73; Hδ21 0.86
H94	114.40	8.303	4.132	3.323/3.294	Hε1-; Hδ2-; Nδ1-; Hδ1-; Nε2-; Hε2-
S95	116.57	8.277		3.997/4.085	Hγ-
A96	126.04	8.488	4.207	1.537	
L97	118.43	8.19	4.087	1.699/1.832	Hγ-; Hδ11 0.829; Hδ21 0.804
Y98	120.61	8.096	4.378	3.019/3.086	Hδ1 6.913; Hε1 6.715; Hη-; Hε2 6.715; Hδ2 6.913
F99	119.10	8.332	4.263	3.214	Hδ1-; HE1-; Hζ-; Hε2-; Hδ2-
K100	118.68	8.194	3.892	1.994	Hγ2 1.391; Hγ1 1.558; Hδ2 1.636; Hδ1 1.689; Hε2 2.935; Hε1-, Nζ-; Hζ1-
Y101	122.51	8.279	4.1	3.106/3.279	Hδ1 6.975; Hε1 6.769; Hη-; Hε2 6.769; Hδ2 6.975
L102	122.94	8.433	3.772	1.428/1.743	Hγ 1.507; Hδ11 0.766; Hδ21 0.792
L103	119.05	8.382	4.005	1.455/1.626	Hγ 1.576; Hδ11 0.73;Hδ21 0.692
S104	114.63	8.003	4.182	3.854	Hγ-
N105	119.75	7.681	4.511	2.392/2.392	Nδ2 127.37; Hδ21 5.947; Hδ22 6.795
Y106	120.38	8.147		3.085/2.997	Hδ1 7.05; Hε1 6.749; Hη-; Hε2 6.749; Hδ2 7.05
S107	114.87	7.95	4.284	3.86/3.926	Hγ-
S108	116.62	7.837	4.452	3.895/3.944	Hγ-
V109	119.38	7.668	4.234	2.18	Hγ11 0.931; Hγ21 0.93
T110	114.99	7.697	4.451	4.301	Hγ1-; Hγ21 1.128

Supplementary Table S4. ^{13}C chemical shift assignments at 25°C.

amino acid in Ste2p	N	CO	Cα	Cβ	Others
G31	-	167.829	43.38	-	
N32	118.86	173.624	53.307	38.88	Cγ-
G33	109.50	172.637	45.847	-	
S34	116.18	173.074	59.314	63.57	
T35	116.72	173.066	63.68	69.56	Cγ2 21.316
I36	122.32	174.792	63.1	38.44	Cγ1 28.3; Cδ1 12.46; Cγ2 16.75
T37	116.08	173.61	64.24	69.65	Cγ2 21.53
F38	121.53	175.226	61.38	38.89	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
D39	117.84	176.103	56.01	37.04	Cγ-; Cδ-
E40	120.83	176.28	58.75	28.06	Cγ 33.046
L41	121.94	176.426	58.21	41.66	Cγ 27.071; Cδ1 23.556; Cδ2-
Q42	117.27	176.784	59.1	28.43	Cγ 33.976; Cδ-
G43	106.82	174.501	47.037	-	
L44	126.12	178.411	58.317	41.9	Cγ 27.029; Cδ1 23.609; Cδ2-
V45	123.54	175.994	67.346	31.7	Cγ1 22.068; Cγ2 20.568
N46	117.76	176.988	56.89	38.55	Cγ-
S47	116.96	173.751	62.21	62.8	
T48	120.14	174.265	67.14	68.92	Cγ2 20.399
V49	122.83	175.596	66.98	31.8	Cγ1 22.203; Cγ2 20.591
T50	114.63	173.969	67.73	69.11	Cγ2 21.109
Q51	118.84	175.82	59.61	28.68	Cγ 34.476; Cδ-
A52	123.20	178.972	55.586	17.59	
I53	121.42	176.662	65.32	37.89	Cγ1-; Cδ1 12.7; Cγ2-
L54	120.07	177.526	58.22	41.66	Cγ-; Cδ1 22.192; Cδ2 24.648
F55	119.59	177.217	61.31	38.92	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
G56	110.72	174.252	47.297	-	
V57	123.30	176.686	66.28	32.1	Cγ1 22.19; Cγ2 20.653
R58	119.73	176.438	59.08	30.05	Cγ 27.44; Cδ 43.26; Cζ-
S59	113.93	173.989	60.82	63.27	
G60	110.72	172.889	46.28	-	
A61	123.97	178.139	54.853	17.98	
A62	122.12	176.894	55.07	17.7	
A63	120.53	177.05	55.09	17.73	
L64	117.12	176.346	57.84	41.87	Cγ-; Cδ 123.513; Cδ2-
T65	113.55	173.268	67.14	68.92	Cγ2 23.478

L66	120.92	176.913	58.42	41.36	Cγ 26.848; Cδ1 23.045; Cδ2 24.104
I67	120.06	175.872	65.43	37.98	Cγ1 28.985; Cδ1 12.37; Cγ2 16.945
V68	120.56	178.611	67.44	31.6	Cγ1 22.208; Cγ2 20.597
V69	124.60	176.67	67.566	31.61	Cγ1 21.935; Cγ2 20.476
W70	124.86	177.006	61.835	28.89	Cγ-; Cδ1-; Cζ2-; Cη2-; Cζ3-; Cε3-; Cδ2-
I71	119.45	177.794	65.594	38	Cγ1-; Cδ1 12.96; Cγ2-
T72	115.05	175.118	66.42	68.85	Cγ2 20.3
S73	117.95	174.053	61.57	62.92	
R74	121.30	175.835	57.66	29.87	Cγ 29.87; Cδ 43.63; Cζ-
S75	114.82	172.921	60.03	63.59	
R76	120.86	174.093	56.81	30.21	Cγ 27.269; Cδ 43.175; Cζ-
K77	119.59	173.901	56.07	32.65	Cγ 25.033; Cδ 28.86; Cε 37.106
T78	114.84		61.51	69.46	Cγ2 21.229
P79	-	175.181			Cγ-; Cδ-
I80	117.48	174.753	63.43	37.95	Cγ1 28.37; Cδ1 12.25; Cγ2-
F81	120.17	175.461	60.49	39.05	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
I82	119.99	175.245	64.24	38	Cγ1-; Cδ1 11.99; Cγ2-
I83	120.01	177.022	64.64	37.91	Cγ1 28.427; Cδ1 12.14; Cγ2 16.757
N84	120.11	174.512	56.65	38.776	Cγ-
Q85	119.42	176.142	59.16	28.696	Cγ 34; Cδ-
V86	120.53	175.658	66.65	31.88	Cγ1 22.022; Cγ2 20.686
S87	114.75	173.462	62.75	62.75	
L88	121.82	176.163	58.11	41.73	Cγ 27.05; Cδ1 23.097; Cδ2 24.172
F89	119.45	175.103	61.29	38.68	Cγ-; Cδ1-; Cε1-; CZ-; Cε2-; Cδ2-
L90	117.90	178.704	58.3	41.8	Cγ-; Cδ1 22.103; Cδ2 25.186
I91	123.61	177.522	65.85	38.12	Cγ1-; Cδ1 12.67; Cγ2-
I92	125.47	176.573	65.873	37.41	Cγ1-; Cδ1 12.46; Cγ2 16.487
L93	121.45	176.911	58.24	41.85	Cγ-; Cδ1 23.084; Cδ2 24.133
H94	114.40	174.824	60.07	27.74	Cγ-; Cε1-; Cδ2-
S95	116.57	173.023	62.17	62.72	
A96	126.04	178.194	55.441	17.78	
L97	118.43	177.125	57.85	41.75	Cγ-; Cδ1 22.612; Cδ2 24.553
Y98	120.61	175.376	61.19	38.09	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
F99	119.10	175.374	61.47	38.55	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
K100	118.68	177.34	59.85	32.09	Cγ 25.113; Cδ 29.335; Cε 42.07
Y101	122.51	176.264	61.48	37.68	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
L102	122.94	177.631	58.39	41.84	Cγ 26.694; Cδ1 24.71; Cδ2 22.848
L103	119.05	177.903	57.543	41.8	Cγ 26.99; Cδ1 22.796; Cδ2 24.186
S104	114.63	173.277	61.06	63.15	

N105	119.75	173.886	55.15	39.46	Cγ-
Y106	120.38	174.705	59.82	38.33	Cγ-; Cδ1-; Cε1-; Cζ-; Cε2-; Cδ2-
S107	114.87	172.548	59.48	63.66	
S108	116.62	172.316	59.11	63.82	
V109	119.38	173.956	62.57	32.46	Cγ1 20.15; Cγ2 20.93
T110	114.99		60.56	70.31	Cγ2 21.26

Supplementary Table S5. Measurements of the dynamic properties of TM1-TM2 in TFE:water(0.1%TFA) (1:1, v:v).

	25°C				45°C	
Residue	T2 (msec)	Temperature Coefficient (HN) (ppm/°C)	H-D Exchange Rate (1/time (h))	H-NOE	T2 (msec)	H-NOE
G31						
N32	414.9±15.5	-6.25		-1.75	909.1±135.1	-1.70
G33	220.8±4.4	-7.45	2860	-0.48	581.4±33.9	-0.84
S34	163.7±3.5	-5.55	530	0.09	409.8±23.6	-0.45
T35	103.3±3.1	-3.65	253	-0.23	251.9±5.1	0.12
I36	63.7±2.8	-4.5	69.4	0.40	159.5±2.8	-0.12
T37	55.9±2.5	-3.65	32.2	0.34	144.9±2.9	0.30
F38	45.7±1.5	-8.15	21.8	0.55	121.7±3.7	0.15
D39	38.8±1.8	-7.5	18.7	0.53	80.6±3.9	0.32
E40	40.2±2.8	-4.15	22.2	0.48	105.4±1.6	0.22
L41	40.5±2.8	-7.3	3.94	0.55	84.0±5.0	0.18
Q42	35.1±3.0	-6.2	3.77	0.60	85.0±1.3	0.20
G43	34.4±2.4	-0.95	7.8	0.92	83.3±2.9	0.24
L44	32.2±3.2	-4.9	3.55	0.81	86.7±3.3	0.32
V45	29.2±2.0	-8.2	1.68	0.59	79.7±2.5	0.29
N46	32.1±2.2	-4.85	6.9	0.79	79.6±2.3	0.28
S47	37.0±2.9	-4.75	6.5	0.67	80.8±2.7	0.38
T48	36.6±1.5	-3.1	2.89	0.70	85.2±2.4	0.37
V49	33.1±2.2	-6.95	1.87	0.91	77.2±2.3	0.33
T50	44.2±2.4	-4.25	0.7	0.56	76.8±2.7	0.36
Q51	35.7±2.3	-2.75	0.383	0.86	81.7±2.0	0.29
A52	30.8±2.1	-2.55	0.7	0.67	75.0±2.0	0.40
I53	33.1±2.4	-6.9	0.093	0.69	73.0±2.7	0.45
L54	28.7±2.7	-5.6	0.254	0.69	68.0±2.3	0.43
F55	27.9±2.0	-4.5	1.18	0.82	75.8±1.8	0.29
G56	34.1±3.1	-3.65	13.6	0.66	78.2±2.1	0.27
V57	28.1±2.5	-5.15	12.7	0.70	70.3±1.9	0.27
R58	31.9±2.4	-7.2	23.8	0.73	77.3±1.6	0.35
S59	31.7±3.2	-5.1	10.9	0.62	94.3±2.2	0.27
G60	29.9±2.3	-1	16.2	0.74	87.2±2.6	0.20
A61	26.5±2.5	-3.9	65.9	0.60	77.7±2.2	0.23
A62	29.6±2.4	-4.4	27.5	0.47	75.2±2.3	0.33
A63	34.6±1.8	-3.25	13.6	0.52	82.8±0.9	0.25
L64	26.2±2.7	-5.1	2.64	0.48	71.2±2.5	0.20

T65	31.2±2.1	-3.25	0.82	0.73	76.5±2.5	0.28
L66	25.4±2.8	-3.55	0.43	0.51	71.7±1.9	0.37
I67	31.3±2.8	-4.35	2.23	0.62	106.3±5.0	0.20
V68	27.2±2.9	-6.4	0.095	0.61	65.8±2.6	0.29
V69	27.8±1.9	-7.2	0.065	0.63	70.9±2.5	0.39
W70	26.4±2.2	-4.9	0.105	0.71	68.0±2.3	0.28
I71	24.8±1.7	-6	0.0001	0.69	65.8±1.9	0.23
T72	27.5±2.1	-4.85	0.333	0.86	69.0±2.9	0.25
S73	33.3±2.2	-4.3	1.51	0.61	74.6±3.9	0.36
R74	34.6±2.4	-2.95	14.4	0.56	89.3±2.1	0.32
S75	41.2±2.6	-3.9	29.6	0.54	103.8±2.3	0.13
R76	44.2±2.2	-2.6	28	0.41	126.3±2.1	0.19
K77	59.2±2.8	-3	61	0.42	130.9±3.6	0.04
T78	55.9±2.8	-2.7	38	0.32	164.5±6.8	0.17
P79						
I80	34.5±1.9	-1.25	10.1	0.56	89.4±2.2	0.16
F81	36.6±2.2	-4.1	8.4	0.65	90.8±1.7	0.21
I82	37.5±2.1	-4.6	4.15	0.65	98.1±2.0	0.19
I83	31.2±3.2	-8.2	0.192	0.61	84.1±2.8	0.20
N84	34.6±2.2	-5.25	9.12	0.68	79.2±2.1	0.25
Q85	41.8±3.0	-1.15	5.9	0.52	90.0±2.6	0.32
V86	36.8±2.3	-8.3	1.15	0.72	81.9±2.3	0.37
S87	38.2±3.4	-4.25	7.3	0.76	83.7±2.3	0.33
L88	34.4±2.0	-3.4	0.592	0.60	81.1±2.6	0.36
F89	37.5±3.4	-3.2	0.047	0.69	83.5±1.6	0.40
L90	34.1±2.7	-5.6	0.28	0.72	76.1±2.8	0.38
I91	32.4±3.7	-4.2	0.006	0.81	78.9±2.9	0.32
I92	38.3±2.4	-4.75	0.0001	0.83	78.3±2.9	0.30
L93	32.4±2.2	-5.05	0.105	0.78	75.6±1.5	0.42
H94	36.6±2.8	-3.55	4.64	0.64	86.8±2.9	0.38
S95	41.8±3.3	-2.5	2.57	0.72	87.6±2.4	0.37
A96	37.6±2.8	-4.55	4.74	0.71	85.3±2.3	0.32
L97	39.4±2.3	-5.75	1.18	0.80	94.4±1.3	0.34
Y98	36.0±2.5	-5.95	0.9	0.74	85.8±2.2	0.39
F99	36.6±2.3	-5.4	1.33	0.55	85.9±2.3	0.34
K100	40.5±2.5	-4.75	1.75	0.76	93.5±2.8	0.21
Y101	42.6±2.9	-5.2	1.75	0.67	95.1±1.7	0.27
L102	41.8±2.6	-6.5	1.43	0.70	97.7±1.1	0.40
L103	37.7±1.9	-7.35	4.17	0.59	97.8±1.9	0.31
S104	44.2±2.4	-5.7	16.1	0.56	119.2±2.4	0.27
N105	55.6±3.1	-0.8	141	0.35	139.3±2.5	0.29

Y106	70.9±2.5	-6.5	160	0.47	159.0±2.8	0.25
S107	133.9±5.2	-5.65	12	0.25	281.7±7.9	-0.17
S108	175.4±2.5	-2.6	670	0.52	389.1±19.7	-0.51
V109	188.3±8.5	-2.25	235	-0.40	505.1±23.0	0.08
T110	384.6±8.9	-3.45	237	-0.82	480.8±16.2	-0.04

Supplementary Table S6. Angle restraints derived from TALOS 45°C

Residue	PHI	PSI	Δ PHI	Δ PSI	DIST	COUNT	CLASS	PHI lower	PHI upper	PSI lower	PSI upper
G31	9999	9999	0	0	0	0	None				
N32	-96	140	33	25	38.12	7	New	-129	-63	115	165
G33	-101	147	44	49	15.08	3	New	-145	-57	98	196
S34	-114	149	38	31	14.97	7	New	-152	-76	118	180
T35	-69	-37	12	16	9.88	5	New	-81	-57	-53	-21
I36	-78	-19	16	34	12.6	7	New	-94	-62	-53	15
T37	-73	-26	17	32	16.6	10	Good	-90	-56	-58	6
F38	-66	-37	13	26	17.65	10	Good	-79	-53	-63	-11
D39	-62	-37	6	10	20.48	10	Good	-68	-56	-47	-27
E40	-64	-35	6	13	22.13	10	Good	-70	-58	-48	-22
L41	-66	-38	6	9	13.56	10	Good	-72	-60	-47	-29
Q42	-65	-41	7	8	18.39	10	Good	-72	-58	-49	-33
G43	-67	-37	6	6	21.59	10	Good	-73	-61	-43	-31
L44	-66	-43	4	8	23.05	10	Good	-70	-62	-51	-35
V45	-64	-43	6	3	10	10	Good	-70	-58	-46	-40
N46	-61	-41	5	7	7.79	10	Good	-66	-56	-48	-34
S47	-72	-36	6	6	11.51	10	Good	-78	-66	-42	-30
T48	-66	-44	6	6	12.2	10	Good	-72	-60	-50	-38
V49	-61	-39	4	5	13.16	10	Good	-65	-57	-44	-34
T50	-65	-44	4	6	10.33	10	Good	-69	-61	-50	-38
Q51	-62	-43	5	6	11.28	10	Good	-67	-57	-49	-37
A52	-63	-38	5	5	10.78	10	Good	-68	-58	-43	-33
I53	-67	-43	8	4	11.66	10	Good	-75	-59	-47	-39
L54	-63	-39	7	7	10.53	10	Good	-70	-56	-46	-32
F55	-65	-39	8	10	19.13	10	Good	-73	-57	-49	-29
G56	-69	-40	8	8	23.21	10	Good	-77	-61	-48	-32
V57	-66	-38	3	12	22.75	10	Good	-69	-63	-50	-26
R58	-64	-36	5	4	11.71	10	Good	-69	-59	-40	-32
S59	-80	-20	20	20	19.32	10	Good	-100	-60	-40	0
G60	81	4	23	29	25.01	5	New	58	104	-25	33
A61	-65	-34	12	13	26.83	10	Good	-77	-53	-47	-21
A62	-64	-42	5	3	15.8	10	Good	-69	-59	-45	-39
A63	-64	-38	5	8	13.91	10	Good	-69	-59	-46	-30
L64	-64	-41	6	4	13.99	10	Good	-70	-58	-45	-37
T65	-61	-45	2	5	13.3	10	Good	-63	-59	-50	-40
L66	-64	-42	6	5	11.57	10	Good	-70	-58	-47	-37
I67	-65	-42	8	8	12.8	10	Good	-73	-57	-50	-34
V68	-65	-37	3	9	13.22	10	Good	-68	-62	-46	-28
V69	-70	-39	8	6	13.14	10	Good	-78	-62	-45	-33
W70	-64	-40	5	4	12.48	10	Good	-69	-59	-44	-36
I71	-63	-43	6	6	10.54	10	Good	-69	-57	-49	-37
T72	-65	-40	5	5	11.96	10	Good	-70	-60	-45	-35
S73	-62	-40	7	9	11.47	10	Good	-69	-55	-49	-31
R74	-71	-30	19	23	11.17	10	Good	-90	-52	-53	-7
S75	-76	-19	21	27	12.91	6	New	-97	-55	-46	8

R76	-97	-4	23	21	13.18	6	New	-120	-74	-25	17
K77	-90	152	40	15	15.87	5	New	-130	-50	137	167
T78	-104	130	33	32	44.4	8	New	-137	-71	98	162
P79	-60	143	7	12	49.63	5	New	-67	-53	131	155
I80	-60	-35	12	18	43.87	9	Good	-72	-48	-53	-17
F81	-73	-26	19	25	15.77	9	New	-92	-54	-51	-1
I82	-64	-36	8	8	16.23	10	Good	-72	-56	-44	-28
I83	-64	-38	7	8	13.45	10	Good	-71	-57	-46	-30
N84	-64	-40	9	5	10.08	10	Good	-73	-55	-45	-35
Q85	-67	-37	7	13	9.63	10	Good	-74	-60	-50	-24
V86	-64	-40	7	7	11.03	10	Good	-71	-57	-47	-33
S87	-61	-39	3	3	13.92	10	Good	-64	-58	-42	-36
L88	-66	-43	7	5	13.89	10	Good	-73	-59	-48	-38
F89	-65	-43	7	6	14.99	10	Good	-72	-58	-49	-37
L90	-65	-39	4	6	11.83	10	Good	-69	-61	-45	-33
I91	-69	-41	7	7	12.81	10	Good	-76	-62	-48	-34
I92	-64	-43	5	5	10.87	10	Good	-69	-59	-48	-38
L93	-62	-41	6	6	11.86	10	Good	-68	-56	-47	-35
H94	-64	-42	7	6	16.28	10	Good	-71	-57	-48	-36
S95	-64	-40	6	4	19.29	10	Good	-70	-58	-44	-36
A96	-65	-42	5	6	15.83	10	Good	-70	-60	-48	-36
L97	-64	-39	6	5	11.48	10	Good	-70	-58	-44	-34
Y98	-66	-42	7	6	11.71	10	Good	-73	-59	-48	-36
F99	-66	-42	8	4	10.82	10	Good	-74	-58	-46	-38
K100	-60	-41	8	12	10.83	10	Good	-68	-52	-53	-29
Y101	-66	-42	8	7	10.61	10	Good	-74	-58	-49	-35
L102	-66	-38	7	9	11.87	10	Good	-73	-59	-47	-29
L103	-65	-35	9	4	12.19	10	Good	-74	-56	-39	-31
S104	-71	-34	11	9	11.3	10	Good	-82	-60	-43	-25
N105	-82	-13	17	26	10.93	6	New	-99	-65	-39	13
Y106	-85	-10	18	33	10.64	8	New	-103	-67	-43	23
S107	-86	143	21	34	10.17	5	New	-107	-65	109	177
S108	-97	148	40	33	8.39	6	New	-137	-57	115	181
V109	-121	142	33	23	13.1	8	New	-154	-88	119	165
T110	9999	9999	0	0	0	0	None	9999	9999	9999	9999

Supplementary Table S7. Angle constraints derived from TALOS at 25°C.

Residue	PHI	PSI	Δ PHI	Δ PSI	DIST	COUNT	CLASS	PHI lower	PHI upper	PSI lower	PSI upper
G31	9999	9999	0	0	0	0	None				
N32	-87	129	31	11	49.31	10	New	-118	-56	118	140
G33	142	-69	86	72	34.47	10	New	56	228	-141	3
S34	-95	-3	15	16	26.48	6	New	-110	-80	-19	13
T35	-83	-20	12	18	18.25	5	New	-95	-71	-38	-2
I36	-109	154	47	26	21.8	5	New	-156	-62	128	180
T37	-76	-34	13	24	32.91	8	New	-89	-63	-58	-10
F38	-62	-46	6	5	33.1	10	Good	-68	-56	-51	-41
D39	-62	-33	4	10	29.71	9	Good	-66	-58	-43	-23
E40	-66	-35	9	14	25.5	8	Good	-75	-57	-49	-21
L41	-64	-40	5	3	15.61	10	Good	-69	-59	-43	-37
Q42	-80	-14	21	23	32.45	9	New	-101	-59	-37	9
G43	95	-16	8	11	35.55	4	New	87	103	-27	-5
L44	-67	-27	6	15	30.98	8	Good	-73	-61	-42	-12
V45	-63	-43	4	8	16.38	10	Good	-67	-59	-51	-35
N46	-63	-43	7	6	17.29	10	Good	-70	-56	-49	-37
S47	-66	-40	6	5	23.06	10	Good	-72	-60	-45	-35
T48	-65	-42	5	7	25.32	9	Good	-70	-60	-49	-35
V49	-64	-39	5	5	24.58	10	Good	-69	-59	-44	-34
T50	-63	-39	5	4	19.22	10	Good	-68	-58	-43	-35
Q51	-62	-42	5	5	15.61	10	Good	-67	-57	-47	-37
A52	-63	-39	4	6	22.36	10	Good	-67	-59	-45	-33
I53	-63	-46	5	3	29.37	10	Good	-68	-58	-49	-43
L54	-60	-38	3	10	28.8	10	Good	-63	-57	-48	-28
F55	-69	-38	7	9	37.39	7	Good	-76	-62	-47	-29
G56	-64	-40	6	6	35.73	6	New	-70	-58	-46	-34
V57	-63	-41	15	15	38.27	10	Good	-78	-48	-56	-26
R58	-64	-38	7	11	25.4	10	Good	-71	-57	-49	-27
S59	-91	4	19	12	39.68	6	New	-110	-72	-8	16
G60	96	-11	5	11	34.72	6	New	91	101	-22	0
A61	-66	-30	6	14	32.43	5	New	-72	-60	-44	-16
A62	-65	-35	4	10	16.97	10	Good	-69	-61	-45	-25
A63	-65	-33	8	14	17.73	10	Good	-73	-57	-47	-19
L64	-69	-34	14	21	20.53	10	Good	-83	-55	-55	-13
T65	-67	-42	10	5	21.13	10	Good	-77	-57	-47	-37
L66	-60	-39	5	3	19.46	10	Good	-65	-55	-42	-36
I67	-65	-43	6	3	16.01	10	Good	-71	-59	-46	-40
V68	-64	-41	5	3	16.08	10	Good	-69	-59	-44	-38
V69	-66	-42	9	7	13.42	10	Good	-75	-57	-49	-35
W70	-61	-41	5	4	12.47	10	Good	-66	-56	-45	-37
I71	-63	-42	6	4	11.16	10	Good	-69	-57	-46	-38

T72	-63	-42	5	5	23.67	10	Good	-68	-58	-47	-37
S73	-66	-40	10	4	34.37	9	New	-76	-56	-44	-36
R74	-78	-18	17	21	34.54	10	New	-95	-61	-39	3
S75	-80	-11	20	22	24.66	5	New	-100	-60	-33	11
R76	-98	150	50	22	19.46	5	New	-148	-48	128	172
K77	-108	138	43	36	25.48	9	New	-151	-65	102	174
T78	-100	125	28	21	61.34	9	New	-128	-72	104	146
P79	-70	151	25	11	70.71	9	New	-95	-45	140	162
I80	-77	156	24	32	56.4	5	New	-101	-53	124	188
F81	-74	-30	17	33	30.45	6	New	-91	-57	-63	3
I82	-68	-39	9	14	27.86	10	Good	-77	-59	-53	-25
I83	-62	-42	4	6	26.46	10	Good	-66	-58	-48	-36
N84	-63	-41	4	3	15.43	9	Good	-67	-59	-44	-38
Q85	-63	-41	5	4	17.95	10	Good	-68	-58	-45	-37
V86	-68	-41	8	4	20.74	10	Good	-76	-60	-45	-37
S87	-64	-36	8	8	22.31	9	Good	-72	-56	-44	-28
L88	-66	-39	8	5	20.2	10	Good	-74	-58	-44	-34
F89	-62	-43	4	7	25.83	10	Good	-66	-58	-50	-36
L90	-63	-42	6	9	32.53	10	Good	-69	-57	-51	-33
I91	-62	-42	4	6	31.53	10	Good	-66	-58	-48	-36
I92	-63	-39	6	5	30.93	10	Good	-69	-57	-44	-34
L93	-60	-41	4	4	23.43	10	Good	-64	-56	-45	-37
H94	-62	-39	4	3	38.03	10	Good	-66	-58	-42	-36
S95	-66	-41	8	4	35.32	9	Good	-74	-58	-45	-37
A96	-64	-37	5	10	27.89	10	Good	-69	-59	-47	-27
L97	-66	-40	6	2	14.05	10	Good	-72	-60	-42	-38
Y98	-67	-41	5	4	15.59	10	Good	-72	-62	-45	-37
F99	-63	-42	6	5	14.76	10	Good	-69	-57	-47	-37
K100	-62	-42	6	4	13.34	10	Good	-68	-56	-46	-38
Y101	-66	-42	4	4	12.14	10	Good	-70	-62	-46	-38
L102	-62	-40	7	2	10.49	10	Good	-69	-55	-42	-38
L103	-62	-36	6	8	13.16	10	Good	-68	-56	-44	-28
S104	-70	-28	9	13	16.86	10	New	-79	-61	-41	-15
N105	-75	-24	9	18	18.61	8	New	-84	-66	-42	-6
Y106	-81	-26	13	17	19.59	6	New	-94	-68	-43	-9
S107	-96	137	45	33	17.57	5	New	-141	-51	104	170
S108	-114	156	36	25	18.34	5	New	-150	-78	131	181
V109	-115	140	34	23	25.48	6	New	-149	-81	117	163
T110	9999	9999	0	0	0	0	None	9999	9999	9999	9999

References

1. Schultz, J., Ferguson, B., and Sprague, G. F., Jr. (1995) Signal transduction and growth control in yeast, *Current opinion in genetics & development 5*, 31-37.

2. Elion, E. A. (2000) Pheromone response, mating and cell biology, *Current opinion in microbiology 3*, 573-581.

3. Dohlman, H. G. (2002) G proteins and pheromone signaling, *Annual review of physiology 64*, 129-152.

4. Gether, U. (2000) Uncovering molecular mechanisms involved in activation of G protein-coupled receptors, *Endocrine reviews 21*, 90-113.

5. Lagerstrom, M. C., and Schioth, H. B. (2008) Structural diversity of G protein-coupled receptors and significance for drug discovery, *Nature reviews 7*, 339-357.

6. Arevalo, E., Estephan, R., Madeo, J., Arshava, B., Dumont, M., Becker, J. M., and Naider, F. (2003) Biosynthesis and biophysical analysis of domains of a yeast G protein-coupled receptor, *Biopolymers 71*, 516-531.

7. Ballesteros, J. A., Jensen, A. D., Liapakis, G., Rasmussen, S. G., Shi, L., Gether, U., and Javitch, J. A. (2001) Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6, *J Biol Chem 276*, 29171-29177.

8. Eilers, M., Hornak, V., Smith, S. O., and Konopka, J. B. (2005) Comparison of class A and D G protein-coupled receptors: common features in structure and activation, *Biochemistry 44*, 8959-8975.

9. Henry, L. K., Khare, S., Son, C., Babu, V. V., Naider, F., and Becker, J. M. (2002) Identification of a contact region between the tridecapeptide alpha-factor mating pheromone of Saccharomyces cerevisiae and its G protein-coupled receptor by photoaffinity labeling, *Biochemistry 41*, 6128-6139.

10. Chen, S., Lin, F., Xu, M., Riek, R. P., Novotny, J., and Graham, R. M. (2002) Mutation of a single TMVI residue, Phe(282), in the beta(2)-adrenergic receptor results in structurally distinct activated receptor conformations, *Biochemistry 41*, 6045-6053.

11. Lin, J. C., Duell, K., and Konopka, J. B. (2004) A microdomain formed by the extracellular ends of the transmembrane domains promotes activation of the G protein-coupled alpha-factor receptor, *Molecular and cellular biology 24*, 2041-2051.

12. O'Dowd, B. F., Hnatowich, M., Regan, J. W., Leader, W. M., Caron, M. G., and Lefkowitz, R. J. (1988) Site-directed mutagenesis of the cytoplasmic domains of the human beta 2-adrenergic receptor. Localization of regions involved in G protein-receptor coupling, *J Biol Chem 263*, 15985-15992.

13. Schandel, K. A., and Jenness, D. D. (1994) Direct evidence for ligand-induced internalization of the yeast alpha-factor pheromone receptor, *Molecular and cellular biology 14*, 7245-7255.

14. Stefan, C. J., and Blumer, K. J. (1994) The third cytoplasmic loop of a yeast G-protein-coupled receptor controls pathway activation, ligand discrimination, and receptor internalization, *Molecular and cellular biology 14*, 3339-3349.

15. Chen, Q., and Konopka, J. B. (1996) Regulation of the G-protein-coupled alpha-factor pheromone receptor by phosphorylation, *Molecular and cellular biology 16*, 247-257.

16. Hicke, L., and Riezman, H. (1996) Ubiquitination of a yeast plasma membrane receptor signals its ligand-stimulated endocytosis, *Cell 84*, 277-287.

17. Dosil, M., Giot, L., Davis, C., and Konopka, J. B. (1998) Dominant-negative mutations in the G-protein-coupled alpha-factor receptor map to the extracellular ends of the transmembrane segments, *Molecular and cellular biology 18*, 5981-5991.

18. Lee, B. K., Lee, Y. H., Hauser, M., Son, C. D., Khare, S., Naider, F., and Becker, J. M. (2002) Tyr266 in the sixth transmembrane domain of the yeast alpha-factor receptor plays key roles in receptor activation and ligand specificity, *Biochemistry 41*, 13681-13689.

19. Son, C. D., Sargsyan, H., Naider, F., and Becker, J. M. (2004) Identification of ligand binding regions of the Saccharomyces cerevisiae alpha-factor pheromone receptor by photoaffinity cross-linking, *Biochemistry 43*, 13193-13203.

20. Huang, L. Y., Umanah, G., Hauser, M., Son, C., Arshava, B., Naider, F., and Becker, J. M. (2008) Unnatural amino acid replacement in a yeast G protein-coupled receptor in its native environment, *Biochemistry 47*, 5638-5648.

21. Umanah, G. K., Son, C., Ding, F., Naider, F., and Becker, J. M. (2009) Cross-linking of a DOPA-containing peptide ligand into its G protein-coupled receptor, *Biochemistry 48*, 2033-2044.

22. Tkacz, J. S., and MacKay, V. L. (1979) Sexual conjugation in yeast. Cell surface changes in response to the action of mating hormones, *The Journal of cell biology 80*, 326-333.

23. Moore, S. A. (1983) Comparison of dose-response curves for alpha factor-induced cell division arrest, agglutination, and projection formation of yeast cells. Implication for the mechanism of alpha factor action, *J Biol Chem 258*, 13849-13856.

24. Manney, T. R. (1983) Expression of the BAR1 gene in Saccharomyces cerevisiae: induction by the alpha mating pheromone of an activity associated with a secreted protein, *Journal of bacteriology 155*, 291-301.

25. Lee, B. K., Khare, S., Naider, F., and Becker, J. M. (2001) Identification of residues of the Saccharomyces cerevisiae G protein-coupled receptor contributing to alpha-factor pheromone binding, *J Biol Chem 276*, 37950-37961.

26. McCaffrey, G., Clay, F. J., Kelsay, K., and Sprague, G. F., Jr. (1987) Identification and regulation of a gene required for cell fusion during mating of the yeast Saccharomyces cerevisiae, *Molecular and cellular biology 7*, 2680-2690.

27. Hauser, M., Kauffman, S., Lee, B. K., Naider, F., and Becker, J. M. (2007) The first extracellular loop of the Saccharomyces cerevisiae G protein-coupled receptor Ste2p undergoes a conformational change upon ligand binding, *J Biol Chem 282*, 10387-10397.

28. Jenness, D. D., Burkholder, A. C., and Hartwell, L. H. (1983) Binding of alpha-factor pheromone to yeast a cells: chemical and genetic evidence for an alpha-factor receptor, *Cell 35*, 521-529.

29. Jenness, D. D., Burkholder, A. C., and Hartwell, L. H. (1986) Binding of alpha-factor pheromone to Saccharomyces cerevisiae a cells: dissociation constant and number of binding sites, *Molecular and cellular biology 6*, 318-320.

30. Naider, F., and Becker, J. M. (1986) Structure-activity relationships of the yeast alpha-factor, *CRC critical reviews in biochemistry 21*, 225-248.

31. Blumer, K. J., Reneke, J. E., and Thorner, J. (1988) The STE2 gene product is the ligand-binding component of the alpha-factor receptor of Saccharomyces cerevisiae, *J Biol Chem 263*, 10836-10842.

32. Raths, S. K., Naider, F., and Becker, J. M. (1988) Peptide analogues compete with the binding of alpha-factor to its receptor in Saccharomyces cerevisiae, *J Biol Chem 263*, 17333-17341.

33. Bajaj, A., Celic, A., Ding, F. X., Naider, F., Becker, J. M., and Dumont, M. E. (2004) A fluorescent alpha-factor analogue exhibits multiple steps on binding to its G protein coupled receptor in yeast, *Biochemistry 43*, 13564-13578.

34. Gao, F. P., and Cross, T. A. (2005) Recent developments in membrane-protein structural genomics, *Genome biology 6*, 244.

35. Hopkins, A. L., and Groom, C. R. (2002) The druggable genome, *Nature reviews 1*, 727-730.

36. Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many drug targets are there?, *Nature reviews 5*, 993-996.

37. McCusker, E. C., Bane, S. E., O'Malley, M. A., and Robinson, A. S. (2007) Heterologous GPCR expression: a bottleneck to obtaining crystal structures, *Biotechnology progress 23*, 540-547.

38. Sarramegna, V., Talmont, F., Demange, P., and Milon, A. (2003) Heterologous expression of G-protein-coupled receptors: comparison of expression systems fron the standpoint of large-scale production and purification, *Cell Mol Life Sci 60*, 1529-1546.

39. Kiefer, H., Krieger, J., Olszewski, J. D., Von Heijne, G., Prestwich, G. D., and Breer, H. (1996) Expression of an olfactory receptor in Escherichia coli: purification, reconstitution, and ligand binding, *Biochemistry 35*, 16077-16084.

40. Bane, S. E., Velasquez, J. E., and Robinson, A. S. (2007) Expression and purification of milligram levels of inactive G-protein coupled receptors in E. coli, *Protein expression and purification 52*, 348-355.

41. David, N. E., Gee, M., Andersen, B., Naider, F., Thorner, J., and Stevens, R. C. (1997) Expression and purification of the Saccharomyces cerevisiae alpha-factor receptor

(Ste2p), a 7-transmembrane-segment G protein-coupled receptor, *J Biol Chem 272*, 15553-15561.

42. Andre, N., Cherouati, N., Prual, C., Steffan, T., Zeder-Lutz, G., Magnin, T., Pattus, F., Michel, H., Wagner, R., and Reinhart, C. (2006) Enhancing functional production of G protein-coupled receptors in Pichia pastoris to levels required for structural studies via a single expression screen, *Protein Sci 15*, 1115-1126.

43. Wedekind, A., O'Malley, M. A., Niebauer, R. T., and Robinson, A. S. (2006) Optimization of the human adenosine A2a receptor yields in Saccharomyces cerevisiae, *Biotechnology progress 22*, 1249-1255.

44. Ohtaki, T., Ogi, K., Masuda, Y., Mitsuoka, K., Fujiyoshi, Y., Kitada, C., Sawada, H., Onda, H., and Fujino, M. (1998) Expression, purification, and reconstitution of receptor for pituitary adenylate cyclase-activating polypeptide. large-scale purification of a functionally active G protein-coupled receptor produced in Sf9 insect cells, *J Biol Chem 273*, 15464-15473.

45. Akermoun, M., Koglin, M., Zvalova-Iooss, D., Folschweiller, N., Dowell, S. J., and Gearing, K. L. (2005) Characterization of 16 human G protein-coupled receptors expressed in baculovirus-infected insect cells, *Protein expression and purification 44*, 65-74.

46. Chelikani, P., Reeves, P. J., Rajbhandary, U. L., and Khorana, H. G. (2006) The synthesis and high-level expression of a beta2-adrenergic receptor gene in a tetracycline-inducible stable mammalian cell line, *Protein Sci 15*, 1433-1440.

47. Hassaine, G., Wagner, R., Kempf, J., Cherouati, N., Hassaine, N., Prual, C., Andre, N., Reinhart, C., Pattus, F., and Lundstrom, K. (2006) Semliki Forest virus vectors for overexpression of 101 G protein-coupled receptors in mammalian host cells, *Protein expression and purification 45*, 343-351.

48. Ishihara, G., Goto, M., Saeki, M., Ito, K., Hori, T., Kigawa, T., Shirouzu, M., and Yokoyama, S. (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors, *Protein expression and purification 41*, 27-37.

49. Klammt, C., Schwarz, D., Fendler, K., Haase, W., Dotsch, V., and Bernhard, F. (2005) Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system, *The FEBS journal 272*, 6024-6038.

50. Carpenter, E. P., Beis, K., Cameron, A. D., and Iwata, S. (2008) Overcoming the challenges of membrane protein crystallography, *Current opinion in structural biology 18*, 581-586.

51. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor, *Science 289*, 739-745.

52. Li, J., Edwards, P. C., Burghammer, M., Villa, C., and Schertler, G. F. (2004) Structure of bovine rhodopsin in a trigonal crystal form, *Journal of molecular biology 343*, 1409-1438.

53. Okada, T., Sugihara, M., Bondar, A. N., Elstner, M., Entel, P., and Buss, V. (2004) The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure, *Journal of molecular biology 342*, 571-583.

54. Stenkamp, R. E. (2008) Alternative models for two crystal structures of bovine rhodopsin, *Acta Crystallogr D Biol Crystallogr D64*, 902-904.

55. Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G., Thian, F. S., Kobilka, T. S., Choi, H. J., Kuhn, P., Weis, W. I., Kobilka, B. K., and Stevens, R. C. (2007) High-Resolution Crystal Structure of an Engineered Human {beta}2-Adrenergic G Protein Coupled Receptor, *Science 318(5854)*, 1258-1265.

56. Warne, T., Serrano-Vega, M. J., Baker, J. G., Moukhametzianov, R., Edwards, P. C., Henderson, R., Leslie, A. G., Tate, C. G., and Schertler, G. F. (2008) Structure of a beta1-adrenergic G-protein-coupled receptor, *Nature 454*, 486-491.

57. Scheerer, P., Park, J. H., Hildebrand, P. W., Kim, Y. J., Krauss, N., Choe, H. W., Hofmann, K. P., and Ernst, O. P. (2008) Crystal structure of opsin in its G-protein-interacting conformation, *Nature 455*, 497-502.

58. Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H. W., and Ernst, O. P. (2008) Crystal structure of the ligand-free G-protein-coupled receptor opsin, *Nature 454*, 183-187.

59. Jaakola, V. P., Griffith, M. T., Hanson, M. A., Cherezov, V., Chien, E. Y., Lane, J. R., Ijzerman, A. P., and Stevens, R. C. (2008) The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist, *Science 322(5905)*,:121-127.

60. Rasmussen, S. G., Choi, H. J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R., Sanishvili, R., Fischetti, R. F., Schertler, G. F., Weis, W. I., and Kobilka, B. K. (2007) Crystal structure of the human beta(2) adrenergic G-protein-coupled receptor, *Nature 450(7168)*, 383-387.

61. Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G., Thian, F. S., Kobilka, T. S., Choi, H. J., Yao, X. J., Weis, W. I., Stevens, R. C., and Kobilka, B. K. (2007) GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function, *Science 318*, 1266-1273.

62. Rosenbaum, D. M., Rasmussen, S. G., and Kobilka, B. K. (2009) The structure and function of G-protein-coupled receptors, *Nature 459*, 356-363.

63. Topiol, S., and Sabio, M. (2009) X-ray structure breakthroughs in the GPCR transmembrane region, *Biochem Pharmacol 78*, 11-20.

64. Lodowski, D. T., Angel, T. E., and Palczewski, K. (2009) Comparative analysis of GPCR crystal structures, *Photochem Photobiol 85*, 425-430.

65. Goncalves, J. A., Ahuja, S., Erfani, S., Eilers, M., and Smith, S. O. (2010) Structure and function of G protein-coupled receptors using NMR spectroscopy, *Prog Nucl Magn Reson Spectrosc 57*, 159-180.

66. Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR, *Science 274*, 1531-1534.

67. Bhattacharya, S., Botuyan, M. V., Hsu, F., Shan, X., Arunkumar, A. I., Arrowsmith, C. H., Edwards, A. M., and Chazin, W. J. (2002) Characterization of binding-induced changes in dynamics suggests a model for sequence-nonspecific binding of ssDNA by replication protein A, *Protein Sci 11*, 2316-2325.

68. Mittag, T., Orlicky, S., Choy, W. Y., Tang, X., Lin, H., Sicheri, F., Kay, L. E., Tyers, M., and Forman-Kay, J. D. (2008) Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor, *Proceedings of the National Academy of Sciences of the United States of America 105*, 17772-17777.

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

70. McIntosh, L. P., Griffey, R. H., Muchmore, D. C., Nielson, C. P., Redfield, A. G., and Dahlquist, F. W. (1987) Proton NMR measurements of bacteriophage T4 lysozyme aided by 15N isotopic labeling: structural and dynamic studies of larger proteins, *Proceedings of the National Academy of Sciences of the United States of America 84*, 1244-1248.

71. Studts, J. M., and Fox, B. G. (1999) Application of fed-batch fermentation to the preparation of isotopically labeled or selenomethionyl-labeled proteins, *Protein expression and purification 16*, 109-119.

72. Cheng, H., Westler, W. M., Xia, B., Oh, B. H., and Markley, J. L. (1995) Protein expression, selective isotopic labeling, and analysis of hyperfine-shifted NMR signals of Anabaena 7120 vegetative [2Fe-2S]ferredoxin, *Archives of biochemistry and biophysics 316*, 619-634.

73. Peterson, F. C., Gordon, N. C., and Gettins, P. G. (2001) High-level bacterial expression and 15N-alanine-labeling of bovine trypsin. Application to the study of trypsin-inhibitor complexes and trypsinogen activation by NMR spectroscopy, *Biochemistry 40*, 6275-6283.

74. Englander, J., Cohen, L., Arshava, B., Estephan, R., Becker, J. M., and Naider, F. (2006) Selective labeling of a membrane peptide with 15N-amino acids using cells grown in rich medium, *Biopolymers 84*, 508-518.

75. Tugarinov, V., Kanelis, V., and Kay, L. E. (2006) Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy, *Nature protocols 1*, 749-754.

76. Xu, R., Ayers, B., Cowburn, D., and Muir, T. W. (1999) Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies, *Proceedings of the National Academy of Sciences of the United States of America 96*, 388-393.

77. Camarero, J. A., Shekhtman, A., Campbell, E. A., Chlenov, M., Gruber, T. M., Bryant, D. A., Darst, S. A., Cowburn, D., and Muir, T. W. (2002) Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR, *Proceedings of the National Academy of Sciences of the United States of America 99*, 8536-8541.

78. Romanelli, A., Shekhtman, A., Cowburn, D., and Muir, T. W. (2004) Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction, *Proceedings of the National Academy of Sciences of the United States of America 101*, 6397-6402.

79. Balambika, R., Inui, T, Sargsyan, H, Arshava, B, Cohen, LS, Ding, FX, Becker, JM, Naider, F (2007) Synthesis of a Double Transmembrane Domain Fragment of Ste2p by Native Chemical Ligation, *International Journal of Peptide Research and Therapeutics 13*, 251-263.

80. Luca, S., White, J. F., Sohal, A. K., Filippov, D. V., van Boom, J. H., Grisshammer, R., and Baldus, M. (2003) The conformation of neurotensin bound to its G protein-coupled receptor, *Proceedings of the National Academy of Sciences of the United States of America 100*, 10706-10711.

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

82. Williamson, P. T., Verhoeven, A., Miller, K. W., Meier, B. H., and Watts, A. (2007) The conformation of acetylcholine at its target site in the membrane-embedded nicotinic acetylcholine receptor, *Proceedings of the National Academy of Sciences of the United States of America 104*, 18031-18036.

83. Vosegaard, T., Kamihira-Ishijima, M., Watts, A., and Nielsen, N. C. (2008) Helix conformations in 7TM membrane proteins determined using oriented-sample solid-state NMR with multiple residue-specific 15N labeling, *Biophys J 94*, 241-250.

84. Oxenoid, K., Kim, H. J., Jacob, J., Sonnichsen, F. D., and Sanders, C. R. (2004) NMR assignments for a helical 40 kDa membrane protein, *Journal of the American Chemical Society 126*, 5048-5049.

85. 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.

86. Poget, S. F., Cahill, S. M., and Girvin, M. E. (2007) Isotropic bicelles stabilize the functional form of a small multidrug-resistance pump for NMR structural studies, *Journal of the American Chemical Society 129*, 2432-2433.

87. Schwaiger, M., Lebendiker, M., Yerushalmi, H., Coles, M., Groger, A., Schwarz, C., Schuldiner, S., and Kessler, H. (1998) NMR investigation of the multidrug transporter EmrE, an integral membrane protein, *European journal of biochemistry / FEBS 254*, 610-619.

88. Chill, J. H., Louis, J. M., Miller, C., and Bax, A. (2006) NMR study of the tetrameric KcsA potassium channel in detergent micelles, *Protein Sci 15*, 684-698.

89. Girvin, M. E., Rastogi, V. K., Abildgaard, F., Markley, J. L., and Fillingame, R. H. (1998) Solution structure of the transmembrane H+-transporting subunit c of the F1F0 ATP synthase, *Biochemistry 37*, 8817-8824.

90. Tian, C., Breyer, R. M., Kim, H. J., Karra, M. D., Friedman, D. B., Karpay, A., and Sanders, C. R. (2005) Solution NMR spectroscopy of the human vasopressin V2 receptor, a G protein-coupled receptor, *Journal of the American Chemical Society 127*, 8010-8011.

91. Tian, C., Breyer, R. M., Kim, H. J., Karra, M. D., Friedman, D. B., Karpay, A., and Sanders, C. R. (2006) Solution NMR spectroscopy of the human vasopressin v2 receptor, a g protein-coupled receptor [j. Am. Chem. Soc. 2005, 127, 8010-8011], *Journal of the American Chemical Society 128*, 5300.

92. Gautier, A., Kirkpatrick, J. P., and Nietlispach, D. (2008) Solution-state NMR spectroscopy of a seven-helix transmembrane protein receptor: backbone assignment, secondary structure, and dynamics, *Angewandte Chemie (International ed) 47*, 7297-7300.

93. Gautier, A., Mott, H. R., Bostock, M. J., Kirkpatrick, J. P., and Nietlispach, D. (2010) Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy, *Nat Struct Mol Biol 17*, 768-774.

94. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D., and Kay, L. E. (1994) A suite of triple resonance NMR experiments for the backbone assignment of 15N, 13C, 2H labeled proteins with high sensitivity, *Journal of the American Chemical Society 116*, 11655-11666.

95. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure, *Journal of molecular biology 222*, 311-333.

96. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, *Biochemistry 31*, 1647-1651.

97. Wishart, D. S., and Sykes, B. D. (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data, *Journal of biomolecular NMR 4*, 171-180.

98. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *Journal of biomolecular NMR 13*, 289-302.

99. Shan, X., Gardner, K., Muhandiram, D., Rao, N., Arrowsmith, C. H., and Kay, L. E. (1996) Assignment of N, C, C, and NH resonances in an N, C, H labeled 64 kDa Trp repressor-operator complex using triple-resonance NMR spectroscopy and H-decoupling, *Journal of the American Chemical Society 118*, 6570-6579.

100. Ikura, M., Spera, S., Barbato, G., Kay, L. E., Krinks, M., and Bax, A. (1991) Secondary structure and side-chain 1H and 13C resonance assignments of calmodulin in solution by heteronuclear multidimensional NMR spectroscopy, *Biochemistry 30*, 9216-9228.

101. Kay, L. E., Xu, G., Singer, A., Muhandiram, D., and Formankay, J. (1993) A gradient-enhanced HCCH-TOCSY experiment for recording side-chain 1H and 13C correlations in H2O samples of proteins, *Journal of Magnetic Resonance, Series B 101*, 333-337.

102. Clore, G. M., Bax, A., Driscoll, P. C., Wingfield, P. T., and Gronenborn, A. M. (1990) Assignment of the side-chain 1H and 13C resonances of interleukin-1 beta using double- and triple-resonance heteronuclear three-dimensional NMR spectroscopy, *Biochemistry 29*, 8172-8184.

103. Baldisseri, D. M., Pelton, J. G., Sparks, S. W., and Torchia, D. A. (1991) Complete 1H and 13C assignment of Lys and Leu sidechains of staphylococcal nuclease using HCCH-COSY and HCCH-TOCSY 3D NMR spectroscopy, *FEBS Lett 281*, 33-38.

104. Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of backbone amide and aliphatic side-chain resonances in 13C/15N-enriched proteins by isotropic mixing of 13C magnetization, *Journal of Magnetic Resonance, Series B 101*, 114-119.

105. Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1993) A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments, *Journal of biomolecular NMR 3*, 225-231.

106. Tugarinov, V., and Kay, L. E. (2003) Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods, *Journal of the American Chemical Society 125*, 13868-13878.

107. Yamazaki, T., Yoshida, M., and Nagayama, K. (1993) Complete assignments of magnetic resonances of ribonuclease H from Escherichia coli by double- and triple-resonance 2D and 3D NMR spectroscopies, *Biochemistry 32*, 5656-5669.

108. Prompers, J. J., Groenewegen, A., Hilbers, C. W., and Pepermans, H. A. M. (1998) Two-Dimensional NMR Experiments for the Assignment of Aromatic Side Chains in 13C-labeled Proteins, *J Magn Reson 130*, 68-75.

109. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA, *Journal of molecular biology 273*, 283-298.

110. Popot, J. L., and Engelman, D. M. (2000) Helical membrane protein folding, stability, and evolution, *Annual review of biochemistry 69*, 881-922.

111. Hunt, J. F., Rath, P., Rothschild, K. J., and Engelman, D. M. (1997) Spontaneous, pH-dependent membrane insertion of a transbilayer alpha-helix, *Biochemistry 36*, 15177-15192.

112. Xie, H., Ding, F. X., Schreiber, D., Eng, G., Liu, S. F., Arshava, B., Arevalo, E., Becker, J. M., and Naider, F. (2000) Synthesis and biophysical analysis of transmembrane domains of a Saccharomyces cerevisiae G protein-coupled receptor, *Biochemistry 39*, 15462-15474.

113. Arshava, B., Taran, I., Xie, H., Becker, J. M., and Naider, F. (2002) High resolution NMR analysis of the seven transmembrane domains of a heptahelical receptor in organic-aqueous medium, *Biopolymers 64*, 161-176.

114. Lazarova, T., Brewin, K. A., Stoeber, K., and Robinson, C. R. (2004) Characterization of peptides corresponding to the seven transmembrane domains of human adenosine A2a receptor, *Biochemistry 43*, 12945-12954.

115. Estephan, R., Englander, J., Arshava, B., Samples, K. L., Becker, J. M., and Naider, F. (2005) Biosynthesis and NMR analysis of a 73-residue domain of a Saccharomyces cerevisiae G protein-coupled receptor, *Biochemistry 44*, 11795-11810.

116. Zhao, J., Zheng, H., and Xie, X. Q. (2006) NMR characterization of recombinant transmembrane protein CB2 fragment CB2(180-233), *Protein and peptide letters 13*, 335-342.

117. Neumoin, A., Arshava, B., Becker, J., Zerbe, O., and Naider, F. (2007) NMR studies in dodecylphosphocholine of a fragment containing the seventh transmembrane helix of a G-protein-coupled receptor from Saccharomyces cerevisiae, *Biophys J 93*, 467-482.

118. Thevenin, D., and Lazarova, T. (2008) Stable interactions between the transmembrane domains of the adenosine A2A receptor, *Protein Sci 17*, 1188-1199.

119. Katragadda, M., Alderfer, J. L., and Yeagle, P. L. (2000) Solution structure of the loops of bacteriorhodopsin closely resembles the crystal structure, *Biochimica et biophysica acta 1466*, 1-6.

120. Therien, A. G., Glibowicka, M., and Deber, C. M. (2002) Expression and purification of two hydrophobic double-spanning membrane proteins derived from the cystic fibrosis transmembrane conductance regulator, *Protein expression and purification 25*, 81-86.

121. Kerman, A., and Ananthanarayanan, V. S. (2005) Expression and spectroscopic characterization of a large fragment of the mu-opioid receptor, *Biochimica et biophysica acta 1747*, 133-140.

122. Zheng, H., Zhao, J., Sheng, W., and Xie, X. Q. (2006) A transmembrane helix-bundle from G-protein coupled receptor CB2: biosynthesis, purification, and NMR characterization, *Biopolymers 83*, 46-61.

123. Martin, N. P., Leavitt, L. M., Sommers, C. M., and Dumont, M. E. (1999) Assembly of G protein-coupled receptors from fragments: identification of functional receptors with discontinuities in each of the loops connecting transmembrane segments, *Biochemistry 38*, 682-695.

124. Nannepaga, S. J., Gawalapu, R., Velasquez, D., and Renthal, R. (2004) Estimation of helix-helix association free energy from partial unfolding of bacterioopsin, *Biochemistry 43*, 550-559.

125. Overton, M. C., and Blumer, K. J. (2002) The extracellular N-terminal domain and transmembrane domains 1 and 2 mediate oligomerization of a yeast G protein-coupled receptor, *J Biol Chem 277*, 41463-41472.

126. Wakarchuk, W. W., Sung, W. L., Campbell, R. L., Cunningham, A., Watson, D. C., and Yaguchi, M. (1994) Thermostabilization of the Bacillus circulans xylanase by the introduction of disulfide bonds, *Protein engineering 7*, 1379-1386.

127. Ko, J. H., Jang, W. H., Kim, E. K., Lee, H. B., Park, K. D., Chung, J. H., and Yoo, O. J. (1996) Enhancement of thermostability and catalytic efficiency of AprP, an alkaline protease from Pseudomonas sp., by the introduction of a disulfide bond, *Biochemical and biophysical research communications 221*, 631-635.

128. Reading, N. S., and Aust, S. D. (2000) Engineering a disulfide bond in recombinant manganese peroxidase results in increased thermostability, *Biotechnology progress 16*, 326-333.

129. Martensson, L. G., Karlsson, M., and Carlsson, U. (2002) Dramatic stabilization of the native state of human carbonic anhydrase II by an engineered disulfide bond, *Biochemistry 41*, 15867-15875.

130. Hagihara, Y., Mine, S., and Uegaki, K. (2007) Stabilization of an immunoglobulin fold domain by an engineered disulfide bond at the buried hydrophobic region, *J Biol Chem 282*, 36489-36495.

131. Camarero, J. A., Fushman, D., Sato, S., Giriat, I., Cowburn, D., Raleigh, D. P., and Muir, T. W. (2001) Rescuing a destabilized protein fold through backbone cyclization, *Journal of molecular biology 308*, 1045-1062.

132. Hubbard, S., and Thornton, J. (1993) NACCESS, Department of Biochemistry and Molecular Biology, University College, London.

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

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

135. Kleid, D. G., Yansura, D., Small, B., Dowbenko, D., Moore, D. M., Grubman, M. J., McKercher, P. D., Morgan, D. O., Robertson, B. H., and Bachrach, H. L. (1981) Cloned viral protein vaccine for foot-and-mouth disease: responses in cattle and swine, *Science 214*, 1125-1129.

136. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes, *Methods in enzymology 185*, 60-89.

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

138. 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, *Journal of the American Chemical Society 126*, 2439-2446.

139. Zheng, H., Zhao, J., Wang, S., Lin, C. M., Chen, T., Jones, D. H., Ma, C., Opella, S., and Xie, X. Q. (2005) Biosynthesis and purification of a hydrophobic peptide from transmembrane domains of G-protein-coupled CB2 receptor, *J Pept Res 65*, 450-458.

140. Kieffer, C., Skalicky, J. J., Morita, E., De Domenico, I., Ward, D. M., Kaplan, J., and Sundquist, W. I. (2008) Two distinct modes of ESCRT-III recognition are required for VPS4 functions in lysosomal protein targeting and HIV-1 budding, *Developmental cell 15*, 62-73.

141. Diefenderfer, C., Lee, J., Mlyanarski, S., Guo, Y., and Glover, K. J. (2009) Reliable expression and purification of highly insoluble transmembrane domains, *Analytical biochemistry 384*, 274-278.

142. Cook, G. A., and Opella, S. J. (2009) NMR studies of p7 protein from hepatitis C virus, *Eur Biophys J 39(7)*, 1097-1104.

143. Cohen, L. S., Arshava, B., Estephan, R., Englander, J., Kim, H., Hauser, M., Zerbe, O., Ceruso, M., Becker, J. M., and Naider, F. (2008) Expression and biophysical analysis of two double-transmembrane domain-containing fragments from a yeast G protein-coupled receptor, *Biopolymers 90*, 117-130.

144. Liu, L. P., and Deber, C. M. (1998) Guidelines for membrane protein engineering derived from de novo designed model peptides, *Biopolymers 47*, 41-62.

145. Martin, N. P., Celic, A., and Dumont, M. E. (2002) Mutagenic mapping of helical structures in the transmembrane segments of the yeast alpha-factor receptor, *Journal of molecular biology 317*, 765-788.

146. Sharp, P. M., and Li, W. H. (1987) The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications, *Nucleic acids research 15*, 1281-1295.

147. Luo, S., Wehr, N. B., and Levine, R. L. (2006) Quantitation of protein on gels and blots by infrared fluorescence of Coomassie blue and Fast Green, *Analytical biochemistry 350*, 233-238.

148. Staley, J. P., and Kim, P. S. (1992) Complete folding of bovine pancreatic trypsin inhibitor with only a single disulfide bond, *Proceedings of the National Academy of Sciences of the United States of America 89*, 1519-1523.

149. Schagger, H., and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, *Analytical biochemistry 166*, 368-379.

150. Baldwin, J. M. (1993) The probable arrangement of the helices in G protein-coupled receptors, *The EMBO journal 12*, 1693-1703.

151. Konopka, J. B., Margarit, S. M., and Dube, P. (1996) Mutation of Pro-258 in transmembrane domain 6 constitutively activates the G protein-coupled alpha-factor receptor, *Proceedings of the National Academy of Sciences of the United States of America 93*, 6764-6769.

152. Dube, P., and Konopka, J. B. (1998) Identification of a polar region in transmembrane domain 6 that regulates the function of the G protein-coupled alpha-factor receptor, *Molecular and cellular biology 18*, 7205-7215.

153. Dube, P., DeCostanzo, A., and Konopka, J. B. (2000) Interaction between transmembrane domains five and six of the alpha -factor receptor, *J Biol Chem 275*, 26492-26499.

154. Kido, M., Yamanaka, K., Mitani, T., Niki, H., Ogura, T., and Hiraga, S. (1996) RNase E polypeptides lacking a carboxyl-terminal half suppress a mukB mutation in Escherichia coli, *Journal of bacteriology 178*, 3917-3925.

155. Grunberg-Manago, M. (1999) Messenger RNA stability and its role in control of gene expression in bacteria and phages, *Annual review of genetics 33*, 193-227.

156. Lopez, P. J., Marchand, I., Joyce, S. A., and Dreyfus, M. (1999) The C-terminal half of RNase E, which organizes the Escherichia coli degradosome, participates in mRNA degradation but not rRNA processing in vivo, *Molecular microbiology 33*, 188-199.

157. Lee, N. (1980) Molecular Aspects of *ara* Regulation. , In *The Operon* (Miller, J. H., and Reznikoff, W. S., Eds.), pp 389-410, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

158. Lee, N., Francklyn, C., and Hamilton, E. P. (1987) Arabinose-induced binding of AraC protein to araI2 activates the araBAD operon promoter, *Proceedings of the National Academy of Sciences of the United States of America 84*, 8814-8818.

159. Grodberg, J., and Dunn, J. J. (1988) ompT encodes the Escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification, *Journal of bacteriology 170*, 1245-1253.

160. Kelly, S. M., and Price, N. C. (2000) The use of circular dichroism in the investigation of protein structure and function, *Current protein & peptide science 1*, 349-384.

161. Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to study proteins by circular dichroism, *Biochimica et biophysica acta 1751*, 119-139.

162. Morrisett, J. D., David, J. S., Pownall, H. J., and Gotto, A. M., Jr. (1973) Interaction of an apolipoprotein (apoLP-alanine) with phosphatidylcholine, *Biochemistry 12*, 1290-1299.

163. Pelton, J. T., and McLean, L. R. (2000) Spectroscopic methods for analysis of protein secondary structure, *Analytical biochemistry 277*, 167-176.

164. Ciaccio, N. A., Moreno, M. L., Bauer, R. L., and Laurence, J. S. (2008) High-yield expression in E. coli and refolding of the bZIP domain of activating transcription factor 5, *Protein expression and purification 62*, 235-243.

165. Training, T. P. i. S. B. a. d. E. i. R. a. Biophysics Core Handbook, (The University of Colorado Health Science Center, D., Ed.).

166. Williams, K. A., Farrow, N. A., Deber, C. M., and Kay, L. E. (1996) Structure and dynamics of bacteriophage IKe major coat protein in MPG micelles by solution NMR, *Biochemistry 35*, 5145-5157.

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

168. Zou, C., Naider, F., and Zerbe, O. (2008) Biosynthesis and NMR-studies of a double transmembrane domain from the Y4 receptor, a human GPCR, *Journal of biomolecular NMR 42*, 257-269.

169. Rost, B., Sander, C., and Schneider, R. (1994) PHD--an automatic mail server for protein secondary structure prediction, *Comput Appl Biosci 10*, 53-60.

170. Snider, C., Jayasinghe, S., Hristova, K., and White, S. H. (2009) MPEx: a tool for exploring membrane proteins, *Protein Sci 18*, 2624-2628.

171. Naider, F., Ding, F. X., VerBerkmoes, N. C., Arshava, B., and Becker, J. M. (2003) Synthesis and biophysical characterization of a multidomain peptide from a Saccharomyces cerevisiae G protein-coupled receptor, *J Biol Chem 278*, 52537-52545.

172. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) 1H, 13C and 15N chemical shift referencing in biomolecular NMR, *Journal of biomolecular NMR 6*, 135-140.

173. Therien, A. G., Grant, F. E., and Deber, C. M. (2001) Interhelical hydrogen bonds in the CFTR membrane domain, *Nature structural biology 8*, 597-601.

174. Melnyk, R. A., Kim, S., Curran, A. R., Engelman, D. M., Bowie, J. U., and Deber, C. M. (2004) The affinity of GXXXG motifs in transmembrane helix-helix interactions is modulated by long-range communication, *J Biol Chem 279*, 16591-16597.

175. Choi, M. Y., Cardarelli, L., Therien, A. G., and Deber, C. M. (2004) Non-native interhelical hydrogen bonds in the cystic fibrosis transmembrane conductance regulator domain modulated by polar mutations, *Biochemistry 43*, 8077-8083.

176. Rath, A., Johnson, R. M., and Deber, C. M. (2007) Peptides as transmembrane segments: decrypting the determinants for helix-helix interactions in membrane proteins, *Biopolymers 88*, 217-232.

177. Neville, D. M., Jr. (1971) Molecular weight determination of protein-dodecyl sulfate complexes by gel electrophoresis in a discontinuous buffer system, *J Biol Chem 246*, 6328-6334.

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

179. Buck, M. (1998) Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins, *Quarterly reviews of biophysics 31*, 297-355.

180. Nelson, J. W., and Kallenbach, N. R. (1989) Persistence of the alpha-helix stop signal in the S-peptide in trifluoroethanol solutions, *Biochemistry 28*, 5256-5261.

181. Bruch, M. D., and Gierasch, L. M. (1990) Comparison of helix stability in wild-type and mutant LamB signal sequences, *J Biol Chem 265*, 3851-3858.

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

183. Lipfert, J., Columbus, L., Chu, V. B., Lesley, S. A., and Doniach, S. (2007) Size and shape of detergent micelles determined by small-angle X-ray scattering, *The journal of physical chemistry 111*, 12427-12438.

184. Gobl, C., Dulle, M., Hohlweg, W., Grossauer, J., Falsone, S. F., Glatter, O., and Zangger, K. Influence of phosphocholine alkyl chain length on peptide-micelle interactions and micellar size and shape, *The journal of physical chemistry 114*, 4717-4724.

185. Chattopadhyay, A., and Harikumar, K. G. (1996) Dependence of critical micelle concentration of a zwitterionic detergent on ionic strength: implications in receptor solubilization, *FEBS Lett 391*, 199-202.

186. Gluck, J. M., Wittlich, M., Feuerstein, S., Hoffmann, S., Willbold, D., and Koenig, B. W. (2009) Integral membrane proteins in nanodiscs can be studied by solution NMR spectroscopy, *Journal of the American Chemical Society 131*, 12060-12061.

187. Longley, R. P., Rose, A. H., and Knights, B. A. (1968) Composition of the protoplast membrane from Saccharomyces cerevisiae, *The Biochemical journal 108*, 401-412.

188. Ejsing, C. S., Sampaio, J. L., Surendranath, V., Duchoslav, E., Ekroos, K., Klemm, R. W., Simons, K., and Shevchenko, A. (2009) Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry, *Proceedings of the National Academy of Sciences of the United States of America 106*, 2136-2141.

189. Mendz, G. L., Brown, L. R., and Martenson, R. E. (1990) Interactions of myelin basic protein with mixed dodecylphosphocholine/palmitoyllysophosphatidic acid micelles, *Biochemistry 29*, 2304-2311.

190. Vinogradova, O., Badola, P., Czerski, L., Sonnichsen, F. D., and Sanders, C. R., 2nd. (1997) Escherichia coli diacylglycerol kinase: a case study in the application of solution NMR methods to an integral membrane protein, *Biophys J 72*, 2688-2701.

191. Wienk, H. L., Wechselberger, R. W., Czisch, M., and de Kruijff, B. (2000) Structure, dynamics, and insertion of a chloroplast targeting peptide in mixed micelles, *Biochemistry 39*, 8219-8227.

192. Columbus, L., Lipfert, J., Jambunathan, K., Fox, D. A., Sim, A. Y., Doniach, S., and Lesley, S. A. (2009) Mixing and matching detergents for membrane protein NMR structure determination, *Journal of the American Chemical Society 131*, 7320-7326.

193. Hilty, C., Wider, G., Fernandez, C., and Wuthrich, K. (2004) Membrane protein-lipid interactions in mixed micelles studied by NMR spectroscopy with the use of paramagnetic reagents, *Chembiochem 5*, 467-473.

194. Pastore, A., Temussi, P. A., Salvadori, S., Tomatis, R., and Mascagni, P. (1985) A conformational study of the opioid peptide dermorphin by one-dimensional and two-dimensional nuclear magnetic resonance spectroscopy, *Biophys J 48*, 195-200.

195. Picone, D., D'Ursi, A., Motta, A., Tancredi, T., and Temussi, P. A. (1990) Conformational preferences of [Leu5]enkephalin in biomimetic media. Investigation by 1H NMR, *European journal of biochemistry / FEBS 192*, 433-439.

196. Trogen, G. B., Edlund, U., Larsson, G., and Sethson, I. (1998) The solution NMR structure of a blue-green algae hepatotoxin, microcystin-RR--a comparison with the structure of microcystin-LR, *European journal of biochemistry / FEBS 258*, 301-312.

197. D'Ursi, A., Albrizio, S., Fattorusso, C., Lavecchia, A., Zanotti, G., and Temussi, P. A. (1999) Solution conformation of a potent cyclic analogue of tuftsin: low- temperature nuclear magnetic resonance study in a cryoprotective mixture, *Journal of medicinal chemistry 42*, 1705-1713.

198. Spadaccini, R., Crescenzi, O., Picone, D., Tancredi, T., and Temussi, P. A. (1999) Solution structure of dynorphin A (1-17): a NMR study in a cryoprotective solvent mixture at 278 K, *J Pept Sci 5*, 306-312.

199. Amodeo, P., Naider, F., Picone, D., Tancredi, T., and Temussi, P. A. (1998) Conformational sampling of bioactive conformers: a low-temperature NMR study of 15N-Leu-enkephalin, *J Pept Sci 4*, 253-265.

200. Albrizio, S., Carotenuto, A., Fattorusso, C., Moroder, L., Picone, D., Temussi, P. A., and D'Ursi, A. (2002) Environmental mimic of receptor interaction: conformational analysis of CCK-15 in solution, *Journal of medicinal chemistry 45*, 762-769.

201. Katragadda, M., Maciejewski, M. W., and Yeagle, P. L. (2004) Structural studies of the putative helix 8 in the human beta(2) adrenergic receptor: an NMR study, *Biochimica et biophysica acta 1663*, 74-81.

202. Yeagle, P. L., Bennett, M., Lemaitre, V., and Watts, A. (2007) Transmembrane helices of membrane proteins may flex to satisfy hydrophobic mismatch, *Biochimica et biophysica acta 1768*, 530-537.

203. Hirota-Nakaoka, N., Hasegawa, K., Naiki, H., and Goto, Y. (2003) Dissolution of beta2-microglobulin amyloid fibrils by dimethylsulfoxide, *Journal of biochemistry 134*, 159-164.

204. Wider, G. (2005) NMR techniques used with very large biological macromolecules in solution, *Methods in enzymology 394*, 382-398.

205. Kobilka, B. K., Kobilka, T. S., Daniel, K., Regan, J. W., Caron, M. G., and Lefkowitz, R. J. (1988) Chimeric alpha 2-,beta 2-adrenergic receptors: delineation of domains involved in effector coupling and ligand binding specificity, *Science 240*, 1310-1316.

206. Schoneberg, T., Liu, J., and Wess, J. (1995) Plasma membrane localization and functional rescue of truncated forms of a G protein-coupled receptor, *J Biol Chem 270*, 18000-18006.

207. Ridge, K. D., Lee, S. S., and Abdulaev, N. G. (1996) Examining rhodopsin folding and assembly through expression of polypeptide fragments, *J Biol Chem 271*, 7860-7867.

208. Arshava, B., Liu, S. F., Jiang, H., Breslav, M., Becker, J. M., and Naider, F. (1998) Structure of segments of a G protein-coupled receptor: CD and NMR analysis of the Saccharomyces cerevisiae tridecapeptide pheromone receptor, *Biopolymers 46*, 343-357.

209. Goodman, M., Naider, F., and Rupp, R. (1971) Conformations of Alanine Oligopeptides in Solution, *Bioorganic Chemistry 1*, 310-328.

210. Naider, F., and Goodman, M. (1977) Conformational Analysis of Oligopeptides by Spectral Techniques, In *Bioorganic Chemistry III, Macro and Multimolecular Systems* (Tamelen, E. E. V., Ed.), p 177, Academic Press, Inc.

211. White, S. H., Ladokhin, A. S., Jayasinghe, S., and Hristova, K. (2001) How membranes shape protein structure, *J Biol Chem 276*, 32395-32398.

212. Ma, D., Liu, Z., Li, L., Tang, P., and Xu, Y. (2005) Structure and dynamics of the second and third transmembrane domains of human glycine receptor, *Biochemistry 44*, 8790-8800.

213. 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.

214. Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *Journal of the American Chemical Society 114*, 10663-10665.

215. Ikura, M., Kay, L. E., and Bax, A. (1990) A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin, *Biochemistry 29*, 4659-4667.

216. Grzesiek, S., and Bax, A. (1992) Improved 3D Triple-Resonance NMR Techniques Applied to a 31-kDa Protein., *J Magn Reson 96*, 432-440.

217. Muhandiram, D., and Kay, L. E. (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity., *J Magn Reson B 103*, 203-216.

218. Kay, L. E., Xu, G. Y., and Yamazaki, T. (1994) Enhanced-sensitivity triple-resonance spectroscopy with minimal H2O saturation, *J Magn Reson A 109*, 129-133.

219. Wittekind, M., and Muller, L. (1993) HNCACB, a high sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonanaces with alpha- and beta-carbon resonances in proteins, *J Magn Reson B 101*, 201-205.

220. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR 4*, 845-858.

221. Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) H-1-H-1 correlation via isotropic mixing of C-13 magnetization, a new 3-dimensional approach for assigning H-1 and C-13 spectra of C-13-enriched proteins, *J Magn Reson 88*, 425-431.

222. Vuister, G. W., and Bax, A. (1992) Resolution enhancement and spectral editing of uniformly 13C-enriched proteins by homonuclear broadband 13C decoupling, *Journal of Magnetic Resonance 98*, 428-435.

223. Thrippleton, M. J., and Keeler, J. (2003) Elimination of Zero-Quantum Interference in Two-Dimensional NMR Spectra *Angewandte Chemie International Edition 42*, 3938-3941.

224. Levitt, M. H., Freeman, R., and Frenkiel, T. (1982) Broadband heteronuclear decoupling *Journal of Magnetic Resonance 47*, 328-330.

225. Bax, A., and Davis, D. G. (1985) MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy, *Journal of Magnetic Resonance 65*, 355-360.

226. Palmer III, A. G., Cavanagh, J., Wright, P. E., and Rance, M. (1991) Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy, *Journal of Magnetic Resonance 93*, 151-170.

227. Davis, A. L., Keeler, J., Laue, E. D., and Moskau, D. (1992) Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients, *Journal of Magnetic Resonance 98*, 207-216.

228. Diercks, T., Coles, M., and Kessler, H. (1999) An efficient strategy for assignment of cross-peaks in 3D heteronuclear NOESY experiments *Journal of biomolecular NMR 15*, 177-180.

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

230. Johnson, B. A., and Blevins, R. A. (1994) NMR View: A computer program for the visualization and analysis of NMR data, *Journal of biomolecular NMR 4*, 603-614.

231. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *Journal of molecular graphics 14*, 51-55, 29-32.

232. Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., Rance, M., and Skelton, N. J. (2007) *Protein NMR Spectroscopy: Principles and Practice*, Second Edition ed., Elsevier, Inc.

233. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins, *Journal of Magnetic Resonance 89*, 496-514.

234. Clubb, R. T., Thanabal, V., and Wagner, G. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1HN, 15N, and 13C′ chemical shifts in 15N---13C-labelled proteins, *Journal of Magnetic Resonance 97*, 213-217.

235. Liu, L. P., and Deber, C. M. (1998) Uncoupling hydrophobicity and helicity in transmembrane segments. Alpha-helical propensities of the amino acids in non-polar environments, *J Biol Chem 273*, 23645-23648.

236. Dutta, K., Cox, C. J., Alexandrov, A., Huang, H., Basavappa, R., and Pascal, S. M. (2002) Letter to the Editor: Sequence-specific chemical shift assignment and chemical shift indexing of murine apo-Mts1, *Journal of biomolecular NMR 22*, 181-182.

237. Jones, C. T., Ma, L., Burgner, J. W., Groesch, T. D., Post, C. B., and Kuhn, R. J. (2003) Flavivirus capsid is a dimeric alpha-helical protein, *Journal of virology 77*, 7143-7149.

238. Burton, R. A., Tsurupa, G., Medved, L., and Tjandra, N. (2006) Identification of an ordered compact structure within the recombinant bovine fibrinogen alphaC-domain fragment by NMR, *Biochemistry 45*, 2257-2266.

239. Ahmed, M. A., Bamm, V. V., Harauz, G., and Ladizhansky, V. (2007) The BG21 isoform of Golli myelin basic protein is intrinsically disordered with a highly flexible amino-terminal domain, *Biochemistry 46*, 9700-9712.

240. Libich, D. S., and Harauz, G. (2008) Backbone dynamics of the 18.5 kDa isoform of myelin basic protein reveals transient alpha-helices and a calmodulin-binding site, *Biophys J 94*, 4847-4866.

241. Pena, P. V., Musselman, C. A., Kuo, A. J., Gozani, O., and Kutateladze, T. G. (2009) NMR assignments and histone specificity of the ING2 PHD finger, *Magn Reson Chem 47*, 352-358.

242. Baker, K. A., Hilty, C., Peti, W., Prince, A., Pfaffinger, P. J., Wider, G., Wuthrich, K., and Choe, S. (2006) NMR-derived dynamic aspects of N-type inactivation of a Kv channel suggest a transient interaction with the T1 domain, *Biochemistry 45*, 1663-1672.

243. Yagi, H., Kajiwara, N., Iwabuchi, T., Izumi, K., Yoshida, M., and Akutsu, H. (2009) Stepwise propagation of the ATP-induced conformational change of the F1-ATPase beta subunit revealed by NMR, *J Biol Chem 284*, 2374-2382.

244. Teng, Q. (2005) *Structural Biology: Practical NMR Applications*, Springer, New York.

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

246. Santoro, J., Gonzalez, C., Bruix, M., Neira, J. L., Nieto, J. L., Herranz, J., and Rico, M. (1993) High-resolution three-dimensional structure of ribonuclease A in solution by nuclear magnetic resonance spectroscopy, *Journal of molecular biology 229*, 722-734.

247. Rodriguez, J. C., Wilks, A., and Rivera, M. (2006) Backbone NMR assignments and H/D exchange studies on the ferric azide- and cyanide-inhibited forms of Pseudomonas aeruginosa heme oxygenase, *Biochemistry 45*, 4578-4592.

248. Werner, K., Lehner, I., Dhiman, H. K., Richter, C., Glaubitz, C., Schwalbe, H., Klein-Seetharaman, J., and Khorana, H. G. (2007) Combined solid state and solution NMR studies of alpha,epsilon-15N labeled bovine rhodopsin, *Journal of biomolecular NMR 37*, 303-312.

249. Baxter, N. J., and Williamson, M. P. (1997) Temperature dependence of 1H chemical shifts in proteins, *Journal of biomolecular NMR 9*, 359-369.

250. Andersen, N. H., Neidigh, J. W., Harris, S. M., Lee, G. M., Liu, Z., and Tong, H. (1997) Extracting Information from the Temperature Gradients of Polypeptide NH Chemical Shifts. 1. The Importance of Conformational Averaging, *Journal of the American Chemical Society 119*, 8547-8561.

251. Stanley, A. M., and Fleming, K. G. (2008) The process of folding proteins into membranes: challenges and progress, *Archives of biochemistry and biophysics 469*, 46-66.

252. Teilum, K., Kragelund, B. B., and Poulsen, F. M. (2002) Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling, *Journal of molecular biology 324*, 349-357.

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

254. 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.

255. Hildebrand, P. W., Preissner, R., and Frommel, C. (2004) Structural features of transmembrane helices, *FEBS Lett 559*, 145-151.

256. Mor, A., Segal, E., Mester, B., Arshava, B., Rosen, O., Ding, F. X., Russo, J., Dafni, A., Schvartzman, F., Scherf, T., Naider, F., and Anglister, J. (2009) Mimicking the structure of the V3 epitope bound to HIV-1 neutralizing antibodies, *Biochemistry 48*, 3288-3303.

257. Moseri, A., Tantry, S., Sagi, Y., Arshava, B., Naider, F., and Anglister, J. An optimally constrained V3 peptide is a better immunogen than its linear homolog or HIV-1 gp120, *Virology 401*, 293-304.

258. Kumar, G. S., Ramasamy, P., Sikdar, S. K., and Sarma, S. P. (2005) Overexpression, purification, and pharmacological activity of a biosynthetically derived conopeptide, *Biochemical and biophysical research communications 335*, 965-972.

259. Hager, D. A., and Burgess, R. R. (1980) Elution of proteins from sodium dodecyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity: results with sigma subunit of Escherichia coli RNA polymerase, wheat germ DNA topoisomerase, and other enzymes, *Analytical biochemistry 109*, 76-86.

260. Luy, B., Kobzar, K., and Kessler, H. (2004) An easy and scalable method for the partial alignment of organic molecules for measuring residual dipolar couplings, *Angewandte Chemie (International ed) 43*, 1092-1094.

261. Freudenberger, J. C., Spiteller, P., Bauer, R., Kessler, H., and Luy, B. (2004) Stretched poly(dimethylsiloxane) gels as NMR alignment media for apolar and weakly polar organic solvents: an ideal tool for measuring RDCs at low molecular concentrations, *Journal of the American Chemical Society 126*, 14690-14691.

262. Luy, B., Kobzar, K., Knor, S., Furrer, J., Heckmann, D., and Kessler, H. (2005) Orientational properties of stretched polystyrene gels in organic solvents and the suppression of their residual 1H NMR signals, *Journal of the American Chemical Society 127*, 6459-6465.

263. Kobzar, K., Kessler, H., and Luy, B. (2005) Stretched gelatin gels as chiral alignment media for the discrimination of enantiomers by NMR spectroscopy, *Angewandte Chemie (International ed) 44*, 3145-3147.

264. Freudenberger, J. C., Knor, S., Kobzar, K., Heckmann, D., Paululat, T., Kessler, H., and Luy, B. (2005) Stretched poly(vinyl acetate) gels as NMR alignment media for the measurement of residual dipolar couplings in polar organic solvents, *Angewandte Chemie (International ed) 44*, 423-426.

265. Lee, B. K., Jung, K. S., Son, C., Kim, H., VerBerkmoes, N. C., Arshava, B., Naider, F., and Becker, J. M. (2007) Affinity purification and characterization of a G-protein coupled receptor, Saccharomyces cerevisiae Ste2p, *Protein expression and purification 56*, 62-71.