Expression, Purification, Circular Dichroism, and NMR Analysis of Triple Transmembrane Domain Containing Fragments of a GPCR

By

Katrina E. Caroccia

Doctoral Graduate Program in Biochemistry
The Graduate Center, City University of New York
Laboratory of Fred Naider
Department of Chemistry
College of Staten Island, City University of New York

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

2013

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. Jeffrey M. Becker
Dr. Robert Bittman
Dr. Sebastien Poget
Dr. Ruth E. Stark
Supervisory Committee

The City University of New York

ABSTRACT

Expression, Purification, Circular Dichroism, and NMR Analysis of Triple Transmembrane Domain Containing Fragments of a GPCR

By: Katrina E. Caroccia

Mentor: Fred Naider

G protein-coupled receptors (GPCRs) are the largest class of signaling molecules in eukaryotes and are important pharmacological targets. Structural characterization of GPCRs is of paramount importance to the discovery of more efficient drugs; however, these studies are hindered by the inherent hydrophobicity, flexibility, and large size of these signaling proteins. Since their flexibility makes crystallization difficult, stabilizing mutations or substitutions are required to facilitate crystal-packing contacts. The size of the receptor/membrane mimetic complex required for solution-state nuclear magnetic resonance (NMR) analysis is too large to enable efficient isotropic tumbling. For these reasons, high-resolution structural information is available for only thirteen of the ~1000 GPCRs identified to date.

Insights into conformational preferences and the three-dimensional (3D) structure of domains of these receptors can be obtained using polypeptide fragments of these proteins. This approach is relevant because functional GPCRs can be reconstituted from fragments. Our lab has been using the yeast α-factor receptor as a paradigm for methods development for GPCR structural characterization. We have studied isolated fragments of this GPCR containing 1 and 2 transmembrane domains (TMs). For my dissertation project, I set out to characterize 3TM containing fragments of Ste2p. The goals of my project were 1) to determine whether 3TM containing fragments form more definied tertiary structures than a 2TM containing counterpart and 2) to determine whether

chemical shifts from smaller fragments could be used to assign larger fragments.

Two 3TM-containing fragments of Ste2p were recombinantly expressed, purified by reverse phase high performance liquid chromatography (RP-HPLC), and subjected to extensive biophysical analysis by circular dichroism (CD) and NMR spectroscopy. A 131-residue fragment containing the first 3TMs of Ste2p, TM1-TM3 (G31-R161) and a 151-residue fragment containing the first, second, and seventh TMs, TM127 (G31-T114,T274-L340) were cloned and expressed as TrpΔLE fusion proteins in *Escherichia coli*. The expressed proteins were subjected to CNBr cleavage to remove the fusion tag and TM1-TM3 and TM127 were purified by reverse-phase HPLC. The cleavage products were isolated in yields of up to 20 mg per liter of culture in a variety of isotopic-labeled forms. The secondary structure of TM1-TM3 and TM127 was determined to be helical in a number of membrane mimetic environments, including 2,2,2-trifluoroethanol (TFE):water and detergent micelles by CD spectroscopy. Preliminary HSQC analysis in 50% TFE:water and detergent micelles revealed that these fragments were suitable for structural analysis by NMR spectroscopy. Complete backbone and side chain assignments and a detailed localization of the secondary structural elements of TM1-TM3 in 50% TFE:water have been achieved. Under these conditions, an NMR structure was determined to have low convergence, and no tertiary contacts were observed. Attempts are currently being made to ascertain tertiary contacts by paramagnetic relaxation enhancement (PRE). NMR structural characterization of both TM1-TM3 and TM127 in detergent micelles is being conducted at the University of Zurich.

To determine the transferability of chemical shifts for small to large fragments of Ste2p, two smaller constructs were analyzed. A 1TM containing fragment of Ste2p, TM1

[Ste2p(G31-T78)] was cloned into a direct expression vector and expressed with an N-terminal histidine tag in *E.coli*. The fragment was purified by RP-HPLC. This construct and a previously expressed TM7CT40 construct [Ste2p(S267-S339)](*1*) were subject to extensive heteronuclear NMR analysis in lysopalmitoylphosphatidylglycerol:dodecylphosphocholine (LPPG:DPC) micelles for backbone and side chain assignment. Backbone and side chain assignments for the TM1 fragment have been completed. Comparison of amide chemical shifts obtained for this fragment to those obtained for the TM1-TM2, TM1-TM3, and TM127 fragments of Ste2p suggests that chemical shifts are transferable for all regions except the flexible GXXXG region. A calculated NMR structure for TM1 reveals a large kink at this region. Comparison of the amide TM7 chemical shifts to those of the TM127 fragment reveals that all obtained chemical shifts are similar for this region. Detailed computational analysis is currently being performed at the Frankfurt Institute of Advanced Studies of Goethe University in order to assess complete chemical shift transferability for these fragments.

To my parents, John and Jacqueline Caroccia:

Your constant love and support have made me the woman I am today. Thank you for setting such a wonderful example of hard work and dedication and for always emphasizing the importance of education.

To my husband, Aldo Fracchiolla:

Thanks for coming along with me on this scientific journey. Most of all, thanks for always believing in me.

ACKNOWLEDGMENTS

Completion of this dissertation would not have been possible without the support and guidance of an entire team of outstanding scientists. First and foremost, I would like to thank my mentor, Dr. Fred Naider. You have instilled in me a deep love for science and have opened doors for me that I didn't know existed. Thank you for always pushing me to reach my potential. You have changed my life and you will always be the best teacher I have ever had. I would like to thank the entire Naider lab, especially Dr. Boris Arshava and Dr. Leah Cohen. Boris, thank you for turning me into a spectroscopist and for always reminding me to "slow down." Leah, thank you for turning me into a molecular biologist and for being such a special friend. Thank you to Dr. Oliver Zerbe and the entire Zerbe lab at the University of Zurich for welcoming me into their lab, teaching me the latest and greatest NMR techniques, and helping me sort through some of the problems I encountered toward the end of this project. A special thanks to Martin Poms for being my Swiss counterpart and helping me put the pieces of the 3TM puzzle together. Last, but certainly not least, I would like to thank the members of my thesis committee, Dr. Jeffrey Becker, Dr. Robert Bittman, Dr. Sebastien Poget, and Dr. Ruth Stark. Thank you for contributing your time and expertise. All of your advice is greatly appreciated.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

°C	Degrees Celcius
3D	Three-dimensional
Ac-MTSL	An acetylated version of *S*-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl) methylmethane sulfonothioate
Amp	Ampicillin
*Bam*HI	*Bacillus amyloliquefaciens* H restriction enzyme I
cAMP	Cyclic Adenosine Monophosphate
CD	Circular Dichroism
CNBr	Cyanogen Bromide
CT	C-terminus
DDM	Dodecyl maltoside
DHPC	1,2-dihexanoyl-*sn*-glycero-3-phosphocholine
DPC	Dodecylphosphocholine
*Dpn*I	*Diplococcus pneumoniae* restriction enzyme I
DSS	2,2-dimethyl-2-silapentane-5-sulfonate sodium salt
EDTA	Ethylenediaminetetraacetic acid
EL	Extracellular Loops
ESI-MS	Electrospray Ionization Mass Spectroscopy
f_H	Fraction Helicity
FP	Fusion Protein
GDP	Guanosine Diphosphate
GPCRs	G Protein-Coupled Receptors
GTP	Guanosine Triphosphate
H-D	Hydrogen-Deuterium

HDPC	Hexadecylphosphocholine
HEPES	4-(2-Hydroxyethyl)-1-piperazine ethanesulfonic acid
*Hind*III	*Haemophilus influenzae* Rd restriction enzyme III
HSQC	Heteronuclear Single Quantum Coherence
IB	Inclusion Bodies
IL	Intracellular Loop
IMP	Integral Membrane Protein
IPTG	Isopropyl β-D1-thiogalactopyranoside
LB	Luria Broth
LMPG	Lysomyristoylphosphatidylglycerol
LPPC	Lysopalmitoylphosphatidylcholine
LPPG	Lysopalmitoylphosphatidylglycerol
LSPG	Lysostearoylphosphatidylglycerol
MAP	Mitogen Activated Protein
MATa	Mating type a
MATα	Mating type α
MOPS	3-(N-morpholino)propanesulfonic acid
MS	Mass Spectrometry
MTSL	*S*-(2,2,5,5-Tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methylmethanesulfonothioate
*Nde*I	*Neisseria denitrificans* restriction enzyme I
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NT	N-terminus
OD_{600}	Optical Density at 600 nm
PCR	Polymerase Chain Reaction

PFO	Perfluorooctanoate
PMSF	Phenylmethanesulfonylfluoride
PRE	Paramagnetic Relaxation Enhancement
RP-HPLC	Reverse Phase High Performance Liquid Chromatography
SDS	Sodium dodecylsulfate
SDS-PAGE	Sodium dodecylsulfate Polyacrylamide Gel Electrophoresis
TALOS	Torsion Angle Likelihood Obtained from Shift and sequence similarity
TFE	2,2,2-Trifluoroethanol
TFE:water	Trifluoroethanol:water mixture
TM	Transmembrane Domain
TM1	Ste2p fragment consisting of 30 residues of the NT and the first TM domain (G31-T78)
TM1-TM3	Ste2p fragment consisting of 30 residues of the NT and the first through the third TMs (G31-R161)
TM127	Ste2p fragment consisting of 30 residues of the NT and the first through the second TMs, an extracellular loop containing eleven residues of EL1 and two residues of EL3, the seventh TM and 40 residues of the CT (G31-T114,T274-L340)
TM7CT40	Ste2p fragment consisting of the third EL, the seventh TM, and 40 residues of the CT
TrpΔLE	Portion of the histidine-tagged TrpΔLE1413

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Background and Significance

G Protein-coupled Receptors

G protein-coupled receptors (GPCRs) are a class of proteins involved in cellular signaling cascades. They are among the most abundant proteins in the human genome, with approximately 1000 GPCRs identified to date. GPCRs are responsive to a variety of ligands including small molecules, peptides, and proteins and have been implicated in signaling pathways ranging from vision and light detection to heart and lung function. They were the subject of the 2012 Nobel Prize in Chemistry.

Despite the lack of sequence homology among GPCRs, hydropathy analysis shows that these proteins have a common structural motif. They are composed of seven transmembrane (TM) domains connected by intra- and extracellular loops (IL and EL, respectively). The N-terminus (NT) of the receptors is always found on the extracellular side of the membrane, while the C-terminus (CT) is found on the cytoplasmic side. The N-terminus and ELs have been implicated in ligand binding. The C-terminus contains an eighth α-helix, is known to interact with heterotrimeric G-proteins, and is also involved in receptor desensitization.

GPCRs have been categorized into five distinct classes based on specific residue conservation, consensus motifs, and the size of the N-terminus of the protein (*2*, *3*). The three major classes are A, B, and C, where class A represents the rhodopsin-like receptors, class B represents the glucagon/secretin-like receptors, and class C represents the metabolotropic receptors. Class D receptors are the fungal pheromone receptors. Class E is comprised of the cAMP receptors from *Dictyostelium discoidium*. Class A

receptors are the most studied class of GPCRs and are characterized by a conserved DRY sequence in TM3, an NPxxY motif in TM7, a disulfide bond, and several conserved, functionally relevant proline residues. Class B are characterized by N-termini of ~100 residues that are enriched in Cys. Class C receptors are characterized by enormous, Venus fly trap-like (500-600 residue) N-termini. The ligand binding sites for class B and C receptors are located in extracellular domains of these receptors (*4*).

Cellular responses to extracellular stimulation of a GPCR signaling pathway are facilitated by the associated heterotrimeric G protein. G proteins are so named due to their interactions with guanine nucleotides, namely GDP and GTP. A G protein is in the inactive state when GDP is bound and is activated upon exchange of bound GDP for GTP. Upon ligand binding to a GPCR, a conformational change occurs in the receptor such that the interaction with the G protein changes. This conformational change causes the dissociation of the βγ subunits of the heterotrimer and the exchange of GDP for GTP on the α subunit, thus activating the G protein and creating a signal cascade. Activated G proteins generally stimulate the production or release of second messengers such as cAMP and calcium. These second messengers propagate the signal, eliciting a cellular response. The signaling cascade is then turned off by the hydrolysis of GTP by the α subunit itself and the reformation of the αβγ heterotrimer (*5*). Signaling is also regulated by GPCR internalization, where the C-terminal tail becomes phosphorylated and the entire receptor is removed from the plasma membrane.

As these receptors have been implicated in a number of important physiological processes, GPCRs are currently major pharmacological targets. Drugs targeting GPCRs currently represent 40-60% of all marketed pharmaceuticals (*6-8*). As a result, structural

characterization of these receptors is extremely important for drug development. High-resolution structures of GPCRs can provide invaluable information regarding the ligand-binding sites of these receptors. Ideally, detailed characterization of binding sites can lead to the development of drugs that will bind specifically to just one receptor. This will lead to the development of drugs with fewer side effects.

Structural Characterization of GPCRs

As previously mentioned, structural information regarding GPCRs is extremely valuable because they are major pharmacological targets, but characterization is difficult due to their hydrophobicity, inherent flexibility, and large size. Despite the large number of identified GPCRs, at the time of writing this dissertation, high-resolution crystallographic structural information was only available for thirteen of these receptors, including rhodopsin/opsin, the β_1 and β_2-adrenergic receptors, the adenosine A_{2A} receptor, the M2 and M3 muscarinic acetylcholine receptors, the CXCR4 chemokine receptor, the dopamine D3 receptor, the histamine receptor, the μ, κ, and δ opioid receptors, the nociceptin receptor, and the sphingosine 1-phosphate type 1 (S1P1) receptor (*9-29*). These structures represent only Class A GPCRs, and the majority of these receptors bind small molecules. Regardless of the functional similarities between these proteins, the crystal structures revealed distinct differences. Slight differences in the orientation of the TM helices among the structures were observed. Most strikingly, differences in the extracellular portions of the receptors were observed, and these differences had functional implications to ligand binding. Structures for receptors in the active and inactive states also revealed subtle differences with functional implications

(*30*). The observance of functionally relevant differences demonstrates the need for further structural characterization, especially for non-Class A receptors.

Techniques for Structure Determination: Benefits and Caveats

Undoubtedly, crystallography provides invaluable, precise structural information on GPCRs. Nevertheless, X-ray structures represent a locked state of the protein and therefore provide little insight into its dynamics. In addition, this method often provides sparse information about disordered regions of receptors, especially those of long loops and tails. Furthermore, stabilizing modifications required to aid in crystallization can possibly compromise the biological relevance of the system being studied. All of the crystal structures solved to date required thermostabilizing mutations (*15, 18, 19*), antibody stabilization of an IL(*11*), or the replacement of an IL with T4 lysozyme (*12, 16, 17, 29*) or an apocytochrome (*28*) in order to facilitate crystal formation. These modifications, in most cases, lead to receptors that are no longer able to signal. To circumvent some of these problems, efforts are being made to examine membrane protein structure using NMR spectroscopy.

Unfortunately, high-resolution solution-state NMR of integral membrane proteins (IMPs) is limited by the size of the complex under investigation due to slow reorientation that results in line broadening and relaxation times that are not appropriate for multidimensional experiments. GPCRs present a major challenge as they must be studied in a membrane environment, further adding to the overall size of the system, and their helical nature limits chemical shift dispersion. To date, NMR investigations have been reported for four heptahelical integral membrane proteins, the vasopressin receptor, the CXCR1 receptor, and sensory rhodopsin and proteorhodopsin, GPCR analogues. Only 80

peaks corresponding to the vasopressin receptor were observed by solution NMR (*31, 32*). Solid and solution state NMR (*33, 34*) of sensory rhodopsin provided insight into the secondary structure and dynamics of the protein and also led to nearly complete assignment of the backbone. Solution-state NMR of proteorhodopsin produced similar results, with 96% of the backbone assigned and 44% of the TM side chain resonances assigned (*35*). These assignments have been used to probe the structural changes of ground state proteorhodopsin (*36*). While high-resolution solution state structures of sensory rhodopsin and proteorhodopsin were published recently (*35, 37*), these proteins are significantly smaller in size than most GPCRs, and application of the solution-state methods used to a full-length GPCR still remains to be demonstrated. The most significant contribution to the field of GPCR NMR has recently come from the Opella lab. Solid-state NMR was used for the determination of the NMR structure of the mammalian CXCR1 receptor in phospholipid bilayers (*38*). This study represents a tremendous breakthrough; however, there were remaining limitations. Backbone assignments are missing for 45 of 350 residues. These missing assignments are from the flexible N- and C-terminal tail regions. Furthermore, the calculated structure was generated using the molecular fragment replacement approach (*39*) using the structure of bovine rhodopsin as a guide (*38*). Therefore, an NMR structure for a full-length, mammalian GPCR based strictly on measured constraints has yet to be determined.

Structural Characterization of GPCR Fragments

Given the challenges that remain with the complete receptor, efforts have been devoted to the expression, purification, and biophysical analysis of fragments of integral membrane proteins including GPCRs. Study of fragments improves the quality and

simplifies NMR spectra due to the reduced molecular weight and residue number. To date, no fragment of an integral membrane protein has been crystallized but high-resolution NMR characterization of several fragments of integral membrane proteins has been conducted.

Evidence for the biological relevance of studying GPCR fragments came from several studies focused on the co-expression of receptor fragments. Ridge and coworkers showed that co-expression of opsin fragments in COS-1 cells produced a functional receptor that bound retinal and activated transducin (*40, 41*). The Dumont group co-expressed fragments of Ste2p in yeast and found that these fragments reconstituted in the membrane to form a functional receptor (*42*). *In vivo* reconstitution has also been successful for the muscarinic and dopamine receptors (*43, 44*) and *in vitro* reconstitution in detergent micelles was reported for bacteriorhodopsin (*45-47*). In addition to structural information, studies on GPCR fragments may provide insights into the folding pathway followed by the complete protein. Fragments of the cannabinoid receptors 1 and 2 (CB1 and CB2), human thromboxane A2 receptor, the human Y4 receptor, and the yeast α factor receptor have all been subject to either chemical or biological synthesis and biophysical analysis (*1, 48-60*). These studies have yielded the secondary and tertiary structures of such fragments and have also helped to explain functional aspects of these receptors.

Ste2p

Our group has been focusing on the NMR analysis of fragments of the yeast α-factor receptor, Ste2p. This protein is a 431 residue, peptide ligand receptor which we are using as a model system for GPCR methods development. As it is a eukaryotic

GPCR, it is similar in size to medically relevant, mammalian GPCRs. The yeast genome is well characterized and easy to manipulate. Functional assays for this receptor are well established, and mutations made in the Ste2p gene can be easily screened for functional implications (*61*). Most importantly, fragments of this receptor can be recombinantly expressed in bacteria, which allows for isotopic labeling required for NMR analysis (*62*). Functionally, Ste2p is an integral component of the mating pathway of *Saccharomyces cerevisiae*. *S. cerevisiae* is found in two haploid forms: MATa and MATα. Both cell types release peptide pheromones: a-factor and α-factor. Ste2p is present in MATa cells and is responsive to α-factor. Upon ligand binding, a MAP kinase cascade is activated, which can lead to transcriptional activation, G1 arrest, morphogenic changes, and cell fusion (*63, 64*).

We began our NMR characterization of Ste2p using peptides corresponding to single transmembrane domains (*1, 58-60, 65*) and then expanded our analyses to fragments containing two TM domains (*56, 57, 66*). We have published the only solution structure for a GPCR fragment containing the first two TMs [TM1-TM2; Ste2p(G31-T110)] in LPPG micelles and in 2,2,2-trifluoroethanol (TFE):water mixtures (*55, 57*). In both cases, the fragment is helical and forms a hairpin. However, the helical hairpin is stable in LPPG and transient in TFE. The formation of tertiary structure, even transient tertiary structure, supports the hypothesis that large domains of a GPCR can fold independently of the remainder of the protein.

To date, almost all structural studies of peptide fragments of GPCRs and other IMPs were conducted on molecules composed of one (*48, 67-69*) or two transmembrane domains (*54, 55, 57, 70*). An analysis of the known X-ray structures of GPCRs makes it

clear that almost all TMs are in contact with more than one other transmembrane domain. Larger fragments of a GPCR may, therefore, adopt a more stable tertiary structure due to additional inter-helical contacts that may provide interaction partners for polar side chains that would otherwise be exposed to lipids in the hydrophobic core. Nevertheless, solution of the structure of the larger fragments will be technically more challenging.

To evaluate advantages and disadvantages of using larger fragments my thesis project set out to study two three transmembrane-containing fragments of Ste2p, TM1-TM3 and TM127. TM1-TM3 [Ste2p(G31-R161)] was found to have a high buried surface area using a template model analysis. This fragment contains 131 residues of Ste2p, including 19 residues from the N-terminal domain, the first TM through the third TM with connecting loops and five residues of the second intracellular loop. TM127 [Ste2p(G31-T114,T274-L340)] was designed based on the observation that TM1 is predicted to interact with both TM2 and TM7 in a computer model of Ste2p (*71*). It contains 30 residues of the N-terminal domain, the first and second TMs with the connecting IL, an extracellular loop containing eleven residues of EL1 and two residues of EL3, TM7 and 40 residues of the C-terminal tail. Both of these 3TM fragments contain functionally relevant regions of the receptor. Residues in TM1 and TM7 have been implicated in ligand binding and receptor dimerization(*72-77*). Residues in EL1 are believed to interact with residues in TM1 and stabilize the inactive state (*78*). Therefore, structural characterization of these two fragments may provide insight into structure-function relationships for this receptor.

Our working hypothesis was that TM1-TM3 and TM127 would fold to more stable tertiary structures than their two transmembrane containing analog because TM2 is

predicted to interact with both TM1 and TM3 in the TM1-TM3 fragment and TM1 is predicted to interact with TM2 and TM7 in the TM127 fragment (*71*). These 3TM fragments were analyzed by NMR spectroscopy to determine whether they adopt a three-dimensional structure and to test the validity of the predicted inter-TM contacts.

Another collaborative goal of my thesis was to determine whether NMR assignments of small fragments could be used as a guide for the assignment of large fragments. The motivation behind this was to determine whether assignments of small fragments could be used to make assignments for a full-length receptor where there is increased overlap and amino acid redundancy. Toward this goal, I expressed and purified an isotopically labeled fragment consisting of residues 31-78 of Ste2p, encompassing 30 residues of the NT-region, the first TM and IL. Extensive NMR analysis in detergent micelles was performed on this 1TM construct and the TM7CT40 construct (*1*). The results are currently being analyzed to assess the transferability of chemical shifts.

This dissertation describes the expression and biophysical characterization of the Ste2p fragments TM1, TM1-TM3, TM127, and TM7CT40. The goals of the project were to 1) express and purify the TM1-TM3, TM127, and TM1 fragments, 2) characterize the secondary structure of the TM1-TM3 and TM127 fragments in membrane mimetic environments using circular dichroism spectroscopy, 3) perform solution-state NMR experiments to determine conditions for structural analysis of TM1-TM3 and TM127, 4) characterize the three-dimensional structure of the TM1-TM3 fragment in trifluoroethanol:water, and 4) perform heteronuclear NMR experiments on the TM1 and TM7CT40 fragment for backbone and sidechain assignment to assess chemical shift transferability.

SPECIFIC AIMS

The main goals of this research were to characterize three-transmembrane domain containing fragments of Ste2p using NMR spectroscopy and to determine the chemical shift transferability of Ste2p fragments.

I. Characterize the 3TM containing fragments of Ste2p: TM1-TM3 and TM127
    1. Clone, express, and purify the Ste2p G31-R131 construct.
    2. Express and purify the Ste2p G31-T114,T274-L340 construct.
    3. Characterize the secondary structure of these fragments using CD spectroscopy.
    4. Optimize conditions for NMR characterization of these fragments in detergent micelles.
    5. Determine the three dimensional structure of the TM1-TM3 fragment in membrane mimetics: TFE/water.

II. Determine the transferability of chemical shifts from small fragments of Ste2p to larger fragments.
    1. Express and purify the Ste2p G31-T78 construct.
    2. Complete backbone assignments for the TM1 and TM7CT40 fragments in 4:1 LPPG:DPC micelles.
    3. Determine the three dimensional structure of the TM1 fragment in detergent micelles.
    4. Compare the assignments for TM1 to TM1-TM2, TM1-TM3, and TM127 and the assignments for TM7 to TM127.

CHAPTER 2

Cloning, Expression and Purification of Regions of Ste2p: Ste2pTM1-TM3 (G31-R161), Ste2p TM127 (G31-T114, T274-L340), and Ste2pTM1 (G31-T78)

Introduction

As evidenced in Chapter 1, expression, purification, and biophysical analysis of IMPs, particularly GPCRs, is not trivial. Therefore, careful planning went into choosing the Ste2p fragments that we wished to study. Previous studies from our lab had been conducted on the first two TMs of Ste2p. This fragment was chosen based on a buried surface area analysis of the Ste2p sequence (*56*). When considering all of the possible 2TM combinations of Ste2p, TM1-TM2 was found to have the highest buried surface area, suggesting that this fragment would have the highest potential to form tertiary contacts. NMR structural analysis of the TM1-TM2 fragment revealed that it does adopt a tertiary structure in both LPPG micelles (*57*) and 50% trifluoroethanol:water (TFE:water) (*55*); however, the structure was found to be transient in organic-aqueous media. Based on these results, we hypothesized that increasing the size of the fragment to include an additional TM would increase the probability of tertiary contact formation and lead to a more defined tertiary structure. This hypothesis was supported by a rhodopsin-templated computer model of Ste2p (*71*), which predicts that TM2 should form helix-packing interactions with both TM1 and TM3 and that TM1 should form contacts with both TM2 and TM7. Therefore, by increasing the number of potential tertiary contacts for TM2 and TM1, respectively, the TM1-TM3 and TM127 fragments should form better defined structures than the TM1-TM2 fragment.

Another major goal of this project was to determine whether chemical shifts from small fragments could be used to assign larger fragments. To this end, NMR assignments

for TM1, TM7, TM1-TM2, TM1-TM3, and TM127 of Ste2p in detergent micelles were desired. As TM1-TM2 and TM7 had been previously cloned, expressed, and purified (*1, 56*), TM1, TM1-TM3, and TM127 were the only constructs that were cloned, expressed, and purified as part of my thesis project. This chapter describes the generation of expression vectors for these three constructs, optimization of expression in *Escherichia coli*, and isolation of desired protein fragments.

Materials and Methods

Cloning of Ste2p TM1-TM3 (G31-R161) and Ste2p TM127 (G31-T114, T274-L340) as TrpΔLE Fusion Proteins

Previous work from our lab showed that expression of Ste2p fragments as TrpΔLE (*79*) fusion proteins leads to high expression levels (*1, 56, 60, 80*). The TrpΔLE fusion tag directs the expressed fragments into inclusion bodies (IB). The tag can then be removed by performing cyanogen bromide (CNBr) cleavage directly on the IB pellets, and the cleaved peptide of interest can be purified by reverse phase high performance liquid chromatography (RP-HPLC). Therefore, the Ste2p TM1-TM3 and TM127 fragments were initially cloned as TrpΔLE fusion proteins.

A methionine-deficient template of Ste2p, pBluescriptSK(-) BEC2(M54L, M69V, M71I, M165I), was generated from the pBluescriptSK(-)BEC2 vector (*78*). To facilitate CNBr cleavage of the expressed fusion proteins, the Met residues in the template Ste2p vector were mutated by double-stranded PCR mutagenesis to other hydrophobic residues that were chosen based on previous mutational analysis of functional receptors (*81*).

PCR amplification of the region Ste2p(G31-R161) was performed using pBluescriptSK(-) BEC2(M54L, M69V, M71I, M165I) as a template. Primers were designed containing the *Hind*III and *Bam*HI restriction sites. The PCR product and the

pSW02 (*60*) vector were digested with *Hind*III and *Bam*HI (NEB). The digested DNA was gel purified and then ligated using T4 DNA ligase. The resulting reaction mixture was transformed into DH5α (Invitrogen) and the DNA was purified and sequenced. The amplified plasmid, pKC01, contained the TrpΔLE leader sequence in phase with Ste2p(G31-R161, M54L, C59S, M69V, M71I; see Figure 2-1).

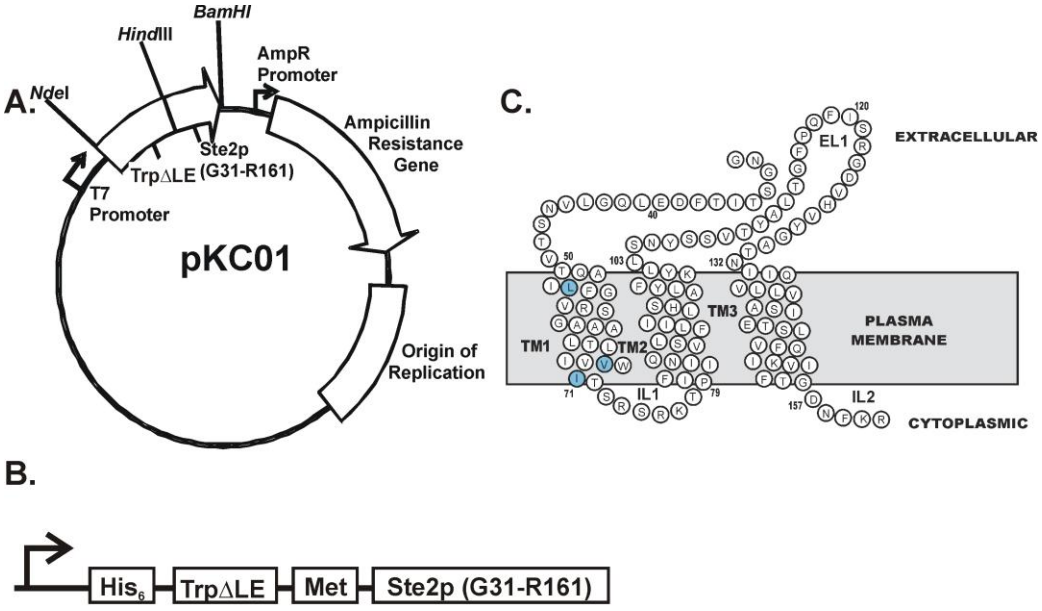

FIGURE 2-1 Cloning of Ste2p TM1-TM3 (G31-R161) Fusion Protein. (A) The pKC01 plasmid was constructed by amplification of the G31-R161 region of the pBluescriptSK(-) BEC2(M54L,M69V,M71I,M165I) plasmid and insertion of this DNA into the pSW02 vector using the HindIII and BamHI restriction sites. (B) E. coli expression of the pKC01 plasmid generated a fusion protein consisting of a six residue Histidine tag, the TrpΔLE fusion tag, a methionine residue, and Ste2p G31-R161. (C) CNBr cleavage of the expressed fusion protein resulted in a 131-residue polypeptide containing 19 residues from the N-terminal domain, the first TM through the third TM with connecting loops and five residues of the second IL. The mutated methionine residues (blue) are indicated in (C).

Construction of the Ste2p TM127 (G31-T114, T274-L340) expression vector was accomplished using a two step process. PCR was performed using primers designed to remove residues G115-G273 and the pBluescriptSK(-) BEC2(M54L, M69V, M71I, M165I) vector as a template. The resulting sequence was PCR amplified using a forward primer containing a *Hind*III site and a sequence complementary to the N-terminal sequence of the TM127 construct and a reverse primer containing a *Bam*HI site and a

sequence complementary to the C-terminal sequence of the TM127 construct. The amplified fragment and the pSW02 vector were digested with *Hind*III and *Bam*HI, gel purified, ligated using T4 DNA ligase, and transformed into DH5α. DNA from some of the resulting colonies was purified and sequenced. The resulting plasmid, pSW127, contained the TrpΔLE leader sequence in phase with Ste2p(G31-T114,T274-L340, M54L, C59S, M69V, M71I, M294L; see Figure 2-2).

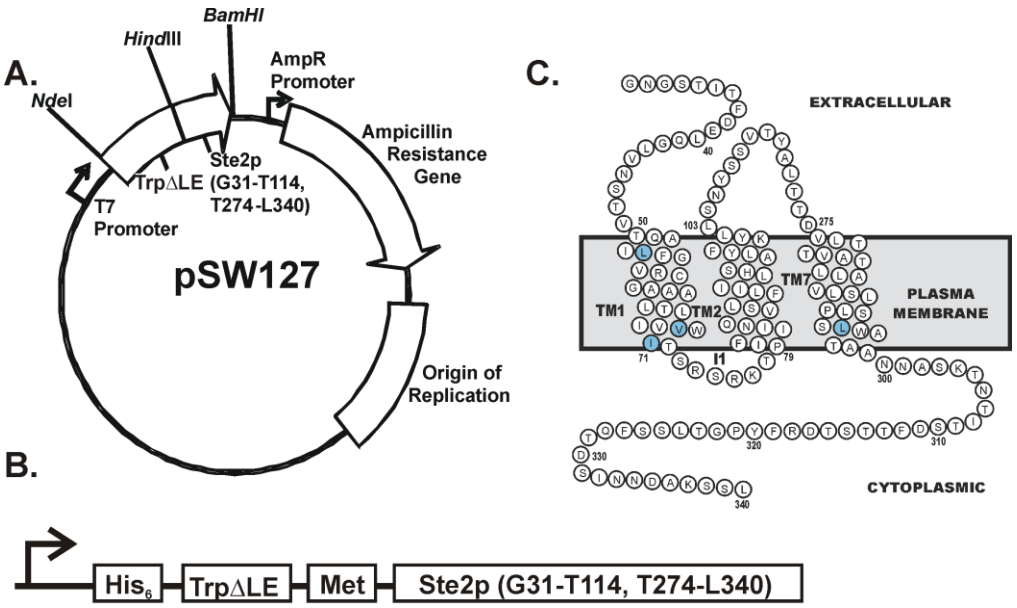

FIGURE 2-2 Cloning of Ste2p TM127 (G31-T114, T274-L340) Fusion Protein. (A) The pSW127 plasmid was constructed by deletion of residues G115-G273 from the pBluescriptSK(-) BEC2(M54L, M69V, M71I, M165I) vector, amplification of the G31-T114,T274-L340 region of the resulting plasmid, and insertion of this DNA into the pSW02 vector using the HindIII and BamHI restriction sites. (B) E. coli expression of the pSW127 plasmid generated a fusion protein consisting of a six residue Histidine tag, the TrpΔLE fusion tag, a methionine residue, and Ste2p G31-T114, T274-L340. (C) CNBr cleavage of the expressed fusion protein resulted in a 151-residue polypeptide containing 19 residues from the N-terminal domain, the first TM through the second TM with connecting loops, an engineered extracellular loop containing residues from EL1 and EL3, TM7 and 40 residues of the C-terminal tail. The mutated methionine residues (blue) are indicated in (C).

Cloning of Ste2p TM1-TM3 (G31-R161) and Ste2p TM1 (G31-T78) into Direct Expression Vectors

Three dimensional structure determination using NMR requires the identification of long range interactions. I planned to use Paramagnetic Relaxation Enhancement (PRE) experiments using a nitroxide radical spin label to find these long range

connectivities in TM1-TM3. In order to accomplish this, a single Cys residue was required in the TM1-TM3 sequence for spin label attachment. The experience of the Naider lab was that CNBr cleavage is unsuccessful in the presence of Cys residues. Therefore, the TM1-TM3 Cys mutant was cloned as a direct expression vector. The G31-R161 region of the Ste2p sequence was amplified from the pKC01 vector using a forward primer containing a His tag and an *Nde*I restriction site and the *Bam*HI reverse primer described above. The resulting PCR product and the pKC01 vector were digested with *Nde*I and *Bam*HI. The digests were gel purified and then ligated using T4 DNA ligase. The resulting reaction mixture was transformed into DH5α and the DNA was purified and sequenced. The amplified plasmid, pRE01, contained a 6 residue His tag in phase with Ste2p(G31-R161, M54L, C59S, M69V, M71I).

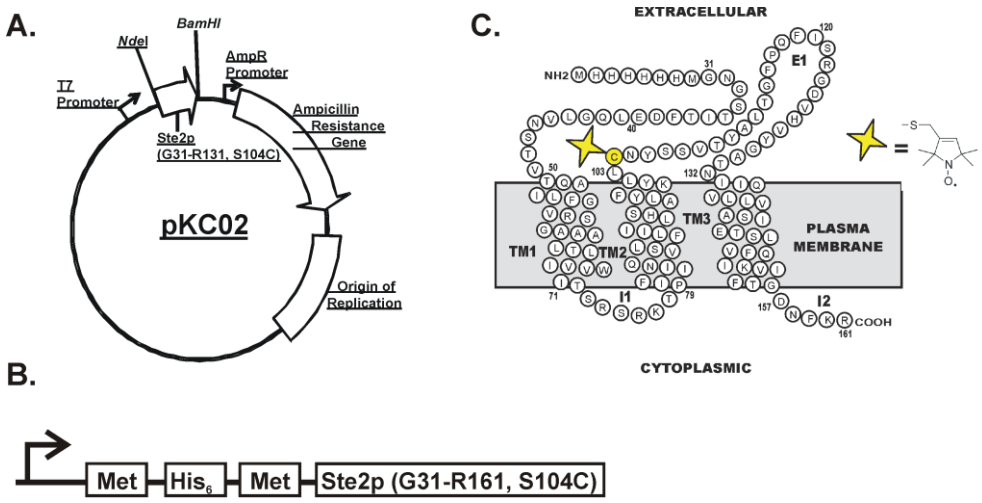

FIGURE 2-3 Direct expression of Ste2p TM1-TM3 (G31-R161, S104C). (A) The pKC02 plasmid was constructed by double stranded mutagenesis of the pRE01 plasmid, which was generated by PCR amplification of the G31-R161 region of the pKC01 plasmid and insertion of this DNA into the pSW02 vector using the *Nde*I and *Bam*HI restriction sites. (B) E. coli expression of the pKC02 plasmid generated a polypeptide consisting of a six residue Histidine tag, a methionine residue, and Ste2p G31–R161 with the S104C mutation. (C) The expressed protein was identical to the TM1TM3 construct generated from CNBr cleavage of the pKC01 expression product, except for the presence of a His tag and a Cys residue at position 104. This Cys was used to attach the MTSL spin label (represented by a yellow star) via a disulfide linkage.

Double-stranded mutagenesis was performed on the pRE01 plasmid to produce two additional expression vectors. One set of mutagenic primers was designed to incorporate a single point mutation in the pRE01 TM1-TM3 sequence such that S104 would be mutated to a Cys. A second set of mutagenic primers was designed to incorporate two stop codons into the pRE01 sequence after T78, resulting in a direct expression vector for Ste2p TM1(G31-T78). PCR mutagenesis was performed using the designed mutagenic primers, and the resulting products were digested with *Dpn*I. The digested DNA was transformed into DH5α, ten colonies were isolated, and the plasmid DNA was extracted from each of the colonies. The extracted DNA was sent for sequencing. The resulting expression vectors for TM1-TM3S104C and TM1 were called pKC02 (See Figure 2-3) and pKC03 (See Figure 2-4), respectively.

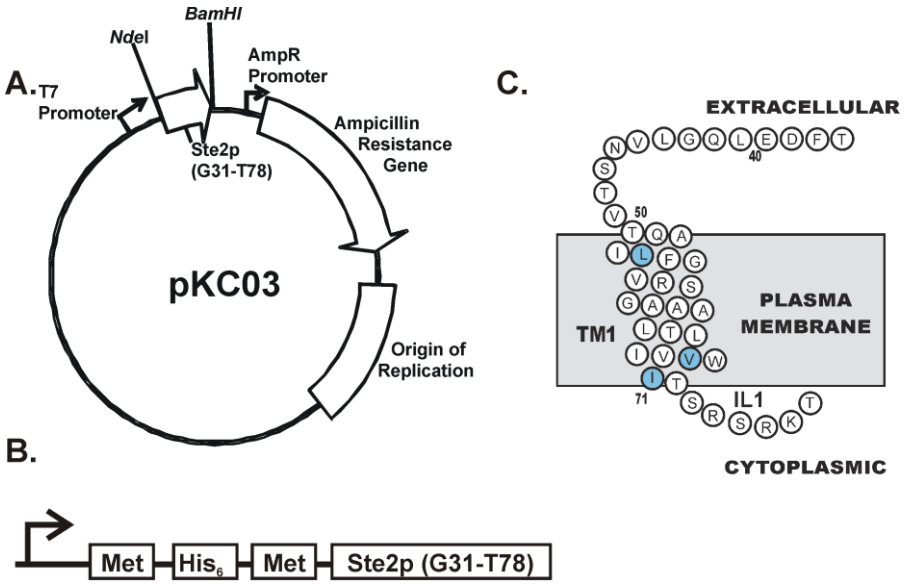

FIGURE 2-4 Direct expression of Ste2p TM1 (G31-T78). (A) The pKC03 plasmid was constructed by double stranded mutagenesis of the pRE01 plasmid, which was generated by PCR amplification of the G31-R161 region of the pKC01 plasmid and insertion of this DNA into the pSW02 vector using the *Nde*I and *Bam*HI restriction sites. (B) E. coli expression of the pKC03 plasmid generated a polypeptide consisting of a six residue Histidine tag, a methionine residue, and Ste2p G31-T78. (C) Purification of the expressed protein resulted in a 56-residue polypeptide containing 19 residues from the N-terminal domain, the first TM, and the first IL. The mutated methionine residues (blue) are indicated in (C).

Optimization of the Expression of Ste2p (G31-R161) TM1-TM3 and (G31-T114 T274-L340) TM127 Fusion Proteins

Expression of the TM1-TM3 fusion protein (TM1-TM3-FP) was tested in the *E.coli* expression strains BL21-AI, BL21-DE3, BL21-DE3 pLysS, BL21 Star DE3 pLysS (Invitrogen), and T7 express (NEB). Fifty milliliter expression tests were conducted in each of the cell lines in Luria Broth (LB) at 37°C in order to determine which cell line would be optimal for large-scale expression. Colonies were selected from the transformation plates and grown overnight in 5 mL of LB containing 200 μg/mL ampicillin (Amp) at 37°C. After overnight incubation, 2.5 mL of each overnight culture were used to inoculate 50 mL of LB/Amp, and the cells were grown at 37°C with shaking at 250 rpm until an optical density at 600 nm (OD_{600}) of approximately 0.4 was reached. Expression was induced with either 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) or 0.5% L-arabinose, and the cultures were incubated at 37°C, 250 rpm for 6 h. One milliliter samples were collected before induction and 2, 4, and 6 h after induction.

The 1 mL samples collected during the expression tests were used to prepare IBs for SDS-PAGE analysis according to a previously developed protocol (*1*). Briefly, the cells were lysed by sonication in lysis mix: lysis buffer (50 mM Trisma base, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0), 100 mM phenylmethanesulfonylfluoride (PMSF), and 10 mg/mL lysozyme). After lysis, the samples were centrifuged at top speed at 4°C for 10 min. The pellets were then sonicated in mild detergent (1% igepal, 1% deoxycholic acid, prepared in lysis buffer) and centrifuged. The final IB pellets were sonicated in SDS loading buffer (4% SDS, 12% glycerol, 50 mM Trisma base pH 6.8, 2% β-mercaptoethanol), and 20 μL of each sample

was run on 12% polyacrylamide gels. BL21-AI cells were observed to be the optimal cell line for expression of TM1-TM3FP.

As expression in minimal medium is required for isotopic labeling, expression of TM1-TM3FP in BL21-AI cells was optimized in M9 minimal medium (*82*) as to induction temperature and inducer concentration. pKC01 was transformed into BL21-AI cells. The transformed cells were grown in LB/amp and incubated at 37°C with shaking at 250 rpm to an OD_{600} of 0.4. Expression was induced with the addition of varying amounts of L-arabinose and IPTG at 37°C and 22°C in both LB and M9 minimal media. One milliliter samples were collected before induction and 2, 4, 6, and 22 h after induction to follow expression. Inclusion bodies of the 1 mL samples were prepared for SDS-PAGE analysis as described above. Optimal conditions were considered to be the inducer concentration, the temperature, and the time that led to the largest amount of protein produced with the lowest background expression as judged by SDS-PAGE.

As expression in BL21-AI cells was found to be optimal for TM1-TM3FP, initial expression attempts for TM127FP were made in this cell line. A 50 mL expression test in LB for BL21-AI cells transformed with pSW127 revealed that the construct did not express well in this cell line. 50 mL expressions were then conducted in M9 minimal medium in Bl21-DE3, Bl21-DE3 pLysS, BL21 Star-DE3 pLysS, and T7 express as described above. BL21-DE3 was initially found to be the optimal cell line for expression of this construct. However, the high expression levels observed in the test expressions with this cell line were not reproducible. Further optimization was performed in BL21 Star DE3 pLysS, and all subsequent expressions were conducted in this cell line using 0.5 mM IPTG as the inducer and induction at 22°C.

Large Scale Expression of Isotopically Labeled Ste2p TM1-TM3 (G31-R161) and TM127 (G31-T114, T274-L340) FP for NMR Analysis

Uniformly [^{15}N]-labeled TM1-TM3FP and TM127FP were expressed in 1 L cultures of BL21-AI cells transformed with pKC01 and BL21 Star-DE3 cells transformed with pSW127, respectively. One liter of LB/amp was inoculated with a 50 mL overnight culture and the cells were grown to an OD_{600} of ~ 0.35 by incubation at 37°C and shaking at 250 rpm. The cells were pelleted by centrifugation at 3000 rpm for 15 min at 4°C and resuspended in M9 minimal media containing $^{15}NH_4Cl$. Expression was induced under optimal conditions described above, the entire culture volume was pelleted in 500 mL aliquots at 5000 rpm, 4°C for 15 min, and inclusion bodies for direct cleavage were prepared as described previously (*56*). Briefly, 500 mL pellets collected after the 22 h expression were resuspended in lysis mix, sonicated, and centrifuged at 18,000 rpm, 4°C for 20 min. The pellets were then washed in the same manner with lysis buffer, mild detergent, and water. After the water wash, the pellets were resuspended in 4 mL of water, aliquoted into 1 mL aliquots in 1.5 mL microfuge tubes, and pelleted by centrifugation at 14,000 rpm, 4°C for 1 h.

[^{15}N, ^{13}C, ^{2}H]-TM1-TM3FP was expressed in a 1 L culture of pKC01 in BL21-AI. The cells were adapted for growth in D_2O through a series of 5 mL overnight cultures in M9 minimal media containing increasing amounts of D_2O (20%, 40%, 60% and 80%, respectively). After the overnight incubation, the cells were pelleted by centrifugation, and resuspended in M9 with an increased D_2O concentration. The final 5 mL 80% D_2O overnight culture was used to inoculate 50 mL of 80% D_2O M9 medium containing $^{15}NH_4Cl$ and uniformly labeled [^{13}C]-glucose as the sole nitrogen and carbon sources, and the cells were grown overnight at 37°C. The entire volume of the overnight culture was

added to 1.2 L of 80% D_2O M9 medium, and the cells were grown at 37°C until an OD_{600} of 0.35 was reached. The cells were then cooled to 22°C, and expression was induced with 0.2% L-arabinose. After 22 h, the cells were pelleted and inclusion bodies were prepared as described above.

The TM1-TM3 fusion polypeptide was also expressed to selectively protonate the methyl carbons of the isoleucine, leucine and valine (I, L, V, respectively) in an otherwise perdeuterated background (*83*). This was accomplished using 2-keto-3-methyl-d,3-d_1-1,2,3,4-^{13}C butyrate and 2-keto-3-d_2-1,2,3,4-^{13}C butyrate as the metabolic precursors for V and L and I, respectively. The pKC01 expression vector was transformed into BL-21 AI cells. One colony of the transformation was used to inoculate 6 mL of LB/amp and the culture was grown at 37°C, 250 rpm to OD_{600} of 1, and the cells were pelleted in 2 mL aliquots. One of the 2 mL pellets was resuspended in 20 mL of M9 minimal medium prepared with H_2O and grown to OD_{600} of 0.6. The 20 mL culture was pelleted and resuspended in 100 mL of M9 minimal medium prepared with $^{15}NH_4Cl$, [^{13}C]-glucose, and D_2O. The 100 mL culture was grown to OD_{600} of 0.4 and then diluted to 200 mL with M9 minimal medium prepared with $^{15}NH_4Cl$, ^{13}C-glucose, and D_2O. The 200 mL culture was grown to OD_{600} of 0.4 and then diluted to 1L with M9 minimal medium prepared with $^{15}NH_4Cl$, [^{13}C]-glucose, D_2O, and the labeled I, L, V metabolic precursors. The cells were again grown to OD_{600} of 0.4, expression was induced with 0.2% L-arabinose, and the culture was incubated at 37°C for 9.5 h. After the incubation, the cells were pelleted by centrifugation, and IBs were prepared as described above.

Direct Expression of Ste2p TM1-TM3(G31-R161) and TM1(G31-T78)

Expression tests were performed for the pRE01 and pKC02 expression vectors in BL21-AI, BL21-DE3, BL21-DE3 pLysS, BL21 Star-DE3 pLysS, and T7 express as described above. Expression results were analyzed by SDS-PAGE and Western blot analysis using anti-His antibody. Little to no expression was observed under these conditions. Fifteen milliliter expression tests were then conducted for pKC02 transformed into BL21-AI, BL21-DE3, BL21-DE3 pLysS, BL21 Star-DE3 pLysS, T7 express, C43, Origami, and Rosetta. Cells were streaked directly from the transformation plates and grown in LB to an OD_{600} of approximately 1. The pellets were resuspended in M9 and expression was induced as described above. Expression was observed in BL21-DE3, and expression was further optimized in this strain.

One liter expressions for the pKC02 vector in M9 minimal medium prepared with $^{15}NH_4Cl$ were conducted under the optimized conditions. The final cultures were pelleted in 500 mL aliquots, and inclusion bodies were prepared using the same protocol used for the 1 L FP expressions.

Fifteen milliliter expression tests for pKC03 transformed into BL21-AI, BL21-DE3, BL21-DE3 pLysS, BL21 Star-DE3 pLysS, T7 express, C43, Origami, and Rosetta were also conducted with growth in LB to an OD_{600} of approximately 1, followed by resuspension in M9 and induction. Expression was observed in BL21-AI. Two 500 mL expressions in M9 medium containing $^{15}NH_4Cl$ and $^{15}NH_4Cl/^{13}C$-glucose, respectively, were conducted with induction at 37°C overnight. The entire culture volume for each expression was pelleted and inclusion bodies were prepared using the same protocol as the large-scale FP expressions.

Purification of Ste2p TM1-TM3 (G31-R161)FP, TM127 (G31-T114, T274-L340)FP, and TM1 (G31-T78), TM1-TM3 and TM127 Peptides

Purifications of the TM1-TM3-FP, the TM127FP, the TM1-TM3S104C direct expression product, and of the TM1-TM3 and TM127 peptides released by CNBr cleavage were conducted by either analytical or preparative RP-HPLC. Analytical HPLC was performed on a Zorbax 300SB-C3 4.6 x 150 mm, 3.5 micron column at 60°C using gradient elution with a water/acetonitrile/2-propanol gradient. The reservoirs contained water/2-propanol/0.1% trifluoroacetic acid (TFA) and acetonitrile/2-propanol/0.1% TFA and the gradient was from 24%-72% acetonitrile with a constant 2-propanol concentration of 20%. Preparative HPLC was performed on a Zorbax 300SB-C3 Prep HT 21.2 x 150 mm, 7 micron column at 60°C using gradient elution with a water/acetonitrile/2-propanol/ 0.1% TFA gradient from 40%-72% acetonitrile with a constant 20% 2-propanol concentration.

Purification of the TM1 direct expression peptide was also performed using RP-HPLC with a water/acetonitrile/2-propanol gradient at 60°C. The reservoirs contained water/2-propanol/0.1% trifluoroacetic acid (TFA) and acetonitrile/2-propanol/0.1% TFA and the gradient was from 36%-72% acetonitrile with a constant 2-propanol concentration of 10%. Preparative HPLC was performed on a Zorbax 300SB-C3 Prep HT 21.2 x 150 mm, 7 micron column at 60°C using gradient elution with a water/acetonitrile/2-propanol/ 0.1% TFA gradient from 36%-81% acetonitrile with a constant 10% 2-propanol concentration.

For preparative purification of the TM1 direct expression polypeptide, a pellet corresponding to one-eighth of the culture volume was solubilized in 700 μL of 100% TFA. The solubilized pellet was sonicated at room temperature until a clear solution was

produced. After sonication, 300 μL of water were added dropwise to the sample while vortexing. The solution was sonicated again, and the entire volume was injected onto the HPLC. Two minutes into the HPLC gradient, 1 mL of 70% TFA was injected to ensure that all of the peptide was removed from the column during the run. The TM1-TM3S104C direct expression construct was purified from pellets corresponding to one-quarter of the total culture volume using the same protocol.

After purification, the HPLC fractions were combined and UV quantitated using Beer's law. Extinction coefficients were calculated using the following equation.

$$\varepsilon = (1490 \times \text{number of Tyr}) + (5500 \times \text{number of Trp}) + (125 \times \text{number of Cys})$$

CNBr Cleavage of the Ste2pTM1-TM3 (G31-R161) FP and the TM127 (G31-T114, T274-L340) FP

CNBr cleavage of the FPs was performed directly on the inclusion bodies. A pellet corresponding to one eighth of the culture was solubilized in 600 μL of 100% TFA. The solubilized pellet was sonicated at room temperature until a clear solution was produced. After sonication, 250 μL of water were added dropwise to the sample while vortexing. The solution was added to 120-140 mg of CNBr, vortexed, and placed in the dark for 1.5-2 h. After the cleavage, the TM1-TM3 or TM127 peptide was purified by preparative RP-HPLC. Following injection of the cleavage reaction, 1 mL of 70% TFA was injected at t = 2 min. This injection was followed by three 2 mL injections of 70% TFA in order to remove the cleaved peptide that stuck to the column. These column washes produced >70% of the total peptide that was recovered. All fractions collected during the purifications were combined, UV quantitated, and lyophilized, as described above.

Mass Spectrometry of Ste2p TM1-TM3 (G31-R161), TM127 (G31-T114, T274-L340), and TM1 (G31-T78)

Mass spectrometry (MS) was used to determine the molecular weight of the expressed protein and the percent of isotope incorporation for the [^{15}N]-, [^{15}N, ^{13}C]-, [^{15}N, ^{13}C, ^{2}H]-, and [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-ILV)]-labeled samples. MS was conducted on the FP and cleaved peptide eluted from the RP-HPLC using an Agilent 1100 LC/MS with an electrospray ion source and an ion trap detector. The percent isotope incorporation was determined as previously described (*56*). ESI-MS was also conducted on the TM1 and TM1-TM3S103C directly expressed polypeptides using the same protocol.

Results

Construction of TM1-TM3, TM127, and TM1 Expression Vectors

Interhelical packing interactions are essential to the overall fold of GPCRs. Crystal structures revealed that each TM helix was in contact with at least two other helices. The Smith group published a comprehensive analysis of a rhodopsin-templated computer model of Ste2p (*71*). In that model TM2 forms interhelical contacts with both TM1 and TM3 and TM1 forms interhelical contacts with both TM2 and TM7. This led us to choose the two three TM-containing constructs of Ste2p used in our study.

The TM1-TM3 and TM127 regions of Ste2p were cloned downstream of the TrpΔLE protein. The sequences of the resulting plasmids were verified to be TrpΔLE-TM1-TM3(G31-R161, M54L, C59S, M69V, M71I) and TM127(G31-T114, T274-L340, M54L, C59S, M69V, M71I, M294L) by DNA sequencing (sequencing data are included as an Appendix to this thesis). The TrpΔLE fusion tag was included to direct the expression of the fragment into inclusion bodies. To facilitate CNBr cleavage of the tag from the fragment, a Met residue was introduced between the TrpΔLE protein and the

Ste2p TM1-TM3 and TM127 sequences, in which all of the native methionine residues had been mutated to other nonpolar amino acid residues (Figures 2-1 and 2-2). These nonpolar residues were chosen based on a previous mutational analysis that indicated they resulted in functional Ste2p (*81*).

Expression vectors for the direct expression of TM1-TM3 and TM1 were also constructed. Double-stranded mutagenesis was performed on the pRE01 plasmid in order to produce the TM1-TM3S104C and TM1 mutants. For TM1-TM3S104C, a single nucleotide was changed in order to express a Cys in position 104 instead of the native Ser. The presence of a Cys at position 104 was desired for nitroxide radical spin label attachment. The paramagnetic spin label *S*-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSL) was attached via a disulfide linkage at this position (yellow star, Figure 2-3). The sequence of the resulting mutagenic plasmid, pKC02, was confirmed to be His_6-Ste2p TM1-TM3 (G31-R161, G31–R161, M54L, C59S, M69V, M71I, S104C). For the TM1 construct, two stop codons were inserted in succession after residue 78 in the native Ste2p sequence. The sequence of the resulting plasmid pKC03 was confirmed to be His_6-Ste2p TM1(G31-T78, M54L, C59S, M69V).

Expression of the Ste2p TM1–TM3 (G31–R161)FP and TM127 (G31-T114, T274-L340)FP

Test expressions for pKC01 revealed that the TM1-TM3FP construct was expressed in a variety of *E.coli* expression strains. Upon inspection of the expression gels, (Figure 2-5) the BL21-AI strain was determined to be the optimal stain for expression of this construct as the intensity of the band at ~24 kDa, which corresponds to the TM1-TM3FP (arrow; Figure 2-5), was highest in this strain. However, the level of background expression in all strains was high.

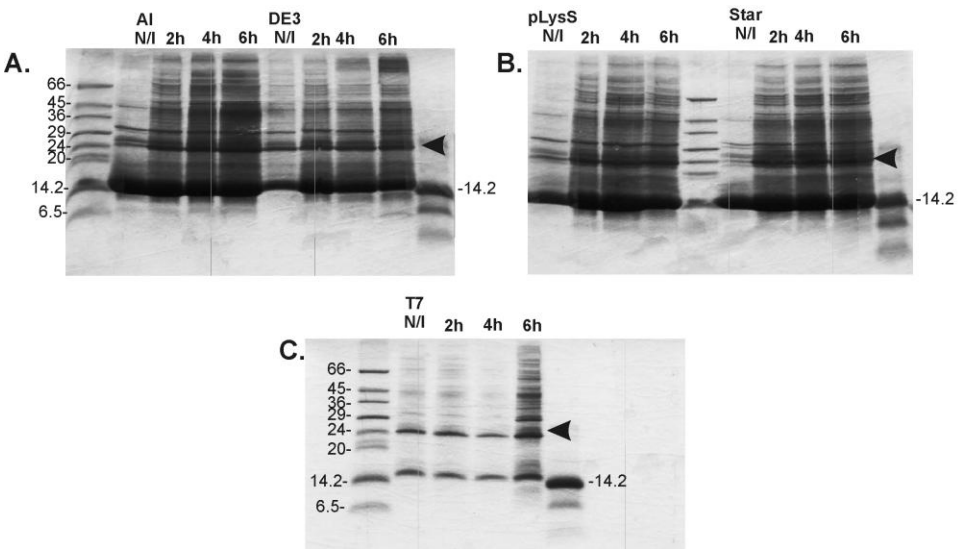

FIGURE 2-5 Optimization of cell lines for expression of Ste2p TM1–TM3 (G31-R161). The pKC01 plasmid was transformed into *E.coli* expression strains. Cells were grown at 37°C for 6h. The optimal cell line was determined to be BL21-AI. The arrow indicates the expected position of the TM1-TM3-FP.

Further optimization of expression in BL21-AI was conducted in order to find the induction temperature and inducer concentration that led to the highest amount of TM1-TM3FP expression with the least amount of background. This optimization was conducted in M9 minimal medium, which is required for isotopic labeling. As judged by the intensity of the fusion protein band and the level of background expression, optimal expression was obtained with 0.2% L-arabinose in M9 at 22°C for 22 h (Figure 2-6). Similar expression was also noted for induction with 0.5% L-arabinose in M9 for 4 h at 37°C. Based on the backgrounds obtained at longer inductions for the 37°C cultures we chose to do our expressions at 22°C. The molecular weight of the TM1-TM3-FP (28,340.65 Da) was confirmed by MS following RP-HPLC purification (Figure 2-8D). These conditions were used for all subsequent expressions.

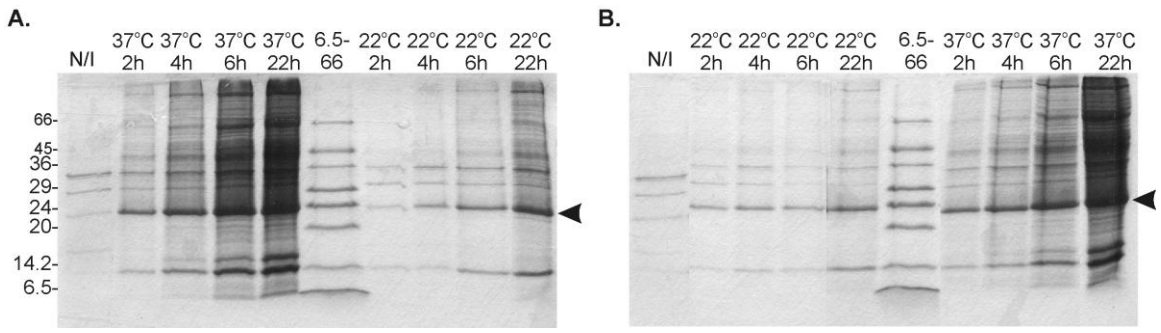

FIGURE 2-6 Optimization of expression of Ste2p TM1–TM3 (G31-R161) fusion protein in BL21-AI. TM1–TM3 expression was optimized in M9 minimal medium using the *E. coli* strain BL21-AI. Expression was induced using 0.2% (A) and 0.5% (B) L-arabinose at both 22°C and 37°C. The arrow indicates the expected position of the TM1-TM3-FP.

Initial expression attempts for TM127FP in BL21-AI were unsuccessful (data not shown). Optimization of expression in M9 minimal medium in a variety of expression strains revealed that expression in BL21-DE3 and BL21 Star DE3 pLysS for 22 h at 37°C produced large bands at the expected molecular weight for TM127 (~29 kDa) with little background expression (Figure 2-7, arrow). As the expected band was more intense for expression in BL21-DE3, this strain was chosen for optimization of large scale expression. However, the expression results were inconsistent and not reproducible in this strain (data not shown). Therefore, subsequent expressions were performed in BL21 Star DE3 pLysS.

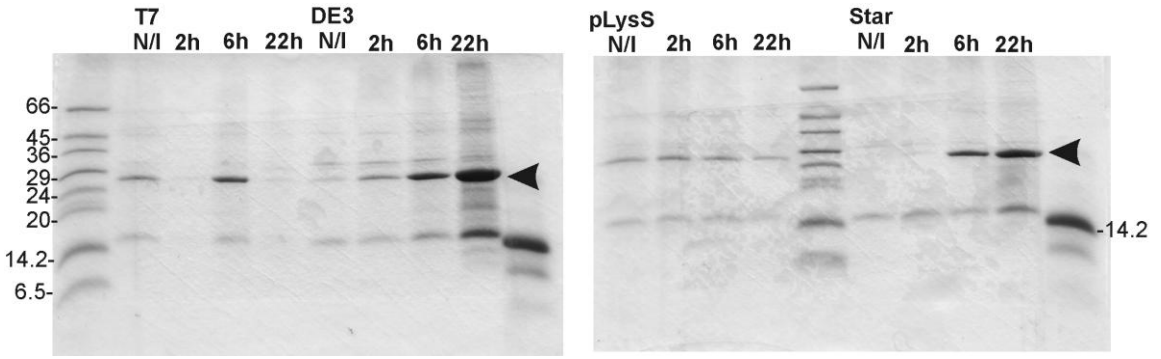

FIGURE 2-7 Optimization of cell lines for expression of Ste2p TM127 (G31-T114,T274-L340). The pSW127 plasmid was transformed into four *E.coli* expression strains. Cells were grown at 37°C for 22h. High expression levels with low background were observed for expression in BL21-DE3 and BL21 Star DE3 pLysS. The arrow indicates the expected position of the TM127FP.

Large-scale Expression and CNBr Cleavage of Ste2p TM1–TM3 (G31–R161) and TM127 (G31-T114, T274-L340) FPs

As high expression was obtained in minimal medium, large-scale expressions were conducted to generate uniformly [^{15}N]-labeled TM1-TM3 and TM127. My first attempts at large scale expression were conducted using pKC01 for expression of the TM1-TM3FP. One-liter cultures of the TM1-TM3FP were grown in M9 medium containing $^{15}NH_4Cl$ as the sole nitrogen source. Expression was induced with 0.2% L-arabinose at 22°C for 22 h and inclusion bodies were prepared. Fifty microliter pellets collected during the IB preparation were analyzed by analytical HPLC. The TM1-TM3 pellets showed a major peak at ~17.5 min corresponding to the fusion protein (Figure 2-8A). The overall homogeneity of the peptide was very good with only minor byproducts seen after the breakthrough peak.

Optimization of CNBr cleavage was performed using the 1 mL TM1-TM3FP IB pellets corresponding to one-eighth of the culture volume, and the cleavage was followed by analytical RP-HPLC (see Figure 2-8B). The best cleavage was obtained when the CNBr reaction proceeded for 1.5-2 h. Longer treatment resulted in loss of product (data not shown). After the reaction was approximately 95% complete as judged by HPLC, the cleavage product was purified by preparative RP-HPLC. The major peak, elution time = 31 min (Figure 2-8B), was collected and subjected to MS analysis by ESI-MS. The calculated mass for 100% [^{15}N]-labeled peptide was 14,588.9 Da and the observed mass of the product was 14,580.5 Da, corresponding to 95% isotope incorporation into TM1-TM3 (Figure 2-8E). It was observed that the cleaved peptide did not fully elute from the column during the first RP-HPLC purification. This is most likely caused by the high hydrophobicity of the fragment and the tendency of the E1 loop of Ste2p to aggregate

(*65*). All of the peptide was removed from the column by increasing the concentration of isopropanol in the purification gradient from 10% to 20% and subsequent washing of the column by injecting 70% TFA. The major peak from these TFA washes had the same elution time as that of the original purification, and ESI-MS analysis confirmed that the peak corresponded to cleaved TM1-TM3 peptide. The peaks collected during the initial purification and subsequent TFA washes were combined, UV quantitated, and lyophilized to yield 15-20 mg by UV, or 30-40 mg by weight of highly homogeneous product as judged by the quality of the HPLC and mass spec profiles (Figures 2-8C and 2-8E) from a 1 L culture. We believe that the discrepancy between the quantitations by UV and weight is due to the use of 20% 2-propanol in the HPLC gradient. A single lyophilization may not remove all of the 2-propanol from the material, leading to an increase in the measured weight of the product. This same discrepancy has been noted with many Ste2p fragments. Purification of the FP was only performed on small fractions of the total culture volume to determine the molecular weight by ESI-MS. Therefore, it is not possible to present yields at each point of the isolation process.

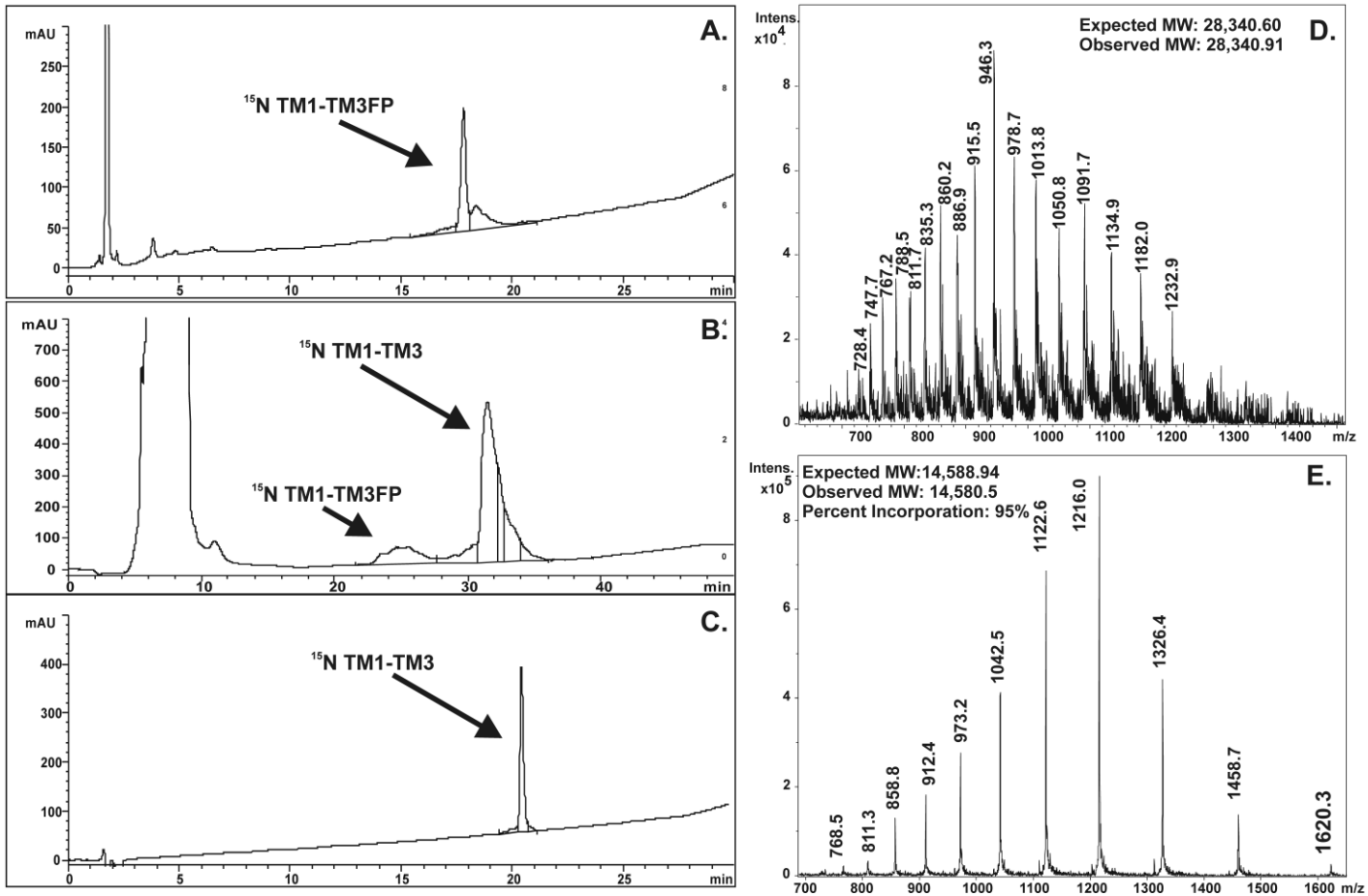

FIGURE 2-8 Cleavage and purification of [^{15}N]- Ste2p TM1-TM3(G31-R161). CNBr cleavage was performed on inclusion bodies isolated from expression of TM1-TM3-FP in M9 minimal medium containing $^{15}NH_4Cl$. (A) Analytical HPLC of the fusion protein before cleavage. The RP-HPLC analysis was performed on a Zorbax 300SB-C3 column (4.6 x 150 mm, 3.5 micron) using a 24%–72% acetonitrile, 20% 2-propanol, 0.1% TFA gradient at 60°C. (B) Preparative HPLC after cleavage for 2 h. Purification was performed on a Zorbax 300SB-C3 Prep HT (21.2 x 150 mm, 7 micron) using a 40–72% acetonitrile, 20% 2-propanol, 0.1% TFA gradient at 60°C. (C) Analytical HPLC of the purified cleavage product. (D) ESI-MS of the purified [^{15}N]-labeled fusion protein to verify mass. (E) ESI–MS of the purified peptide to determine incorporation level of [^{15}N].

Large-scale expression of [^{15}N]-TM127 was conducted in BL21 Star DE3 pLysS using 0.5 mM IPTG for induction at 37°C for 22 h. IBs were prepared and subject to CNBr cleavage and purification using the protocol developed for the TM1-TM3 construct. Analysis of a 50 μL IB pellet revealed that the TM127FP also showed excellent homogeneity of the desired product with a retention time of 27 min (Figure 2-9A). This retention time was approximately 10 min later than the elution of the TM1-TM3FP using the same HPLC gradient. As the TM127 construct contains 40 residues of the hydrophilic C-terminal tail, this result was surprising. Preparative purification of the TM127 peptide, released by CNBr cleavage, also resulted in elution 10 minutes later than the TM1-TM3 construct (Figure 2-9B). ESI-MS analysis of the purified cleavage product revealed that the final product was very pure, and that the incorporation of the [^{15}N] isotopic label was approximately 96%. This labeled polypeptide was purified in yields of up to 11 mg/L of culture volume as determined by UV.

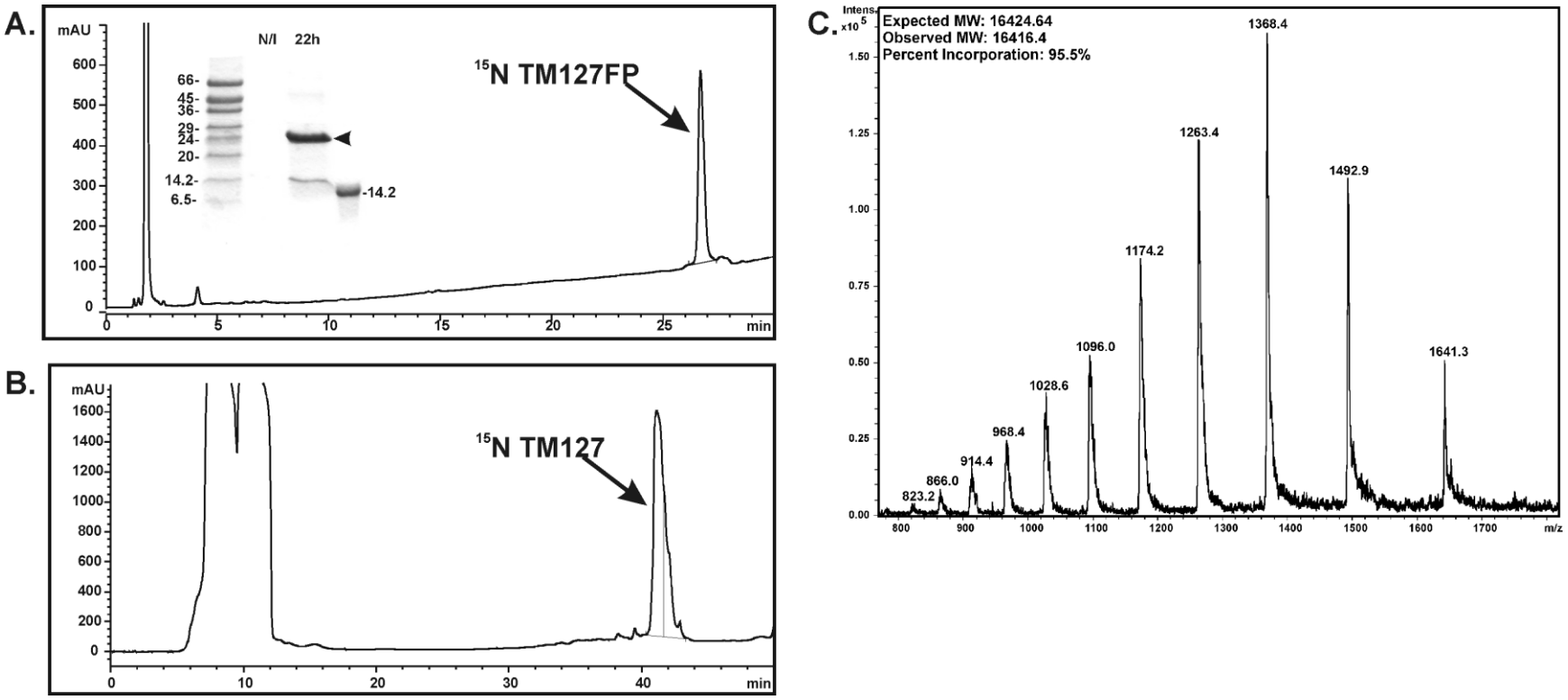

FIGURE 2-9 Cleavage and purification of [^{15}N]- Ste2p TM127 (G31-T114,T274-L340). CNBr cleavage was performed on inclusion bodies isolated from expression of TM127-FP in M9 minimal medium containing $^{15}NH_4Cl$. (A) Analytical HPLC of the fusion protein before cleavage. The RP-HPLC analysis was performed on a Zorbax 300SB-C3 column (4.6 x 150 mm, 3.5 micron) using a 24%–72% acetonitrile, 20% 2-propanol, 0.1% TFA gradient at 60°C. (B) Preparative HPLC after cleavage for 2 h. Purification was performed on a Zorbax 300SB-C3 Prep HT (21.2 x 150 mm, 7 micron) using a 40–72% acetonitrile, 20% 2-propanol, 0.1% TFA gradient at 60°C. (C) ESI–MS of the purified peptide to determine incorporation level of [^{15}N]

Inclusion bodies prepared for the [^{15}N, ^{13}C, ^{2}H]- and [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-ILV)]-labeled TM1-TM3 were processed and subjected to CNBr cleavage and purification as described above (Figures 2-10 and 2-11). The yield of the [^{15}N, ^{13}C, ^{2}H]-large-scale expression was 5.3 mg by UV, significantly lower than that achieved from expression in light water. MS analysis was performed to determine the percent of [^{2}H] incorporation. The incorporation of the [^{13}C] and [^{15}N] nuclei was assumed to be 95%, and the expected molecular weight for 100% [^{2}H]-labeled peptide was calculated to be 16,078.94 Da. The measured average molecular weight was determined to be 15,702 Da, with ~56% [^{2}H]-incorporation (Figure 2-10D). Cleavage and purification of the [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-ILV)]-labeled TM1-TM3 resulted in a total of 12 mg by UV of pure product with 95% [^{2}H] incorporation assuming 100% [^{15}N]- and [^{13}C]- incorporation and excluding all methyl I, L, V protons from the total proton count (Figure 2-11D). The purity of the final products based on their analytical HPLC and MS profiles was sufficient for subsequent NMR analysis (*vida infra*).

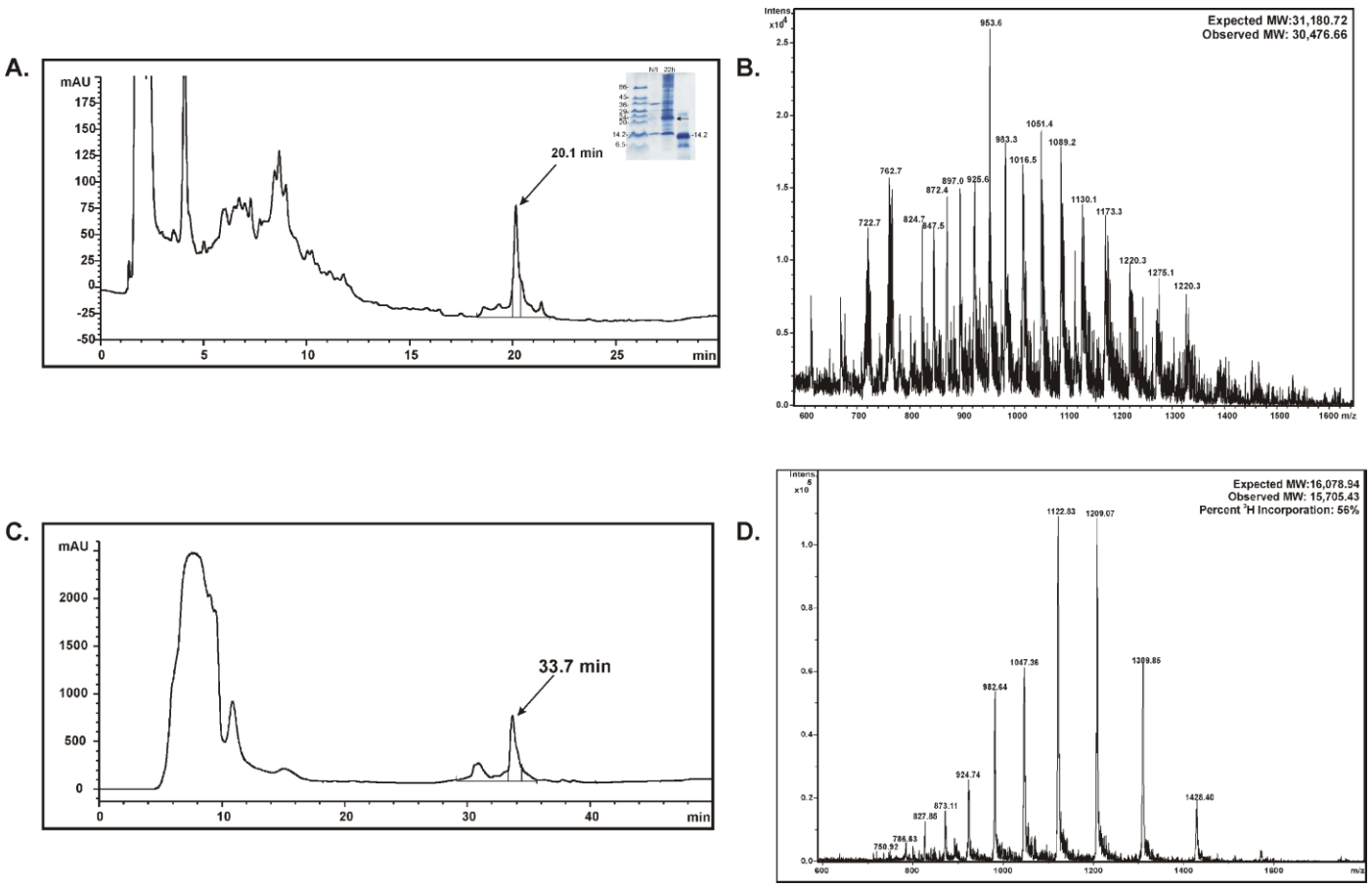

FIGURE 2-10 Cleavage and purification of [^{15}N,^{13}C,^{2}H]- Ste2p TM1-TM3(G31-R161). CNBr cleavage was performed on inclusion bodies isolated from expression of TM1-TM3-FP in M9 minimal medium prepared with $^{15}NH_4Cl$, ^{13}C-glucose, and D_2O. (A) Expression (inset) and analytical HPLC of the fusion protein before cleavage. The RP-HPLC analysis was performed on a Zorbax 300SB-C3 column (4.6 x 150 mm, 3.5 micron) using a 24%-72% acetonitrile, 20% 2-propanol, 0.1% TFA gradient at 60°C. (B) ESI-MS of the purified [^{15}N,^{13}C,^{2}H]-labeled fusion protein to verify mass (C) Preparative HPLC after cleavage for 2 h. Purification was performed on a Zorbax 300SB-C3 Prep HT (21.2 x 150 mm, 7 micron) using a 40–72% acetonitrile, 20% 2-propanol, 0.1% TFA gradient at 60°C. (D) ESI-MS of the purified [^{15}N,^{13}C,^{2}H]-labeled fusion protein to verify mass. (E) ESI–MS of the purified peptide to determine incorporation level of [^{2}H]. Incorporation of the [^{15}N] and [^{13}C] isotopes was assumed to be 95%.

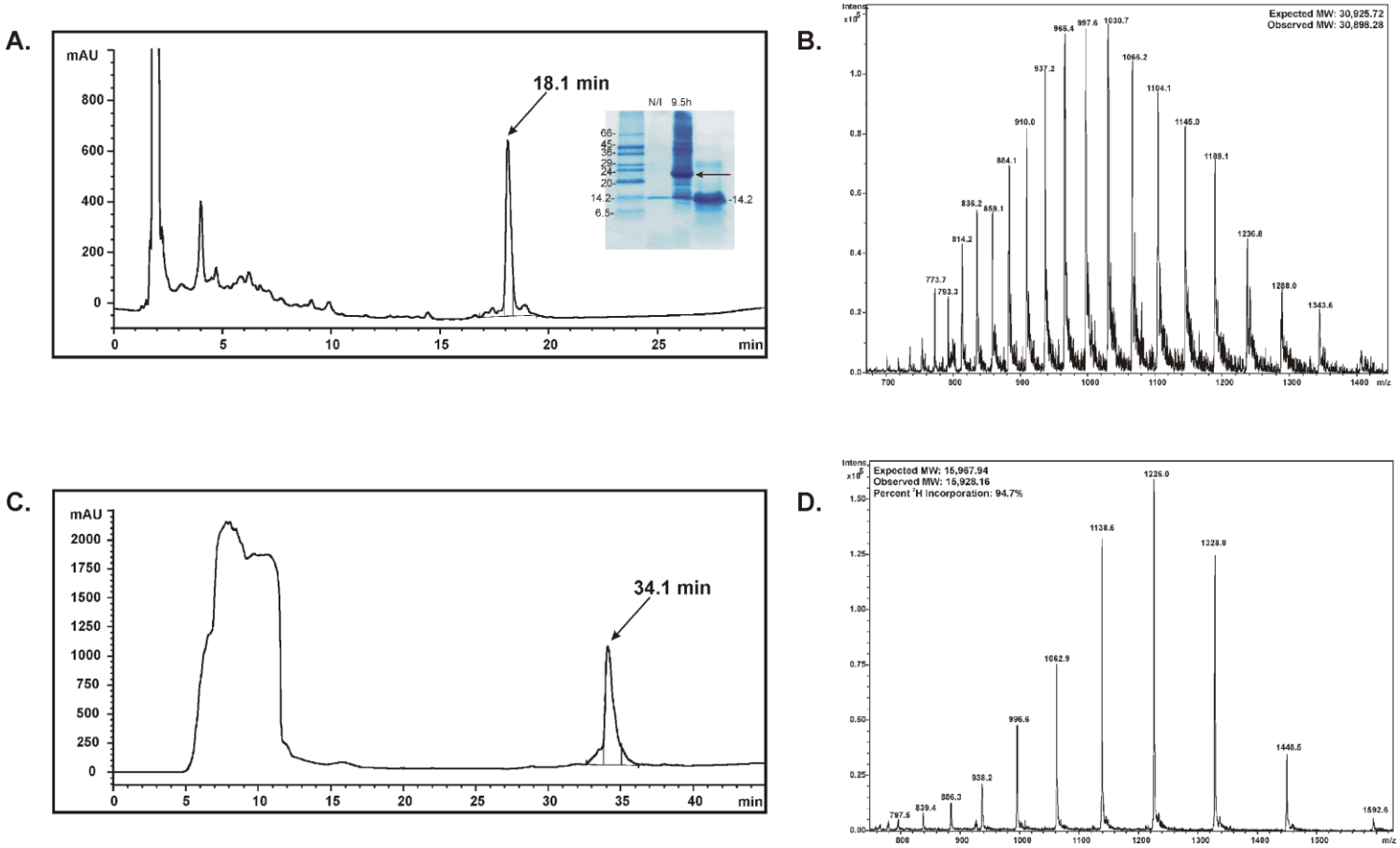

FIGURE 2-11 Cleavage and purification of selectively methyl protonated [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-ILV)]- Ste2p TM1-TM3(G31-R161). CNBr cleavage was performed on inclusion bodies isolated from expression of TM1-TM3-FP in M9 minimal medium prepared with $^{15}NH_4Cl$, ^{13}C-glucose, 2-keto-3-methyl-d,3-d_1-1,2,3,4-^{13}C butyrate, 2-keto-3-d_2-1,2,3,4-^{13}C butyrate and D_2O. (A) Expression (inset) and analytical HPLC of the fusion protein before cleavage. The RP-HPLC analysis was performed on a Zorbax 300SB-C3 column (4.6 x 150 mm, 3.5 micron) using a 24%–72% acetonitrile, 20% 2-propanol, 0.1% TFA gradient at 60°C. (B) ESI-MS of the purified [^{15}N,^{13}C,^{2}H]-labeled fusion protein to verify mass (C) Preparative HPLC after cleavage for 2 h. Purification was performed on a Zorbax 300SB-C3 Prep HT (21.2 x 150 mm, 7 micron) using a 40–72% acetonitrile, 20% 2-propanol, 0.1% TFA gradient at 60°C. (D) ESI-MS of the purified [^{15}N,^{13}C,^{2}H]-labeled fusion protein to verify mass. (E) ESI–MS of the purified peptide to determine incorporation level of [^{2}H].

Direct Expression of the Ste2p TM1-TM3(G31-R161, S104C) and TM1 (G31-T78) Constructs

Direct expression of TM1-TM3S104C was required as CNBr cleavage has been found to be unsuccessful in the presence of Cys residues[62]. Initial attempts at directly expressing the TM1-TM3 mutant using the optimization protocol described for the TrpΔLE fusion constructs were unsuccessful. Optimization of expression in a variety of expression strains revealed that the mutant was directly expressed in BL21-DE3 when the induction OD_{600} was increased from 0.4 to 1 (data not shown). SDS-PAGE and Western blot analysis of optimization tests in this cell line to determine the optimal induction time and temperature revealed that the desired protein was expressed at 37°C; 6 h and 22 h post induction, and at 30°C; 22 h post induction (Figure 2-12A). Almost no protein was observed at 22°C. In all three cases, the band at the expected molecular weight of ~15 kDa is not clearly evident on the Coomassie stained gel and a blurry band is obtained on the anti-His Western Blot. HPLC analysis was performed on a 15 mL IB pellet prepared from the 37°C, 22 h pellet. The initial injection resulted in a very low intensity peak with a retention time of ~20 minutes (data not shown). Injection of 70% TFA after the initial IB injection resulted in a high intensity band at the same retention time (Figure 2-12B). ESI-MS analysis of this fraction revealed that it had a molecular weight of 15,407, which is 15 Da larger than the expected molecular weight. It was decided that this molecular weight discrepancy could be disregarded, and two 1 L expressions in M9 prepared with $^{15}NH_4Cl$ were conducted.

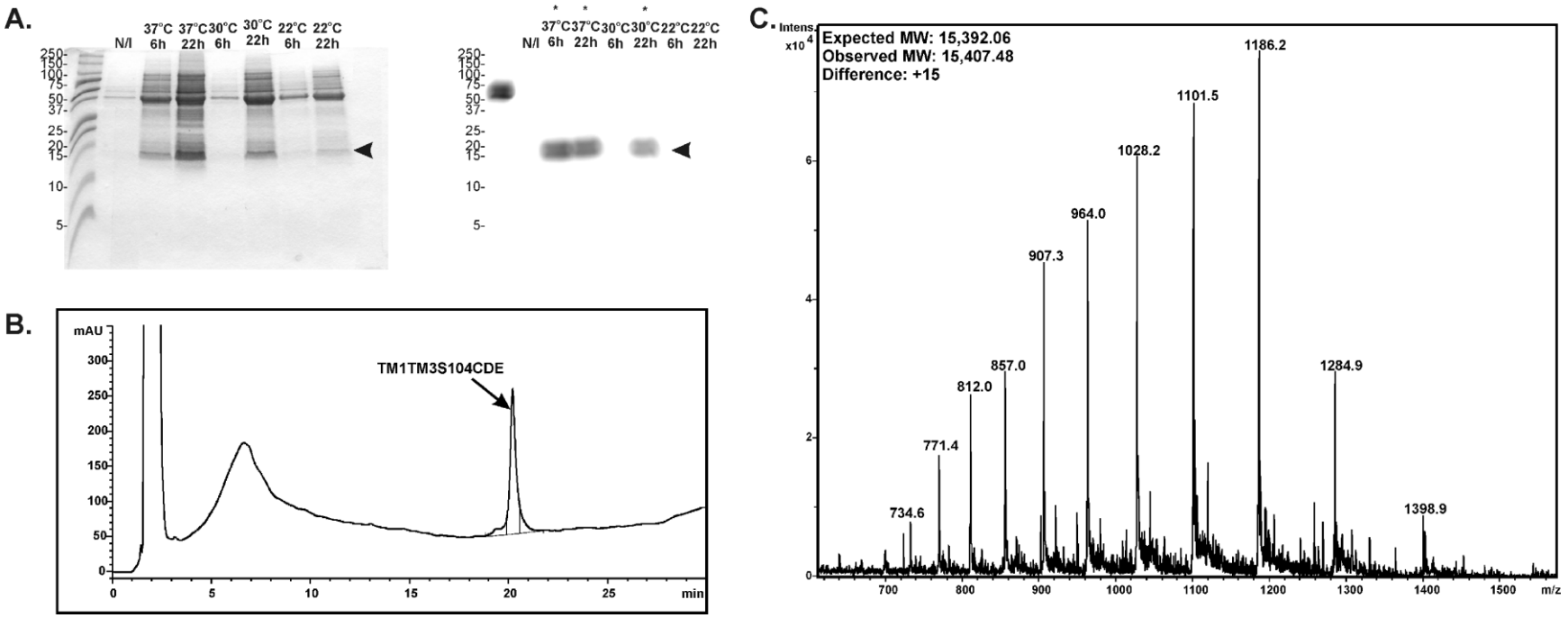

FIGURE 2-12 Optimization of direct expression of Ste2p TM1TM3 (G31-R161, S104C). The pKC02 plasmid was transformed into the BL21-DE3 expression strain. Expression was conducted at 37°C, 30°C, and 22°C, and samples were collected before and 6 and 22 h after induction. SDS-PAGE and Western Blot analysis using anti-His antibodies (A, left and right, respectively) revealed that the desired protein was produced at 37°C 6h and 22h post induction and at 30°C 22 h post induction. HPLC analysis of an IB pellet produced from the 37°C, 22h expression revealed that the desired product eluted from the column after injection of 70% TFA (B). ESI-MS analysis of the eluted polypeptide (D) revealed that the product had a molecule weight 15 Da greater than the expected molecular weight

SDS-PAGE and Western blot analysis of the large scale TM1-TM3S104C expressions revealed the same pattern as the optimization experiments. The desired band was not clearly evident in the Coomassie stained gel, and a blurry Western band was obtained (Figure 2-13A, inset). Preparative purification of a 1 mL IB pellet corresponding to one-eighth of the culture volume revealed that the expression level was low. Injection of 70% TFA did not produce a high intensity band as in the optimization experiments (data not shown). The remaining pellets were resuspended in water, and combined to produce pellets representing one-quarter of the culture volume. Injection of these larger pellets followed by injection of 70% TFA allowed for the purification of TM1-TM3S104C (Figure 2-13). ESI-MS analysis of fractions collected during the purification contained peaks with the desired mass to charge ratios (data not shown), but the intensity of these spectra was very low. UV quantitation of the combined fractions was unsuccessful, as an accurate amplitude of the absorbance peak could not be determined. After lyophilization, the total yield by weight for 2 L of culture volume was approximately 5 mg. This material was later used for MTSL attachment and PRE experiments.

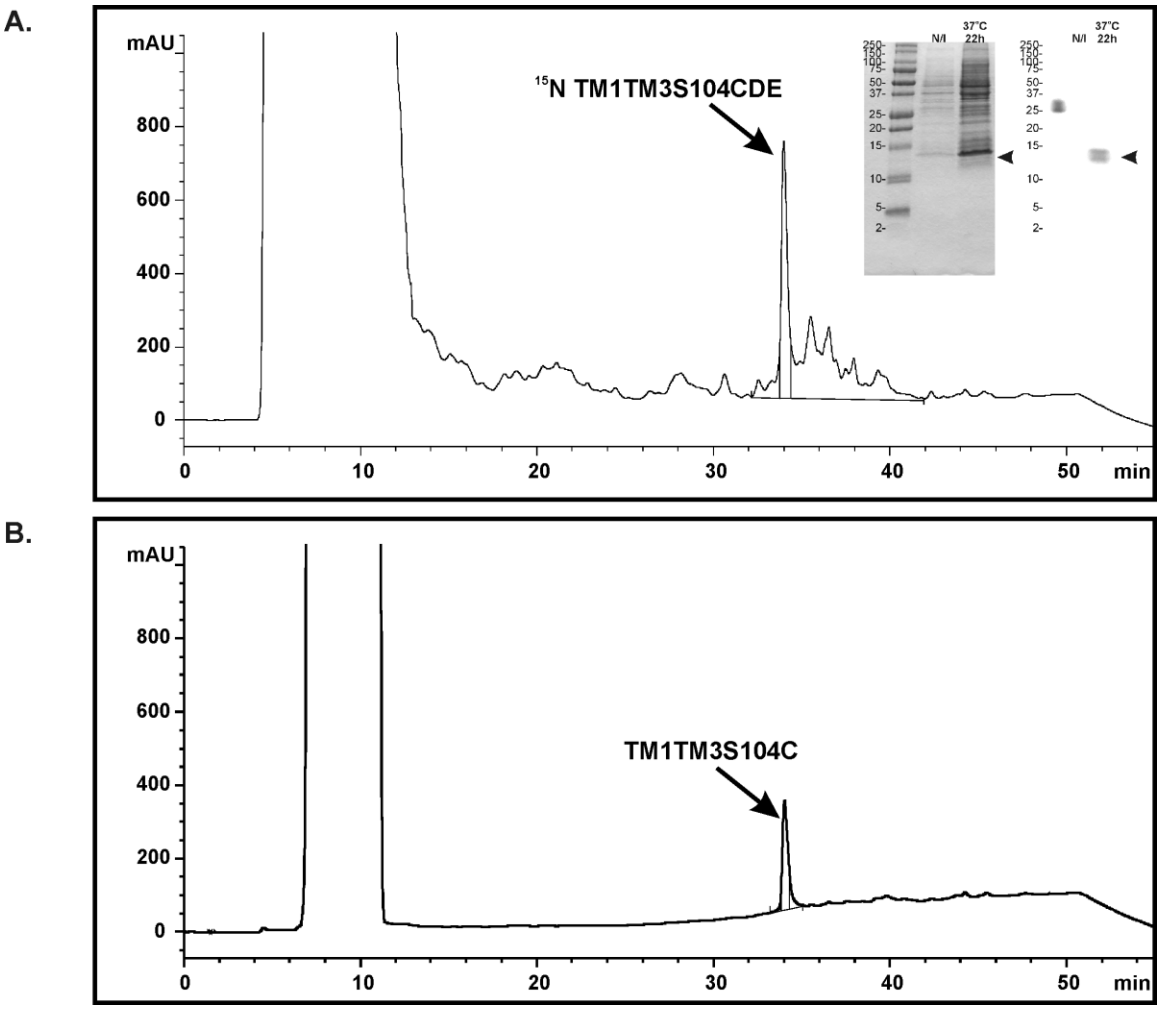

FIGURE 2-13 Large-sale expression of Ste2p TM1TM3 (G31-R161, S104C). The pKC02 plasmid was transformed into the BL21-DE3 expression strain and expressed in 2 1L cultures in M9 minimal medium prepared with $^{15}NH_4Cl$. Expression was conducted at 37°C for 22h. SDS-PAGE and Western Blot analysis using anti-His antibodies (A, inset, left and right, respectively) revealed that the desired protein was produced. Preparative HPLC purification of IB pellets corresponding to one-quarter of a total 1 L culture volume revealed that the desired product eluted from the column at approximately 34 minutes (A). Additional material was eluted from the column after injection of 70% TFA (B).

Optimization of direct expression of TM1 revealed that this construct was expressed in high yields in BL21-AI when expression was induced with 0.5% L-arabinose in M9 minimal medium at an OD_{600} of 0.4 and the cells were incubated at 37°C for 22 h. Large-scale expressions were conducted in 500 mL of M9 medium prepared with $^{15}NH_4Cl$ and with $^{15}NH_4Cl$ and [^{13}C]-glucose for NMR analysis. Purification of the resulting inclusion bodies yielded highly homogeneous products, as evidenced by the ^{15}N purification in Figure 2-14. ESI-MS analysis of the fractions collected during the [^{15}N] purifications revealed that the eluent had a molecular weight of 6087, which represented ~100% [^{15}N]-incorporation (Figure 2-14B). Similar analysis of the [^{15}N,^{13}C]-labeled material revealed that the eluent had a molecular weight of 6348 (data not shown). Assuming 100% [^{15}N]-incorporation, this represents 98% [^{13}C]-incorporation.

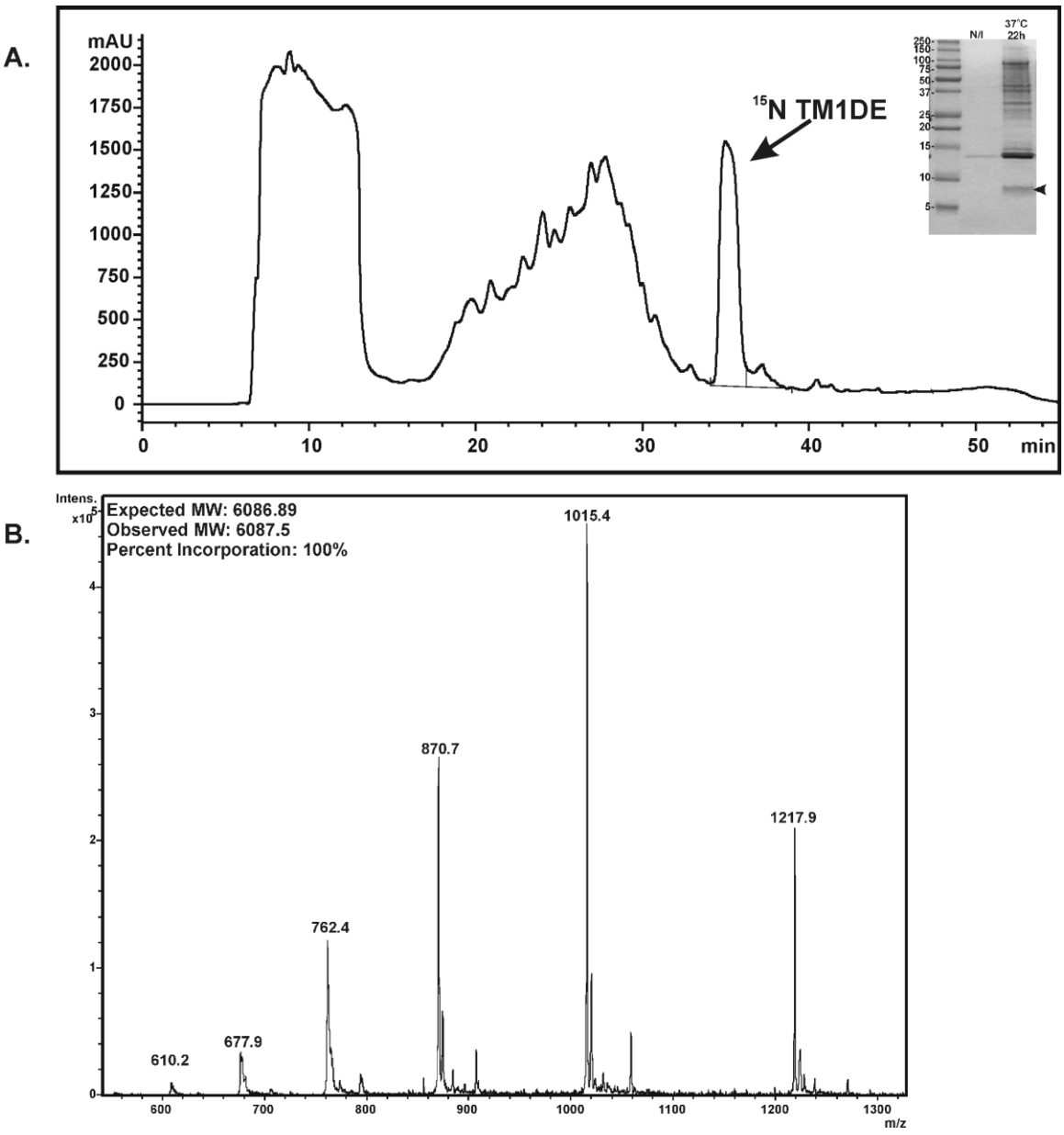

FIGURE 2-14 Large-sale expression of Ste2p TM1 (G31-T78). The pKC03 plasmid was transformed into the BL21-AI expression strain and expressed a 500 mL culture of minimal medium prepared with $^{15}NH_4Cl$. Expression was conducted at 37°C for 22h. SDS-PAGE (A, inset) revealed that the desired protein was produced. Preparative HPLC purification of IB pellets corresponding to one-fourth of the total 500 mL culture volume revealed that the desired product eluted from the column at approximately 34 minutes (A). ESI-MS analysis of the eluted product revealed that the polypeptide had the desired molecular weight with 100% [^{15}N]-incorporation (B).

Conclusions

Fragments encompassing the first three TMs and the first, second, and seventh TMs of Ste2p have been chosen for biophysical analysis and structure determination under the assumption that large GPCR fragments will form more defined tertiary contacts than previously studied smaller fragments. Ste2pTM1-TM3(G31-R161) and TM127(G31-T114, T274-L340) have been cloned as a Met deficient TrpΔLE fusion proteins, and expression has been optimized in *E.coli*. Conditions for CNBr cleavage and purification by RP-HPLC have been established, and multimilligram quantities of TM1-TM3 and TM127 have been obtained. Our yields of more than 10 mg/L of purified product are a vast improvement over previous studies of biosynthesized GPCR fragments where final yields were no more the 3 mg/ L (*70, 84, 85*). Furthermore, these studies were conducted on fragments containing no more than 2 TMs. Production of isotopically labeled fragments of this size in such large yields is a significant contribution to this field.

Direct expression of the TM1-TM3S104C and TM1 constructs was also accomplished. Direct expression of the 3TM construct containing Cys resulted in yields comparable to those from previous studies in the literature described above. Direct expression of the TM1 construct resulted in high yields of approximately 11 mg/L of culture volume. One troubling aspect of the TM1-TM3S104C expression was that the results were not reproducible, and the purified polypeptide is unstable. Sufficient material for one PRE experiment was obtained, and I have not been able to produce additional material to date. The TM1 polypeptide is extremely stable and the results of the expressions are highly reproducible. This difference would suggest that the TM1-TM3S104C construct is hindered either by the presence of the single Cys residue or by

the presence of the aggregation prone first extracellular loop. I have recently found that direct expression of TM1-TM3 from the natural start codon is reproducible and results in higher yields. I am currently working to produce more material for PRE experiments using this system.

The polypeptides described in this chapter were generated in a number of isotopic labeling schemes required for NMR analysis. Percent isotope incorporations for ^{15}N and ^{13}C were found to be 95% or greater in all cases, which is considered more than acceptable for NMR analysis. Partially deuterated polypeptides were desired for the large 3TM constructs as deuteration affects relaxation and can result in spectra with sharper lines. We were able to control the percent of [^{2}H] incorporation by controlling the amount of [^{2}H] present in the M9. Use of nondeuterated glucose and D_2O resulted in [^{2}H] incorporations of ~55%. Use of deuterated glucose and D_2O resulted in [^{2}H] incorporations of ~90%. Such high isotope incorporations and the ability to control the percent incorporation by media manipulation suggests robust expression conditions for all polypeptides studied.

CHAPTER 3

Biophysical Characterization of Ste2pTM1-TM3 (G31-R161) and TM127 (G31-T114, L274-L240) Using Circular Dichroism and Heteronuclear NMR

Introduction

Biophysical characterization of Ste2pTM1-TM3 (G31-R161) and Ste2pTM127 (G31-T114, T274-L340) was undertaken in order to determine the secondary structure of the fragments and NMR sample conditions suitable for structure determination. As TM1-TM3 and TM127 are membrane protein fragments, all studies conducted on these constructs required a membrane mimetic medium. Traditionally, membrane proteins have been studied in either organic:aqueous solvent mixtures or in detergent micelles. Organic:aqueous solvent mixtures mimic the low dielectric constant of the cell membrane. Detergent micelles mimic both the dielectric constant and the physical make up of the cell membrane. Previous studies on fragments of Ste2p had been conducted in mixtures of trifluroethanol:water (TFE:water) as well as various detergent micelles (*1, 55-58, 60, 65, 86*). Therefore, initial biophysical characterization of TM1-TM3 was conducted in TFE:water and detergent micelles.

The first attempts at characterization of TM1-TM3 and TM127 were conducted using circular dichroism (CD) spectroscopy. CD spectroscopy is an excellent method for initial characterization of the secondary structure of polypeptides since it is inexpensive and efficient. Very little material is needed in order to collect data, and the experiments are not time consuming. Analysis of the shape of a CD spectrum can quickly yield a qualitative elucidation of the secondary structure of the protein. For helical proteins such as TM1-TM3 and TM127, the mean residue ellipticity at 222 nm can be used to quantitatively determine the fraction helicity (f_H) of a protein of interest (*67, 87*).

Previous studies on Ste2p fragments have shown that the quality of the CD spectrum and the determined f_H in a given membrane mimetic can be used as a guide for NMR sample conditions (*56*). Our laboratory has suggested that samples of a membrane protein that produce CD spectra indicative of helical secondary structures are predictive of solvent conditions that are a reasonable starting point for initial NMR characterization(*56*). In contrast conditions that yield CD spectra for a membrane protein fragment indicative of a high content of β-sheet structures are indicative of aggregation and should be avoided. Therefore, CD spectroscopy was used to determine the secondary structure of my 3TM constructs in various membrane mimetics and as a guide to determining sample conditions that would be suitable for NMR analysis of TM1-TM3 and TM127.

Once conditions had been identified by CD as candidates for NMR, the samples were scaled up to concentrations suitable for [^{15}N, ^{1}H]-HSQC (*88*) analysis. [^{15}N, ^{1}H]-HSQC spectra that exhibit a majority of the expected cross peaks, well resolved crosspeaks, and wide peak dispersion in the [^{1}H]-dimension are generally considered to represent suitable conditions for NMR backbone assignment and 3D structure determination. Assignment of backbone resonances for large GPCR fragments is not a trivial task as the fragments are mostly α-helical and are composed of aliphatic residues, characteristics which cause significant spectral overlap. Consequently, peak visibility and separation are crucial for complete backbone assignment. Furthermore, as a large [^{1}H] dispersion is indicative of protein folding into a tertiary structure, spectra with the widest dispersions should represent conditions where the protein of interest adopts the highest degree of structure. Therefore, I sought to identify sample conditions that

produced [^{15}N, ^{1}H]-HSQC spectra with limited spectral overlap and wide dispersion in the [^{1}H]-dimension.

CD analysis of TM1-TM3 and TM127 produced several solvent conditions that were candidates for NMR sample preparation. [^{15}N,^{1}H]-HSQC analysis under these conditions led to the identification of two membrane mimetic environments for TM1-TM3 and one membrane mimetic environment for TM127 that were suitable for NMR structure determination. This chapter describes the detailed biophysical analysis conducted in order to determine conditions for structure determination.

Materials and Methods

Circular Dichroism Spectroscopy of Ste2pTM1-TM3(G31-R161) and Ste2pTM127(G31-T114, T247-L340)

CD spectroscopy was used to determine the secondary structure of the TM1-TM3 and TM127 polypeptides. Experiments were carried out in both organic-aqueous and micellar environments. Purified TM1-TM3 and TM127 were solubilized in 50% TFE:water. The concentration of the peptide stock solutions were determined by UV absorbance spectroscopy. The absorbance at 280 nm and a molar extinction coefficient of 12,950 M^{-1} cm^{-1} for TM1TM3 and 18,450 M^{-1} cm^{-1} for TM127 (calculated using the equation in Chapter 2) were used to determine the concentration by Beer's Law. This stock solution was portioned and lyophilized so that 20 μM samples could be prepared in the desired volume.

For organic:aqueous investigations, the lyophilized peptide was solubilized in 500 μL of 25, 50, 75 or 95% TFE. All TFE mixtures used were volume to volume. The desired volume of TFE was added to the lyophilized peptide, and the sample was sonicated briefly until it was homogeneous. After sonication, the desired volume of

water was added, and the sample was sonicated again. For detergent micelle investigations, 20 mM stock solutions of various detergents (lysomyristoylphosphatidylglycerol (LMPG), lysoplamitoylphosphatidylcholine (LPPC), lysopalmitoylphosphatidylglycerol (LPPG) lysostearoylphosphatiylglycerol (LSPG), dodecyl maltoside (DDM), 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC), dodecylphospholcholine (DPC), hexadecylphosphocholine (HDPC), sodium dodecylsulfate (SDS), and perfluorooctanoate (PFO)) were prepared in 20 mM sodium phosphate buffer with a pH of 5.0. Four hundred microliters of the buffer were added to an appropriate weight of each detergent. The mixture was sonicated at 50°C for 15 min at ~50 W using a S3000 sonicator (Farmingdale, NY) with a cup horn. Two hundred microliters of each detergent stock solution were added to a lyophilized portion of peptide, and the detergent-peptide sample was sonicated again under the same conditions. The remaining 200 μL were used as a blank.

Additional experiments were conducted to determine the optimal buffer pH and composition for TM1-TM3. For pH investigations, TM1-TM3 was solubilized in LMPG micelles prepared in sodium phosphate buffer at pH 4.5, 5.5, 6.5, 7.5, and 8.5 as described above. For buffer composition investigations, TM1-TM3 was solubilized in LMPG micelles prepared in 20 mM sodium phosphate with and without 100 mM sodium chloride, 20 mM 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid (HEPES) buffer with and without sodium chloride, 20 mM 3-(N-morpholino)propanesulfonic acid (MOPS) buffer with and without sodium chloride, or 20 mM Tris buffer with and without sodium chloride at pH of 5.6.

CD investigations for TM1-TM3 were also conducted in LMPG mixed micelles

containing HDPC, DPC, or DDM with varying LMPG. LMPG was solubilized in sodium phosphate buffer as described above. This solubilized detergent solution was then added to DHPC, DPC, or DDM and the mixture was sonicated to yield a solution with a final total detergent concentration of 20 mM. The sonicated mixture was then added to lyophilized peptide and sonicated.

The spectra were recorded on an Aviv model 410 CD instrument (Aviv Biomedical, Lakewood, NJ). For the organic-aqueous studies, a quartz cuvette with a pathlength of 1 mm was used. For the detergent micelle studies, a quartz cuvette with a pathlength of 0.2 mm was used. The spectra were collected over a wavelength range of 260 nm to 185 nm in increments of 1 nm. An average of three scans was taken in all cases, and the blank spectra were subtracted from the spectra containing the peptide. CD spectra of the solvent were run at the same time. The CD of the peptide was determined after subtracting the signal of the blank from the signal of the polypeptide sample under identical conditions. The circular dichroism was then expressed as mean residue molar ellipticity (deg cm^2 $dmol^{-1}$). Percent helicity was determined using the following formula, where f_H is the helical content, θ_{222} is the mean residue ellipticity at 222 nm, and θ^0_{222} and θ^{100}_{222} are the mean residue ellipticities for 0% and 100% helicity and are estimated to be 2,000 and 30,000 deg cm^2/dmol, respectively (*67, 87*).

$$f_H = ([\theta]_{222} - [\theta]^0_{222}) / ([\theta]^{100}_{222} - [\theta]^0_{222})$$

Sample Preparation for NMR Spectroscopy of Ste2p TM1–TM3 (G31–R161)

For NMR investigations in organic-aqueous media, lyophilized peptide was solubilized in 175 μL of TFE and sonicated at room temperature for 15 min at ~50W. After sonication, 175 μL of water containing 0.1% TFA were added slowly and the

sample was sonicated again. The sample was transferred to a Shigemi NMR tube (Shigemi, Allison Park, PA), and [^{15}N, ^{1}H]-HSQC analysis was performed at 25°C, 35°C, and 45°C in order to determine the optimal temperature for backbone assignment and structure determination. The sample concentrations were determined by UV-quantitation, as described above, and were generally 0.1-0.5 mM.

For investigations in detergent micelles, various detergents were weighed out such that the final concentration of detergent in 350 μL would be approximately 100 mM. The detergent was solubilized in sodium phosphate (NaH_2PO_4) buffer, pH 5.6, prepared with 10% D_2O and sonicated at 50°C for 15 min at ~50W. The sonicated solution was added to solid TM1-TM3 that had been lyophilized two times: first from acetonitrile/water/2-propanol following HPLC isolation and finally from trifluoroethanol:water. The resulting detergent-peptide solution was sonicated as described above. Five microliters each of 200 mM NaN_3 and 10 mM 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS) were added to the peptide detergent solution, and the sample was transferred to a Shigemi NMR tube. Sample preparation was optimized to increase sample stability by altering the sample buffer. Samples were prepared in 20 mM sodium phosphate, 20 mM sodium phosphate/100 mM sodium chloride, 20 mM HEPES, and 20 mM HEPES/100 mM sodium chloride. [^{15}N,^{1}H]-HSQC experiments were conducted at 45°C on a three-channel Varian NMR-S 600MHz NMR spectrometer (Varian NMR Instrument, Palo Alto, CA) with a z-axis pulsed-field-gradient and a Varian 5-mm [^{1}H, ^{15}N, ^{13}C, ^{2}D] cryo-probe. Sample quality was assessed by counting the peaks resolved in the HSQC spectrum and following the stability of the spectrum with time both at room temperature and 45°C.

Results

Characterization of the Secondary Structure of Ste2p TM1-TM3 (G31-R161)

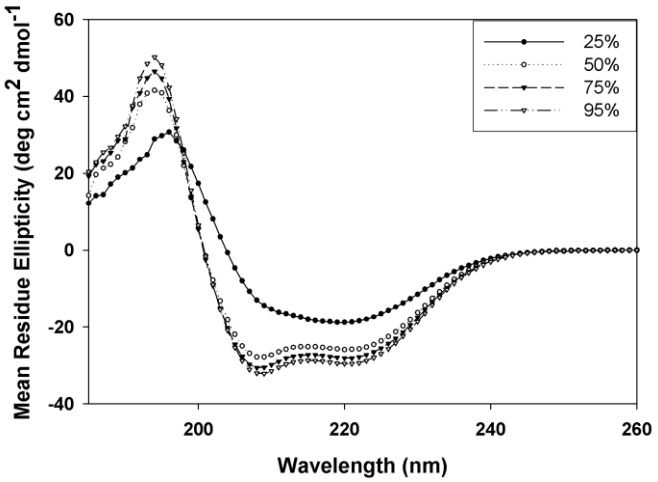

FIGURE 3-1: Circlar Dichroism analysis of TM1-TM3 in an organic:aqueous membrane mimetic. 20 μM samples of TM1-TM3 in 25, 50, 75, and 95% TFE:water were prepared as described in materials and methods. The presented spectra represent a solvent-subtracted average of three scans.

CD spectroscopy was used to determine whether the TM1-TM3 fragment adopted a stable helical structure in different membrane mimetic environments. The analysis was conducted in organic-aqueous solvents and in detergent micelles. In TFE, the CD spectra of the fragment exhibited a maximum at 192 nm and two minima at 208 nm and 222 nm, which are indicative of a helical peptide (Figure 3-1). The percent helicity was directly proportional to the concentration of TFE, with mean residue molar ellipticity at 222 nm increasing from -18,000 deg cm^2 $dmol^{-1}$ in 25% TFE:water to -29,000 deg cm^2 $dmol^{-1}$ in 95% TFE:water. The helicity observed in 25% TFE:water corresponded to ~60% helicity, whereas in 95% TFE:water the peptide was nearly 100% helical (Table 3-1). Moreover, the shape of the CD spectrum in 25% TFE/75% water was typical of a partially aggregated peptide with a mixture of secondary structures.

TABLE 3-1: Circular Dichroism of Ste2p TM1-TM3(G31-R161)

	Composition	f_H
TFE:water		
	25%	0.58
	50%	0.85
	75%	0.92
	95%	0.98
Detergent Micelles		
	LMPG	0.49
	LPPC	0.22
	LPPG	0.51
	LSPG	0.49
	DDM	0
	DHPC	0.02
	DPC	0.32
	HDPC	0.43
	SDS	0.43
LMPG:HDPC Mixed Micelles		
	LPMG	0.48
	LMPG:HDPC (1:1)	0.47
	LMPG:HDPC (2:1)	0.45
	LMPG:HDPC (4:1)	0.42
	LMPG:HDPC(1:2)	0.48
	HDPC	0.50
LMPG:DPC/DDM Mixed Micelles		
	LMPG	0.31
	LMPG:DPC (1:1)	0.17
	LMPG:DPC (1:4)	0.06
	LMPG:DPC (4:1)	0.39
	DPC	0.07
	LMPG:DDM (1:1)	0.18
	LMPG:DDM (1:4)	0
	LMPG:DDM (4:1)	0.34
	DDM	0
LMPG Buffer Test[1]		
	pH 4.5	0.32
	pH 5.5	0.29
	pH 6.5	0.24
	pH 7.5	0.24
	pH 8.5	0.23
	Sodium Phosphate	0.38
	Sodium Phosphate+	0.35
	HEPES	0.34
	HEPES+	0.38
	MOPS	0.32
	MOPS+	0.41
	Tris	0.30
	Tris+	0.29

[1]Buffers marked with + contained 100 mM sodium chloride.

Line shapes similar to those observed in TFE/water were obtained for CD experiments on TM1-TM3 conducted in LMPG, LPPC, LPPG, LSPG, DDM, DHPC, DPC, HDPC, and SDS micelles prepared in sodium phosphate buffer at pH 5.6, although the minima at 208 and 222 nm were not as well defined (Figure 3-2). The mean residue ellipticity ranged from -16,000 deg cm^2 $dmol^{-1}$ in LPPG (Figure 3-2A) micelles to -700 deg cm^2 $dmol^{-1}$ in DDM micelles (Figure 3-2B) with fraction helicity ranging from approximately 50% in LPPG micelles to 0% in DDM micelles (Table 3-1). The differences in fraction helicities observed in the tested micelles are due to the solubility of the TM1-TM3 polypeptide, and although a wide range of fraction helicities is observed, a general trend is evident. TM1-TM3 shows little to no solubility and in turn low helicity in micelles with short aliphatic chains (≤12 carbon atoms). The highest calculated helicities were for the lysophosphatidylglycerol samples, which all exhibited helicities of ~50% (Table 3-1). SDS showed a fraction helicity of 43% and was the only short chain detergent that compared to the lysophosphatidylglycerol samples. Although LPPG showed the highest fraction helicity, TM1-TM3 showed limited solubility in this detergent at higher concentrations. For this reason, all further sample optimization experiments were done using LMPG, which had the next highest f_H.

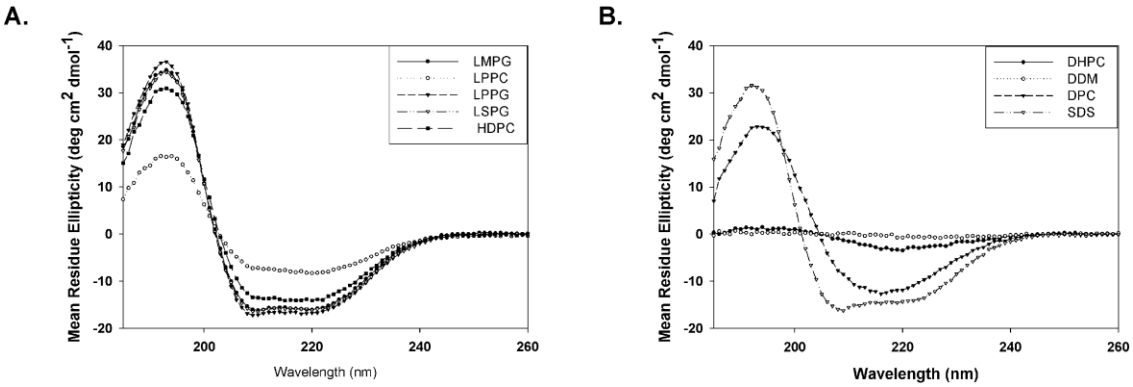

FIGURE 3-2: Circular Dichrosim analysis of TM1-TM3 in detergent micelles. Lyophilized TM1-TM3 was solubilized in LMPG, LMPC, LPPG, LSPG, HDPC (A), or DHPC, DDM, DPC, SDS (B) micelles prepared in 20 mM sodium phosphate buffer, pH 5.6. The final peptide concentration was 20 μM. The presented spectra represent a solvent-subtracted average of three scans.

Mixed micelles are better mimics of the heterogeneous nature of the native biological membrane. Therefore, experiments were conducted in LMPG mixed micelles prepared with varying ratios of HDPC, DPC, or DDM (Figure 3-3). TM1-TM3 was soluble in LMPG:HDPC mixed micelles of all compositions. These micelles showed high f_H values of 40-50% with reasonable line shapes (Figure 3-3A). Solubility was diminished in LMPG:DPC and LMPG:DDM mixed micelles (Figure 3-3B and C, respectively), with almost no solubility observed when the short chain component was the dominant species. When LMPG was the main component, f_H values greater than 30% were obtained (Table 3-1). Given the fact that it is very difficult to determine the concentration of the peptide with a high degree of certitude, we believe that peak shape together with the CD intensity are the most important indicators of the quality of the sample. From this perspective, for example, (LMPG:DDM 4:1) seems equivalent to pure LMPG (Figure 3-3C).

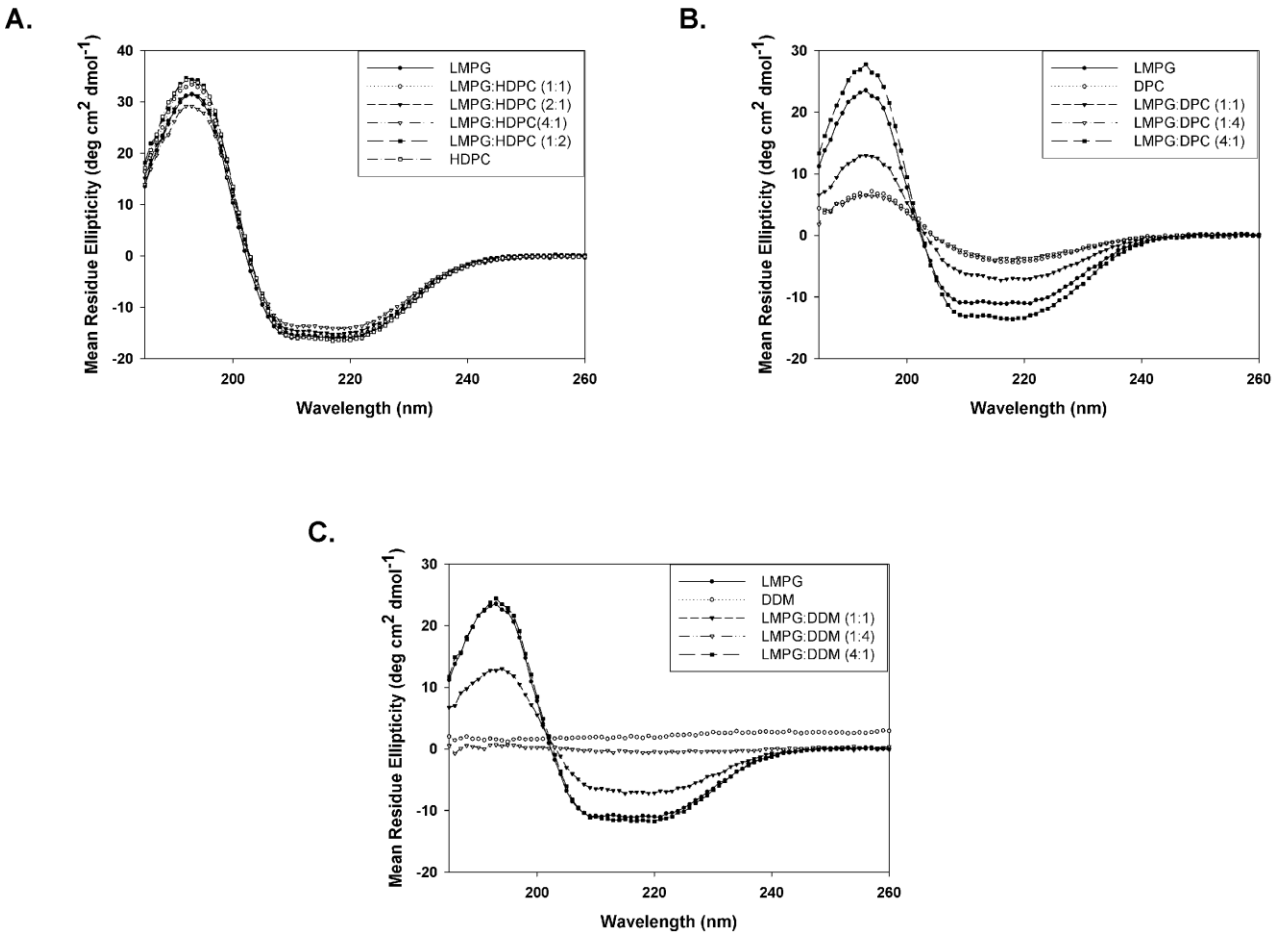

FIGURE 3-3: Circular Dichroism analysis of TM1-TM3 in LMPG:HDPC (A), LMPG:DPC (B), and LMPG:DDM (C) mixed micelles. Lyophilized TM1-TM3 was solubilized in 20 mM sodium phosphate buffer, pH 5.6 to create 20 μM peptide samples (see materials and methods). The presented spectra represent a solvent-subtracted average of three scans.

Upon analysis of all of the CD spectra, it was decided that the optimal detergent micelle for TM1-TM3 was composed of LMPG as the polypeptide showed reasonable solubility in this detergent, and the CD spectra showed acceptable lineshapes and a high fraction helicity. Further optimization of the buffer conditions for micelle preparation were conducted in order to determine the optimal pH and buffering agent (Figure 3-4). The f_H values were found to increase with decreasing pH, with pH 4.5 exhibiting the highest value of 30% (Figure 3-4A, Table 3-1). As this value is approximately 1000 fold more acidic than physiological pH, this would not be an ideal pH for NMR studies. For this reason, pH 5.5 was chosen as the pH for NMR investigation in LMPG micelles as this was the highest pH, which produced a f_H of 30%. The fraction helicity did not change drastically with buffer composition, with f_H values between 30-40%, with the highest value obtained for micelles prepared in MOPS buffer plus sodium chloride (Figure 3-4B, Table 3-1). As low conductivity buffer conditions were desired and sodium chloride has a high conductivity, any buffer prepared with the addition of sodium chloride was not chosen for the final buffer conditions for NMR investigations. HEPES buffer produced the highest f_H without the addition of salt and was chosen for further investigation.

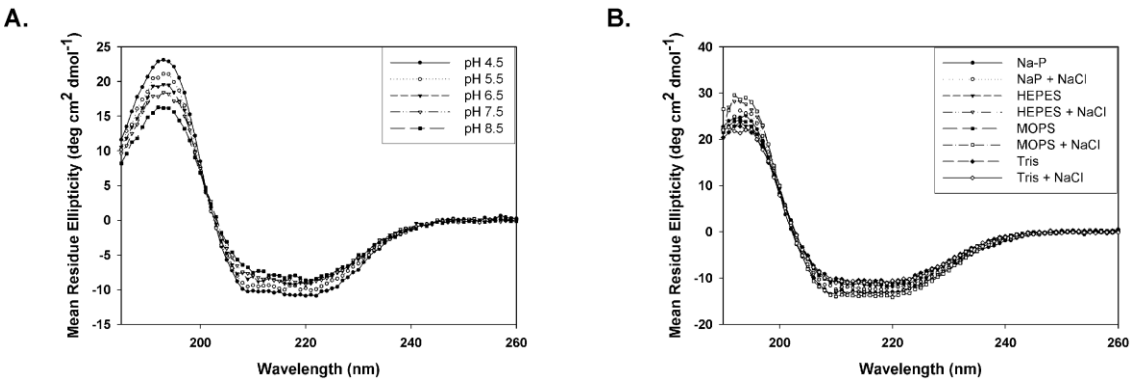

FIGURE 3-4: Optimization of buffer composition for biophysical analysis of TM1-TM3. Circular Dichroism analysis was conducted in LMPG micelles prepared in sodium phosphate buffer at pH 4.5, 5.5, 6.5, 7.5, and 8.5 (A) and in sodium phosphate, HEPES, MOPS, and Tris with and without the addition of sodium chloride at pH 5.6 in order to determine the optimal buffer conditions for NMR analysis. The presented spectra represent a solvent-subtracted average of three scans.

In TFE:water, the fragment exhibited approximately 50%-90% helicity with increasing concentrations of TFE. Considering the helix-connecting loops and the amino terminus, out of the 131 residues in Ste2pTM1-TM3 at least 40-50 are in non-TM regions. Therefore, it is highly unlikely that samples with CD helicities exceeding 80% would represent the native state of this fragment (Figure 3-1A). The 40%-50% helicities determined in detergent micelles are much more likely to match the secondary structure content in a biologically relevant state of this fragment. Although the helicity measured in 25% TFE:water is closest to those observed in detergent micelles because of the shape of the spectrum and the likelihood that the sample solubility was low, and it therefore would be difficult to achieve concentrations necessary for NMR experiments, 50% TFE:water was chosen for NMR analysis. This represented the lowest concentration of TFE where distinct minima at 208 nm and 222 nm where observed in the CD spectrum.

Characterization of the Secondary Structure of Ste2p TM127 (G31-T114, T274-L340)

CD investigations for TM127 were also conducted both in TFE:water mixtures and in various detergent micelles. Surprisingly, TM127 showed limited solubility in

TFE:water mixtures and was only completely soluble with TFE concentrations of 75% or higher. CD spectra in 75 and 95% TFE:water (Figure 3-5A) were consistent with a highly α-helical polypeptide with f_H values larger than 80% (Table 3-2). As TM127 contains 30 residues of the N-terminal tail and 40 residues of the C-terminal tail, f_H values this large would not represent a native-like state for this fragment. It has been shown that portions of the N- and C-terminal tails (*55, 58*) and of EL1 (*78*) have a tendency to assume helical structure. Even if these structured portions of the disordered regions were included in the predicted helicity calculation, CD determined f_H values above 70% would still not be representative of a native-like state. Accordingly, we decided not to carry out further structural analysis on TM127 in TFE/water as membrane mimetic environment.

Analysis of TM127 in detergent micelles prepared in sodium phosphate buffer, pH 6.5, produced similar results as those obtained for the TM1-TM3 fragment (Figure 3-5B). The fragment showed limited solubility in the short chain detergents DDM, DPC and PFO. Long chain detergents with glycerol head groups were preferred, with mean residue ellipticity at 222 nm increasing from -13,000 to -15,000 deg cm^2 $dmol^{-1}$as a function of the chain length. The detergent which showed the largest f_H (0.46) was LSPG (Table 3-2), and all further investigations of this fragment were conducted in this detergent.

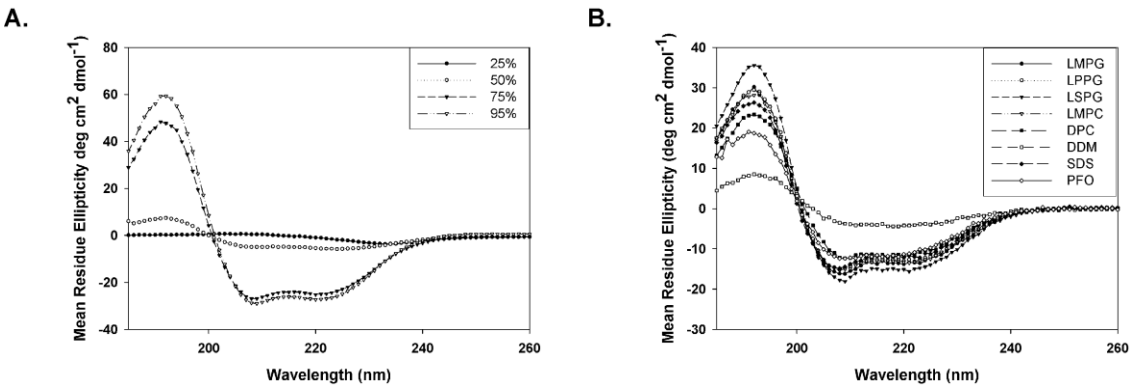

FIGURE 3-5: Circular Dichroism analysis of TM127 in organic:aqueous and detergent micelle membrane mimetics. Lyophilized aliquots of TM127 were solubilized in either 25, 50, 75, and 95% TFE:water (A) or LMPG, LPPG, LSPG, LMPC, DPC, DDM, SDS, and PFO detergent micelles (B) as described in materials and methods. The final peptide concentration in all cases was 20 μM. The presented spectra represent a solvent-subtracted average of three scans.

TABLE 3-2: Circular Dichrosim of Ste2p TM127(G31-T114,T274-L340)

	Composition	f_H
TFE:water		
	25%	0
	50%	0.13
	75%	0.81
	95%	0.89
Detergent Micelles		
	LMPG	0.40
	LPPG	0.40
	LSPG	0.46
	LMPC	0.40
	DPC	0.33
	DDM	0.08
	SDS	0.36
	PFO	0.32

Optimization of NMR Sample Conditions for Ste2p TM1-TM3 (G31-R161)

NMR samples for TM1-TM3 were prepared in aqueous TFE (1:1) as described in the Materials and Methods section. Upon slow addition of water to the peptide in TFE and sonication, a clear solution was produced. [^{15}N, ^{1}H]-HSQC spectra were measured on a 0.18 mM [^{15}N]-TM1-TM3 sample at 25°C, 35°C, and 45°C (see Figure 3-6). Many of the resolved resonances at 25°C were of low intensity, suggesting that correlations originating from these atoms would not be observed in 3D experiments. In contrast the peak intensities were more uniform in TFE/water (1:1) at 35°C, and 45°C. The best peak

dispersion and narrowest peak widths were obtained at 45°C, where 122 out of 128 expected peaks were observed (Figure 3-6C).

NMR sample preparation in detergent micelles was complicated by the low peptide solubility in water and the limited stability of the peptide-detergent complex. As the NMR investigations of the TM1-TM2 fragment were conducted in LPPG micelles and since TM1-TM3 exhibited the highest helicity in this detergent as judged by CD, it was chosen for initial sample preparations. Although the 20 μM TM1-TM3 LPPG CD sample was clear, several sample preparations at 0.1 mM peptide concentrations led to turbid solutions. In contrast, a 0.15 mM sample of TM1-TM3 in LMPG micelles in 20 mM sodium phosphate, pH 5.6, produced a clear solution after sonication. The [^{15}N, ^{1}H]-HSQC spectrum at 45°C displayed many of the peaks expected for the GPCR fragment (Figure 3-7A). However, a second HSQC spectrum measured after keeping the sample five days at room temperature revealed that a significant number of cross-peaks had been lost (Figure 3-7C). Moreover, the resolution of the spectrum was much poorer than that of Ste2p TM1-TM3 in organic aqueous medium (Figure 3-6). Addition of 100 mM sodium chloride resulted in retention of most of the HSQC cross-peaks even after incubation at 45°C for five days (Compare Figures 3-7B and 3-7D) indicating that salt stabilizes the polypeptide/micelle complex. However, the spectral dispersion was small compared to that in TFE:water under these conditions.

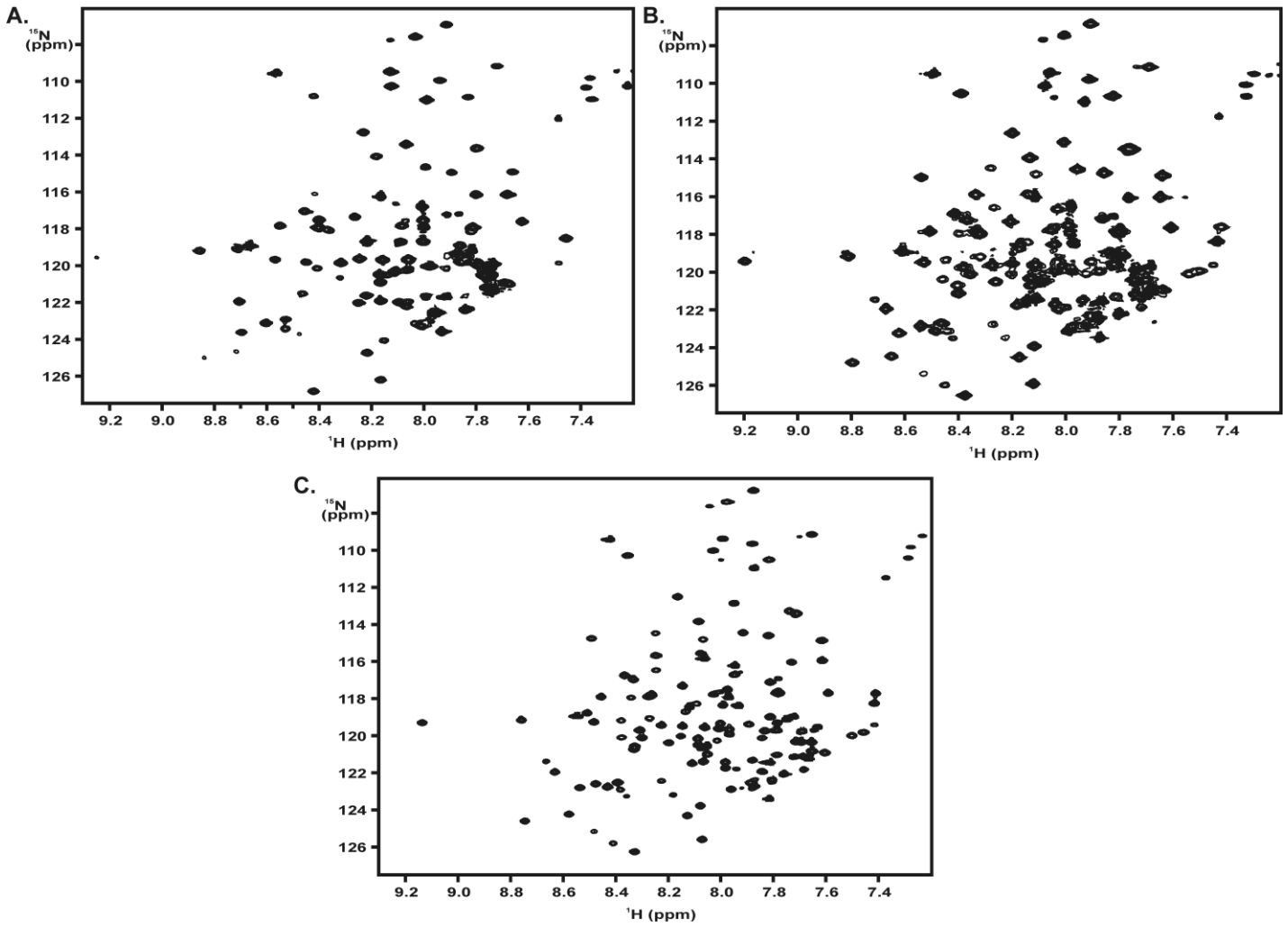

FIGURE 3-6 Determination of optimal temperature for NMR analysis of TM1-TM3. [^{15}N, ^{1}H]-HSQC spectra of Ste2p TM1-TM3 (G31-R161) in 50% TFE:water at 25°C (A), 35°C (B), and 45°C (C). The peptide concentration was 0.18 mM.

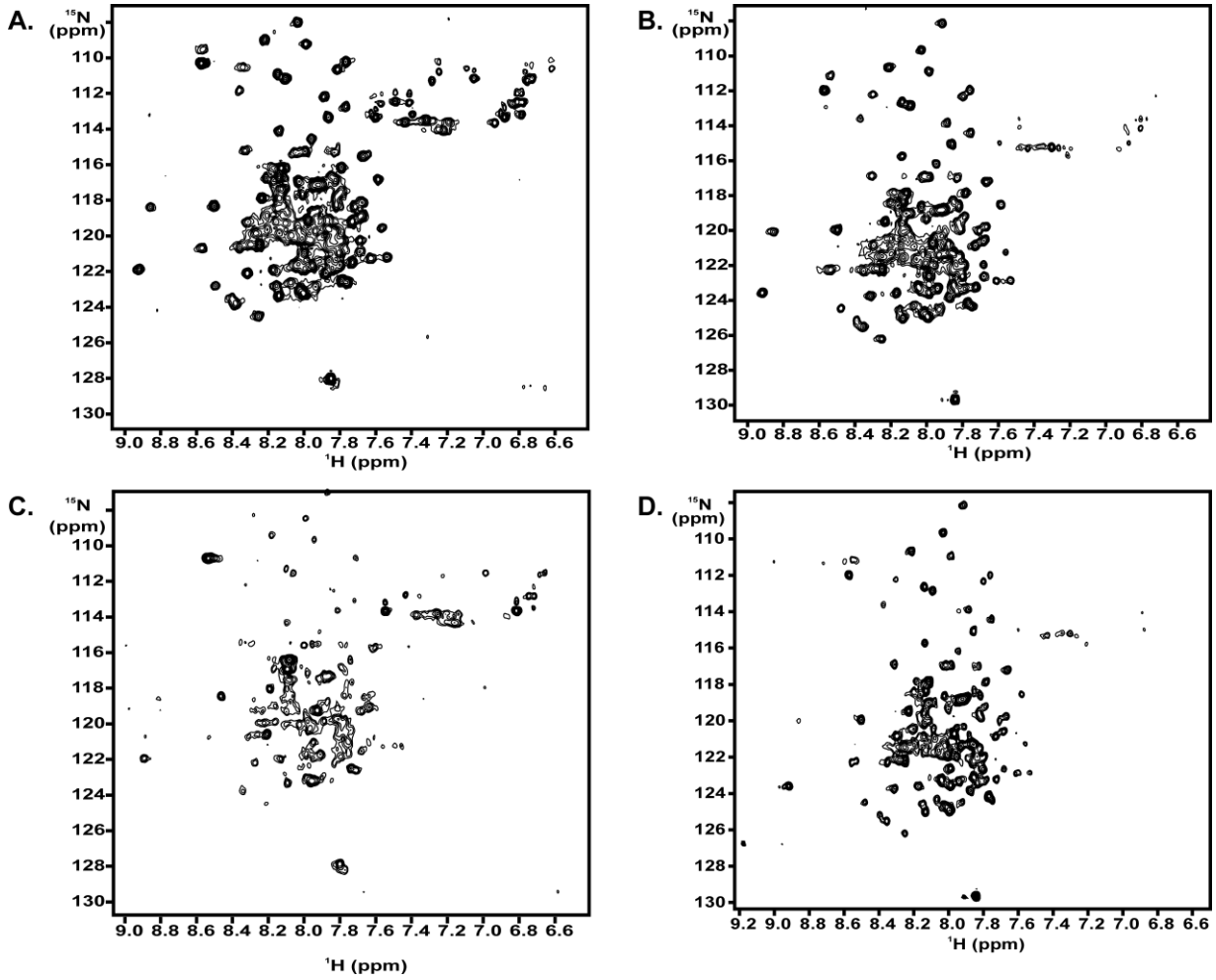

FIGURE 3-7 Effect of salt on the stability of TM1-TM3 in LMPG micelles. [^{15}N, ^{1}H]-HSQC analysis of Ste2p TM1-TM3 (G31-R161) peptide in LMPG micelles prepared in sodium phosphate buffer. 0.15 mM uniformly [^{15}N]-labeled TM1-TM3 peptide was solubilized in 20 mM LMPG micelles prepared in 20 mM sodium phosphate buffer, pH 5.6, (A and C) and 20 mM sodium phosphate 100 mM sodium chloride buffer, pH 5.6 (B and D). Spectra were collected for freshly prepared samples (A and B) and after five days of incubation at 45°C (C and D).

As solvent conductivity has been linked to NMR sensitivity (*89*), attempts were made to increase the number of observable peaks using a lower conductivity buffer. CD investigations revealed that TM1-TM3 was soluble and helical in a number of low conductivity buffers, with HEPES buffer giving the best results (Table 3-1). Sodium phosphate has a conductivity of 11 mS/cm and a sensitivity factor of 0.35 while HEPES has a conductivity of 0.06 mS/cm and a sensitivity factor of 0.70 (*89*). 0.15 mM [^{15}N]-TM1-TM3 samples were prepared in LMPG micelles in HEPES and HEPES with sodium chloride, pH 5.6. [^{15}N, ^{1}H]-HSQC spectra in both buffer conditions displayed peak dispersion and separation similar to those of the sodium phosphate samples. Spectra recorded after 5 days of incubation at 45°C revealed that the HEPES samples were very stable and the addition of sodium chloride did not significantly affect sample stability (data not shown).

To assess the suitability of the LMPG/HEPES sample conditions for multidimensional experiments, a new sample with a peptide concentration of 0.43 mM was prepared and subjected to [^{15}N, ^{1}H]-HSQC analysis. After sonication, the sample was completely clear and produced a high quality [^{15}N, ^{1}H]-HSQC spectrum after 21 min (Figure 3-8A). Sample incubation at 45°C for 5 days did not cause peptide precipitation and the spectrum showed at least comparable peak separation (Figure 3-8B). Further incubation revealed that the sample was stable even after 12 days at 45°C. This latter spectrum maintained the original observable peaks, however, a significant loss of peak intensity was observed (Figure 3-8C). Based on these findings micelles were determined to be suitable to determine a high-resolution solution structure of TM1-TM3.

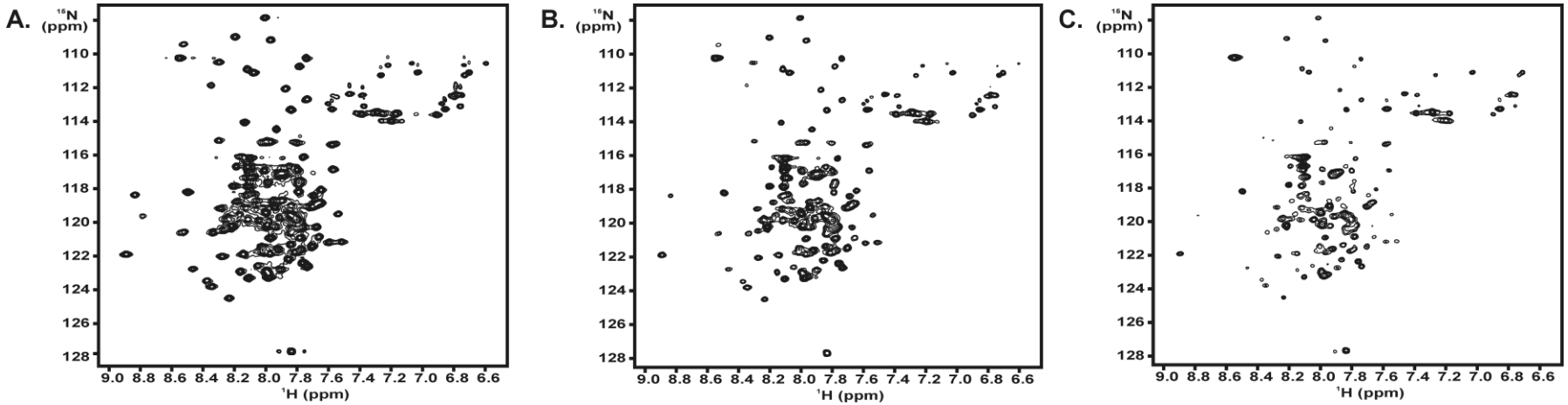

FIGURE 3-8 Spectral quality for TM1-TM3 in LMPG micelles in a low conductivity buffer. [^{15}N, ^{1}H]-HSQC spectra of Ste2p TM1-TM3 (G31-R161) in LMPG micelles prepared in HEPES buffer. (A) Spectrum of a freshly prepared sample. (B) Spectrum after 5 days of incubation at 45°C. (C) Spectrum after 12 days of incubation at 45°C. 0.43 mM uniformly [^{15}N]- labeled TM1-TM3 peptide was solubilized in 20 mM LMPG micelles prepared in HEPES buffer.

Optimization of NMR Sample Conditions for Ste2p TM127 (G31-T114, T274-L340)

Initial [^{15}N,^{1}H]-HSQC analysis was conducted in 75% TFE:water to ascertain whether a high-quality spectrum could be obtained for this construct. A 0.8 mg sample of [^{15}N]-TM127 was solubilized in 75% TFE:water as described above and subjected to [^{15}N,^{1}H]-HSQC analysis at 25ºC, 35ºC, and 45ºC. As with TM1-TM3, the optimal temperature for HSQC experiments was found to be 45ºC (Figure 3-9), as this temperature produced the largest number of expected crosspeaks. However, even at this optimized temperature only 99 of 146 expected crosspeaks were clearly visible and the signal intensity was not uniform. Although these results were not promising, LSPG detergent micelle investigations were attempted.

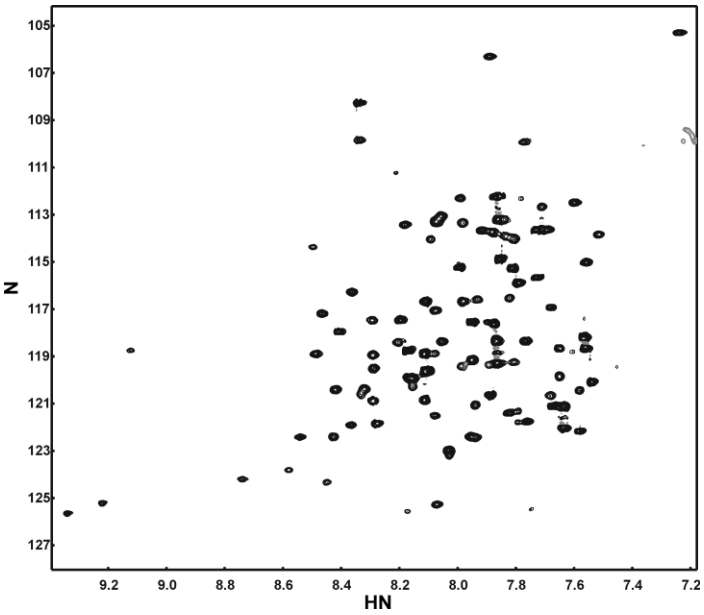

FIGURE 3-9 [^{15}N, ^{1}H]-HSQC spectra of Ste2p TM127 (G31-T114, T274-L340) in 75% TFE:water at 45°C. The peptide concentration was 0.18 mM.

[^{15}N]-TM127 was solubilized in LSPG micelles prepared in sodium phosphate buffer , pH 6.5, as described above to create a 0.15 mM sample. The sample was optically clear and not turbid. Unfortunately, due to unforeseen technical problems with the NMR, the sample had to be stored on a 45ºC heat block for three weeks before HSQC analysis could be performed. The spectrum was severely overlapped and very few

strong, well-resolved signals were observed (Figure 3-10A). The sample preparation was repeated when the instrument was operational, but the results could not be reproduced. The sample was extremely cloudy and turbid. The [^{15}N,^{1}H]-HSQC spectrum showed one largely overlapped region with several well-resolved resonances. The well-resolved resonances were most likely due to the residues in the unstructured tail regions of this construct. At this point, I halted all further analysis of the TM127 construct.

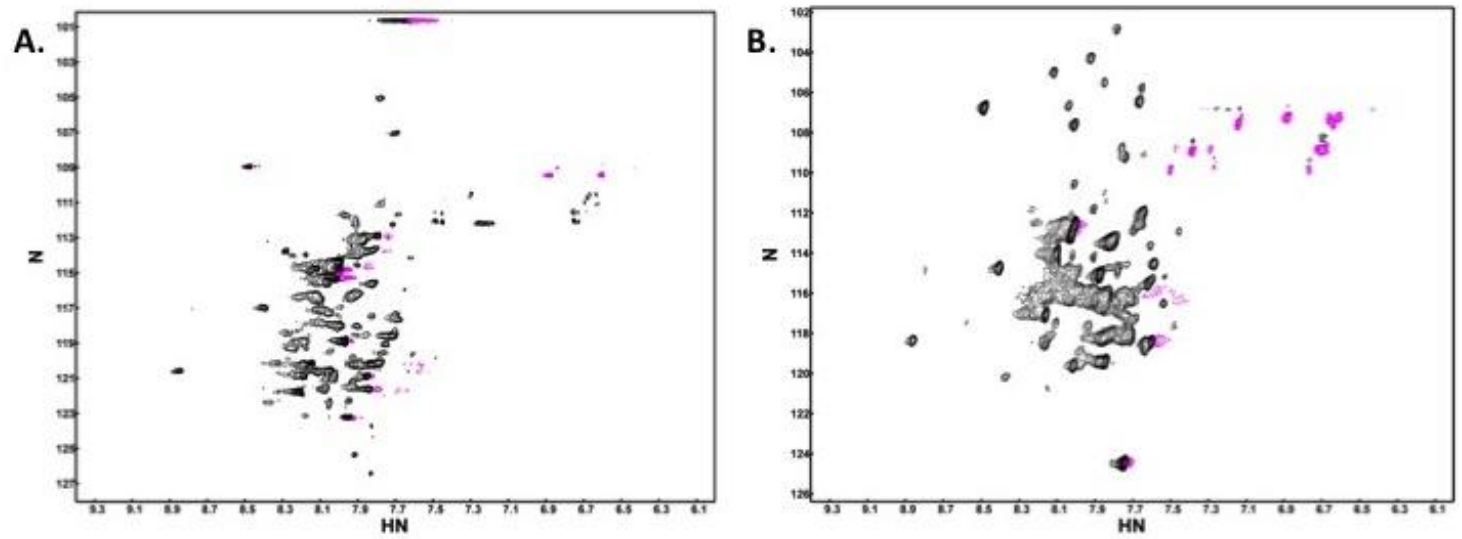

FIGURE 3-10 [^{15}N, ^{1}H]-HSQC spectra of Ste2p TM127 (G31-T114, T274-L340) in LSPG micelles prepared in 20 mM sodium phosphate buffer at pH 6.5 at 45°C. An optically clear sample was placed on a 45°C heat block before HSQC analysis was performed (A). A second identical sample preparation led to a cloudy, turbid solution (B).

Distribution of the 3TM Fragment Workload

As the structural characterization of Ste2p fragments is a collaborative project with the Zerbe group at the University of Zurich, it was decided that all further characterization of TM1-TM3 and TM127 in micelles would be conducted in Zurich while characterization of TM1-TM3 in TFE:water would be done in New York. Since detergent signals significantly overlap sidechain resonances in [^{13}C, ^{1}H]-HSQC experiments, the Zerbe group set out to optimize conditions for sample preparation in deuterated detergent. Deuterated detergent needs to be custom synthesized and is quite expensive. As the TM1-TM2 structure was done in LPPG, the Zerbe group had access to deuterated LPPG and began optimization of sample preparation in this detergent. Using a modified version of the Killian protocol (*90*), Martin Poms was able to produce optically clear samples in for both TM1-TM3 and TM127 in LPPG:DPC mixed micelles with a 4:1 molar ratio that produced [^{15}N, ^{1}H]-HSQC spectra with superior peak separation and dispersion compared to the LMPG and LSPG samples (Figure 3-11). Unfortunately, Martin has run into difficulties with sample stability for TM1-TM3. The samples remained optically clear, but the NMR spectral resolution deteriorated over time. Upon analysis of the NMR samples by Size Exclusion Chromatography (SEC), Martin discovered that these samples contained aggregated TM1-TM3. Addition of cholesterol sulfate to the mixed micelle samples improved sample stability enough to conduct 3D heteronuclear experiments for backbone assignment. I have supplied him with TM127 in a variety of isotope-labeled forms. He is currently working on sidechain assignment and structure calculation for both the Ste2pTM1-TM3 and Ste2pTM127 fragments in detergent micelles.

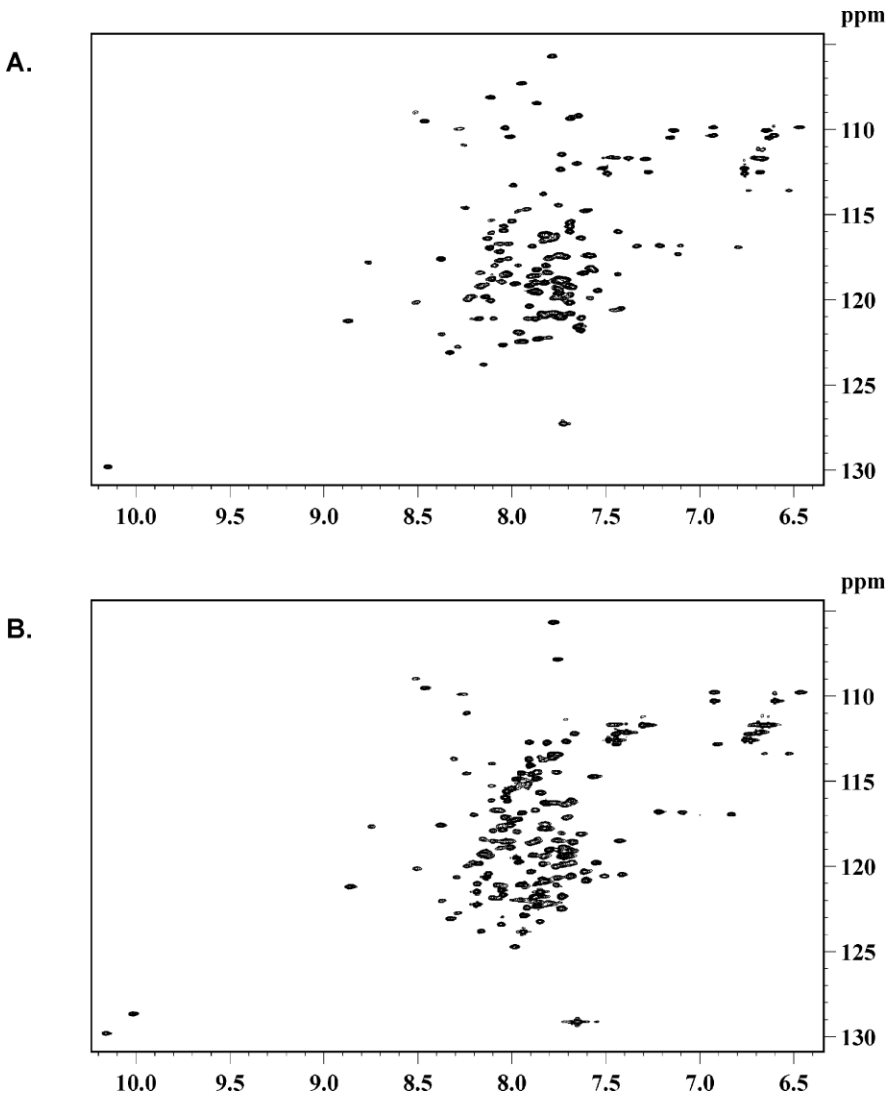

FIGURE 3-11 [^{15}N,^{1}H]-HSQC spectra of Ste2p TM1-TM3 (G31-R161) (A) and TM127 (G31-T114, T274-L340) (B) in 4:1 LPPG:DPC mixed micelles prepared in 40 mM sodium phosphate buffer at pH 6.5 at 45°C. These experiments were performed by Martin Poms at the University of Zurich.

Conclusions

The TM1-TM3 and TM127 fragments of Ste2p have been subjected to biophysical analysis by CD spectroscopy and [^{15}N, ^{1}H]-HSQC experiments in order to determine the secondary structure of the fragments and the optimal sample conditions for NMR backbone assignment and structure determination. The fragments were determined to assume an α-helical structure in both organic aqueous and detergent micelles.

Calculated f_H values between 40-50% were considered to represent the most native-like sample conditions. Analysis of both spectral shape and calculated f_H led to the identification of 50% TFE:water containing 0.1% TFA as the optimal sample conditions for organic:aqueous structure determination for Ste2p TM1-TM3. TFE:water mixtures were not found to be a suitable membrane mimetic for TM127 as the fragment was only found to be soluble at high TFE concentrations. Due to the helix inducing nature of TFE, these conditions would not represent a native-like state for the receptor fragment and indeed we observed f_H values considerably higher than expected. Although LMPG and LSPG micelles were initially identified as suitable conditions for TM1-TM3 and TM127 structure determination, further sample preparation optimization led to the identification of 4:1 LPPG:DPC mixed micelles as the sample conditions for structure determination, with cholesterol sulfate added to the TM1-TM3 samples to improve sample stability. 3D heteronuclear experiments for backbone assignment under these two conditions have been conducted, and assignments are underway. Based on my CD and preliminary NMR studies I decided to undertake a detailed structural analysis for Ste2p TM1-TM3 in TFE/water. My findings are reported in the next chapter.

CHAPTER 4

NMR Structure Determination in 50% TFE:Water

Introduction

NMR spectroscopy was used to determine the 3D structure of the TM1-TM3 fragment of Ste2p. The results of the biophysical analysis discussed in Chapter 3 established that the optimal temperature for backbone assignment and subsequent structure determination in 50% TFE:water was 45°C. This temperature produced a well-resolved spectrum with the majority of the 128 expected crosspeaks visible and a relatively wide dispersion in the [^{1}H] dimension (7.4 - 8.8 ppm).

NMR structure determination requires chemical shift assignment of all backbone atoms and the majority of side chain atoms. A series of 3D heteronuclear experiments in conjunction with [^{15}N, ^{1}H]- and [^{13}C, ^{1}H]-HSQC experiments are used to assign chemical shifts for all atoms. Through bond NMR experiments that transfer magnetization through the protein sequence are used for this task. Experiments originating from the amide proton (HNCO, HNCA, HN(CA)CO, and HNCACB) are used to assign all backbone resonances and all crosspeaks in the [^{15}N, ^{1}H]-HSQC spectrum (*88, 91-97*). Experiments originating from side chain aliphatic protons are used to assign side chain resonances. The HCCH-TOCSY experiment (*98, 99*) transfers magnetization between side chain protons and carbons. Side chain assignments using this experiment are based on prior assignment of alpha and beta carbons. The HCCC(CO)NH experiment (*100*) also transfers magnetization between side chain protons and carbons, but the final transfer links the side chain resonances to the amide proton, allowing for assignment of highly overlapped regions of the [^{13}C,^{1}H]-HSQC (*91, 98*). Assignment of side chain resonances

using this experiment is based on prior assignment of alpha and beta carbons as well as the amide proton. Use of the [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-ILV)]-labeling scheme and the (HM)CM(CBCA)NH and (HM)CM(CGCBCA)NH (*83*) experiments aid in the assignment of I, L, and V side chains. These pulse sequences transfer magnetization from the methyl protons through all side chain carbons, with final magnetization transfer to the amide proton. This links side chain methyl groups to the amide protons, with assignment of side chain resonances based on prior assignment of the alpha and beta carbons as well as the amide proton. Use of [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-ILV)]-labeled protein allows for the assignment of only I, L, V methyl groups because these are the only protonated methyl groups in the protein. As the methyl region of the [^{13}C,^{1}H]-HSQC is very crowded, especially for a protein of this size, this selective methyl assignment procedure greatly simplifies overall methyl carbon assignments. Combination of the HCCH-TOSCY, HCCC(CO)NH, and (HM)CM experiments should allow for the assignment of all crosspeaks in the [^{13}C,^{1}H]-HSQC.

Backbone assignments can be used to determine structural features such as torsion angles, secondary structure, and relative mobility. Torsion angles are calculated using the TALOS+ program (*101*). This program uses a database of 200 proteins whose structures have been solved by both NMR and X-Ray crystallography to predict the torsion angles of a given residue based on the chemical shift. These angles are later used in structure calculation. The secondary structure of a protein can be qualitatively assessed by NMR using chemical shift differencing analysis (*102*), T_2 relaxation experiments, and H-D exchange experiments. Chemical shift differencing involves subtracting standard random coil chemical shifts from measured chemical shift values.

Positive differences for CO and CA and negative differences for HA indicate helical structure. T_2 relaxation experiments are conducted on [^{15}N]-labeled protein. A series of [^{15}N,^{1}H]-HSQC experiments are collected with increasing relaxation times. The data are compiled and relaxation times for each residue are determined. For H-D exchange analysis, a fully protonated [^{15}N]-labeled protein is solubilized in a fully deuterated solvent. A series of [^{15}N, ^{1}H]-HSQC spectra are measured, the volumes of the amide crosspeaks are monitored over time, and exchange rates are calculated based on the compiled data. Residues with amide protons involved in hydrogen bonds or that are shielded from the solvent have longer exchange times than unstructured residues. Combination of data from these analyses gives a rough estimate of the structured region of the protein.

The relative flexibility of a protein can also be assessed using a simple 2D NMR experiment, the HNOE experiment (*103*). In this experiment, a set of two [^{15}N, ^{1}H]-HSQC experiments are collected. The first experiment is a standard HSQC experiment with an increased delay at the beginning of the pulse sequence. In the second experiment, a proton-decoupling sequence is added during the increased delay. This decoupling essentially creates a dipole-dipole interaction between the ^{15}N atoms and the irradiated protons. This is manifest as a Nuclear Overhauser Effect (NOE) between the amide proton and nitrogen. Because all of the nitrogen-proton distances for the amide protons are the same, the distance component of the NOE is removed, and the amplitude of the NOE depends only on the correlation time, which is an indicator of mobility. Measuring the ratio of the peak amplitude with and without the additional pulse for all resulting HSQC peaks gives an indication of the relative mobility for every residue in the

protein sequence. Residues with amplitude ratios of 1 are essentially immobile residues. The more the ratio for a given residue deviates from one, the more mobile the residue.

Complete structure determination requires knowledge of connectivities between residues in addition to connectivities within a residue. Identifying such interactions is accomplished using through space experiments, namely experiments that employ NOE, and Paramagnetic Relaxation Enhancement (PRE) experiments. NOE experiments can be used to identify interactions between protons that are within 5 Å of one another. [^{15}N]-edited NOESY experiments (*104*) identify interactions involving protons attached to nitrogen atoms while [^{13}C]-edited NOESY experiments (*105, 106*) identify interactions involving protons attached to carbon atoms. Interactions are identified by chemical shift assignments as described in detail below. PRE experiments require the attachment of a paramagnetic spin label to the polypeptide of interest. This spin label affects the relaxation of residues within 6-20Å of the unpaired electron. The phenomenon is manifest as a loss of intensity or complete loss of a crosspeak in the [^{15}N,^{1}H]-HSQC spectrum (*107-109*). Comparison of [^{15}N,^{1}H]-HSQC spectra of the protein with the radical spin-label and of the same protein with a reduced version of the label allows for identification of residues that are in close proximity to the spin-label. Effects seen for residues which are not close in sequence to the spin-label attachment site are an indication of tertiary folding and can be used as restraints in 3D structure calculation.

Backbone and side-chain chemical shifts, TALOS+ calculated torsion angles, and NOESY data are compiled and supplied to structure calculation software for 3D structure calculation. The UNIO software suite (*110-114*) is used for automatic NOESY assignment and 3D structure calculation. The NOESY spectra and chemical shift

assignments are supplied to this program, and the NOESY connectivities are automatically assigned by the ATNOS-CANDID component of the program based on chemical shifts. This output is then used in the 3D structure calculation by the program CYANA.

This chapter describes the backbone and sidechain assignment of all TM1-TM3 residues in 50% TFE:water. The use of chemical shifts to qualitatively assess the TM1-TM3 structure using secondary shift analysis, H-D exchange experiments, T_2 relaxation experiments, and HNOE experiments are also discussed. Finally, attempts toward 3D structure determination in 50% TFE:water are presented.

Materials and Methods

Three-Dimensional Heteronuclear NMR Experiments for Backbone and Sidechain Assignments of Ste2p TM1-TM3(G31-R161) in 50% TFE:Water (0.1% TFA) at 45 ℃

Triple-resonance NMR experiments were conducted for backbone and sidechain assignment of the TM1-TM3 Ste2p fragment in 50% TFE-d_2:water. All samples were prepared as described in Chapter 3. Briefly, the HPLC purified TM1-TM3 peptide was solubilized in 50% TFE:water and the concentration was determined by UV spectroscopy. Samples were portioned from this stock solution and lyophilized. After lyophilization, the desired volume of TFE-d_2 was added and the sample was sonicated for 15 min at ~50 W. An equal volume of water containing 0.1% TFA was added, and the sample was sonicated again. The sample was then transferred to a Shigemi tube for NMR analysis. A 0.29 mM [^{15}N,^{13}C,^{2}H]-TM1-TM3 (56% ^{2}H) sample was used for [^{15}N,^{1}H]-HSQC, HNCA, HNCO, HN(CA)CO, and HNCACB experiments. A 0.56 mM [^{15}N]-TM1-TM3 sample was used for recording the [^{15}N]-resolved NOESY experiment. These experiments were conducted on a three-channel Varian NMR-S 600 MHz NMR

spectrometer with a z-axis pulsed-field-gradient and a Varian 5mm [^{1}H, ^{15}N, ^{13}C, ^{2}D] cryo-probe at the College of Staten Island. A 0.5 mM [^{15}N,^{13}C]-TM1-TM3 sample was used for [^{13}C]-HSQC, HCCH-TOCSY, [^{13}C]-NOESY, and HCCC(CO)NH analysis. These experiments were performed on a three-channel Bruker AV-700 700 MHz NMR spectrometer equipped with a CRYO TXI inverse triple resonance cryoprobe at the University of Zurich. A 0.35 mM sample of [^{15}N,^{13}C,^{2}H(^{1}H(methyl)-ILV)]-labeled TM1-TM3 was used for (HM)CM(CGCBCA)NH and (HM)CM(CBCA)NH analysis. These experiments were performed on a four-channel Bruker 800 MHz NMR spectrometer equipped with a CRYO TCI triple resonance cryoprobe at the New York Structural Biology Center.

Backbone assignments were completed using HNCO, HN(CA)CO, and HNCACB spectra using the software NMRView5. For backbone assignments, a crosspeak in the [^{15}N,^{1}H]-HSQC spectrum was first linked to a carbonyl resonance in the HNCO spectrum. The nitrogen plane in the 3D HNCO spectrum that corresponded to the nitrogen chemical shift in the [^{15}N, ^{1}H]-HSQC was scanned for a HN-CO crosspeak with the correct HN chemical shift. CA and CB crosspeaks were identified in the 3D HNCACB using the same method, and the identity of the residue was determined by comparing the CA and CB chemical shifts to standard CA/CB shifts for the twenty amino acids. The next consecutive residue was then identified by scanning the HN(CA)CO spectrum for a second, more intense, crosspeak at the same HN chemical shift. The carbonyl chemical shift of this crosspeak was used to identify all HNCO crosspeaks at that chemical shift. The nitrogen plane for each possible HNCO crosspeak was identified, and the HNCACB spectrum was used to decide which HNCO peak should be

chosen. A crosspeak was determined to be correct when the corresponding HNCACB plane exhibited weak crosspeaks with CA and CB chemical shifts identical to the previously assigned residue. This process was repeated until all of the backbone resonances for the protein had been assigned.

Side chain resonances were assigned using the HCCH-TOCSY, HCCC(CO)NH , (HM)CM(CGCBCA)NH, and (HM)CM(CBCA)NH experiments using NMRView 5 and CARA. Assignment of the (HM)CM experiments was done in NMRView5. The HN CM projections of the 3D experiments were used to assign each CM resonance to an I, L, V residue based on previously assigned HN chemical shifts. Assignment of resonances in the HCCH-TOCSY and HCCC(CO)NGH experiments was done in CARA and correlated to 2D [^{15}N]- and [^{13}C]-HSQC experiments. When using the HCCH-TOCSY spectrum for assignment, the [^{13}C]-HSQC spectrum was viewed in a Polyscope window. The HCCH-TOCSY spectrum was then viewed as a strip in the Polyscope window. The crosshair was placed over a given resonance in the [^{13}C]-HSQC, and all TOCSY crosspeaks corresponding to this resonance were displayed in the HCCH-TOCSY strip. Each TOCSY crosspeak was assigned to a given atom by comparing the chemical shift and crosspeak pattern to standard values for the twenty amino acids. Residue identification was based on CA and CB chemical shifts previously assigned using the HNCACB experiment. When using the HCCC(CO)NH experiment for assignment, the [^{15}N]-HSQC spectrum was viewed in a Polyscope window, and the HCCC(CO)NH spectrum was viewed as a strip in this window. The crosshair was placed over a given resonance in the [^{15}N]-HSQC, and all HCCC(CO)NH crosspeaks corresponding to this resonance were displayed in the strip. Assignments of all crosspeaks in the strip were

made based on standard chemical shifts for the twenty amino acids, and residue identification was based on previously assigned amide proton chemical shift. The HCCH-TOCSY, HCCC(CO)NH, and (HM)CM experiments were used simultaneously rather than exclusively because they all provided benefits. The resolution of the HCCH-TOCSY is superior to that of the HCCC(CO)NH experiment, and aliphatic proton and carbon chemical shifts could be determined accurately using this experiment. However, the [^{13}C]-HSQC spectrum was very crowded and largely overlapped, and assignment of sidechain resonances using the CA and CB chemical shifts was extremely difficult. As the chemical shifts for amide protons are more disperse for those of CA and CB, residue identification was more accurate using the HCCC(CO)NH experiment. Selective methyl-carbon assignment using the (HM)CM experiments allowed for easier assignment of the overcrowded methyl carbon region of the [^{13}C,^{1}H]-HSQC.

All assigned backbone and sidechain resonances were used to generate an assignment list for automated NOESY assignment.

Secondary Structure Localization in Ste2p TM1-TM3 (G31-R161) in 50% TFE:Water (0.1% TFA) at 45 ℃

Chemical shift difference analysis of carbonyl, α-carbon and α-proton nuclei was performed to determine the secondary structure of the TM1-TM3 fragment. Random coil chemical shifts (determined in water) were subtracted for every assigned nucleus (*102*). As standard shifts determined in TFE are not available, we conducted our analysis using standard shifts determined in water. Since these standard random coil values were determined for soluble proteins in water, they are appropriate for a protein in its native environment. TFE:water is considered to mimic the native environment for TM1-TM3 in my investigation. Therefore, we are assuming that comparison of the TM1-TM3 shifts

determined in TFE:water to the random coil shifts determined in water is a valid approach. The overall helix content was determined by excluding all residues for which the difference was below 1 ppm in the case of CO and CA or larger than -0.1 ppm in the case of HA.

T_2 relaxation experiments were performed on a 0.5 mM [^{15}N]-TM1-TM3 sample solubilized in 50% TFE-d_2:water. A series of eight [^{15}N,^{1}H]-HSQC experiments were performed with varying relaxation times (Table 4-1). The experiments were run in the order specified in the table in order to minimize sample heating. The data were compiled and the rate analysis function of NMRView5 was used to calculate the relaxation time for each residue. A plot of relaxation time vs. residue number was constructed.

Table 4-1: T_2 Relaxation Parameters for Ste2p TM1-TM3(G31-R161) in 50% TFE:water

Experiment Number	Relaxation Time (ms)
1	0
2	210
3	10
4	150
5	30
6	110
7	50
8	70

H-D exchange analysis was performed on a 0.5 mM [^{15}N]-TM1-TM3 sample solubilized in fully deuterated 50% TFE-d_3:D_2O containing 0.1% deuterated TFA. [^{15}N,^{1}H]-HSQC experiments were performed at 40 minute intervals for 5 hours. Additional spectra were collected daily with the final spectrum collected eight days after the original sample preparation. The rate analysis function of NMRView5 was used to calculate the exchange time for each assigned residue, and a logarithmic plot of the

inverse of the exchange rate vs. residue number was constructed.

Assessment of the Relative Flexibility of Ste2p TM1-TM3 (G31-R161) in 50% TFE:water

A 0.5 mM sample of [^{15}N]-TM1-TM3 was solubilized in 50% TFE-d_2:water and subject to HNOE analysis at 45°C and 30°C. Two datasets for each temperature were created. The first was used as a reference and was essentially a [^{15}N,^{1}H]-HSQC. The second was a [^{15}N,^{1}H]-HSQC with the use of the HNOE pulse. The amplitude and volume for each crosspeak in each data set was calculated using CARA. The ratio of the HNOE and reference amplitudes was calculated and the resulting data were smoothed by averaging neighboring residue ratios in groups of three. The smoothed data were plotted as a function of residue number.

Paramagnetic Relaxation Enhancement Experiments for Ste2p TM1-TM3(G31-R161)

The TM1-TM3S104C direct expression product described in Chapter 2 was used for PRE experiments. The purified protein was split into two samples and lyophilized. Both samples were solubilized in 50% TFE:water. A five-fold molar excess of either MTSL or an acetylated version of MTSL (Ac-MTSL) was added to each sample for disulfide attachment with the Cys residue. The Ac-MTSL was chosen as an approach to generate a "reduced" version of MTSL for use as a control for the PRE experiments. This approach was used because, in our hands, the nitroxide radical version of MTSL could not be regenerated after reduction with acid. Reduction of nitroxide labeled proteins has been used by others to generate a control for PRE experiments (*108*). However, in my studies, because the spin-labeled material was difficult to produce, we did not want to reduce the radical if we could not regenerate it.

The reactions with MTSL and Ac-MTSL were incubated at room temperature

overnight. The progress of the reaction was tracked by HPLC and ESI-MS, but the intensity of the HPLC peak was found to decrease over time. Both samples were lyophilized and resuspended in 350 μL of 50% TFE-d_2:water. [^{15}N,^{1}H]-HSQC spectra were collected at 45°C, 35°C, and 30°C. The spectral quality for the Ac-MTSL was found to be unacceptable for further analysis. As a result, a sample of [^{15}N]-TM1-TM3 without the S104C mutation was used for reference spectra. NMRView5 was used to calculate the amplitude and volume of all crosspeaks in the MTSL and reference spectra. All values for overlapping peaks were removed, and the ratio of MTSL to the reference volume or amplitude was calculated. The data were normalized so that the majority of the ratios were approximately one and then smoothed by averaging neighboring residue ratios in groups of three. The smoothed data were plotted as a function of residue number.

NOESY Assignment and Structure Calculation for Ste2p TM1-TM3(G31-R161)

Backbone chemical shift assignments were used to calculate torsion angles for TM1-TM3 using TALOS+. This program scanned the TM1-TM3 sequence as a series of tripeptides and compared the HN, HA, CA, CB, CO and N chemical shifts to a database in order to determine the most likely phi and psi angles for each residue. The ATNOS-CANDID (*110, 111*) component of the UNIO suite was used for automated assignment of the [^{13}C]- and [^{15}N]-NOESY crosspeaks based on the assigned backbone and sidechain chemical shifts. The program picked peaks in each of the NOESY spectra, and the chemical shifts of the picked peaks were then compared to the provided backbone and sidechain chemical shift assignments. Each NOESY crosspeak was assigned to a given residue based on chemical shift matching within a user-defined threshold. The program

generated an assignment file, and transferred all of the provided data along with the assignment file to the CYANA component of the software suite. The TM1-TM3 sequence, assignments, calculated TALOS+ angles, and [^{15}N]- and [^{13}C]-edited NOESY spectra were input to the UNIO software suite (*110, 111, 113, 114*) for NOESY assignment and CYANA (*112*) structure calculation. A seven cycle CYANA iteration was used to calculate 80 structures, and the 20 lowest energy structures for TM1-TM3 were analyzed. Each cycle adjusted the NOE restraints according to the value of the target function and the number of violations, with the last cycle providing the lowest target function, the smallest number of violations, and the lowest energy structures. A low target function is an indication of a small number of violations of torsion angles, distant restraints and Van der Waals contacts.

Results

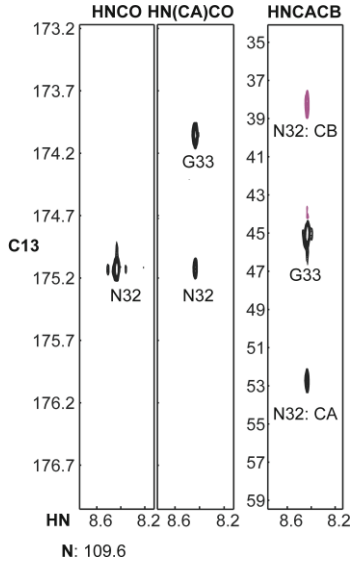

Figure 4-1: Illustration of backbone assignment procedure for Ste2p TM1-TM3 (G31-R161) in 50% TFE:water. The amide nitrogen plane (109.6 ppm) for G33 was selected in all 3D spectra. A CO-HN crosspeak is observed in the HNCO spectrum, with the HN shift matching that of G33. The HN(CA)CO spectrum contains two crosspeaks at the G33 amide proton shift: a weak crosspeak with a CO shift identical to the N32 shift identified in the HNCO spectrum and a stronger crosspeak corresponding to the CO of G33. The HNCACB spectrum reveals three crosspeaks at the G33 amide proton shift. The chemical shift pattern is consistent with an asparagine-glycine sequence.

Backbone Assignments for Ste2p TM1-TM3 (G31-R161) in 50% TFE:Water (0.1% TFA)

Three-dimensional heteronuclear NMR experiments were conducted in 50% TFE:water at 45°C, as these conditions were found to be optimal by the [^{15}N,^{1}H]-HSQC

analysis described in Chapter 3. HSQC, HNCO, HN(CA)CO, HNCACB, HCCH-TOCSY, and HCCC(CO)NH experiments were used for backbone assignments (HN, N, CO, CA, CB, HA, HB). The amide proton (HN) and nitrogen (N) chemical shifts were picked in the [^{15}N,^{1}H]–HSQC spectrum, and assigned to specific residues using 3D triple-resonance experiments. The carbonyl resonances (CO) were assigned using HNCO and HN(CA)CO experiments, and the α and β-carbons (CA and CB, respectively) were assigned from the HNCACB experiment (See Figure 4-1 for an illustration of the backbone assignment process). The α and β-protons were assigned based on the CA and CB shifts using the HCCH-TOCSY experiment.

All backbone resonances were assigned except for Gly 31, Pro79, and Pro117 (Figure 4-2, Appendix B). A comparison of the obtained chemical shifts with those from TM1-TM2 indicates that for most residues in the TM1-TM2 portion of the sequence the Δδ values are quite similar and close to zero. Significant differences are limited to residues at the C terminus of TM2 (Figure 4-3). This comparison suggests that small fragments may be useful in the assignment of larger fragments. It also suggests that the TM1-TM2 and TM1-TM3 fragments may adopt a similar structure in 50% TFE:water.

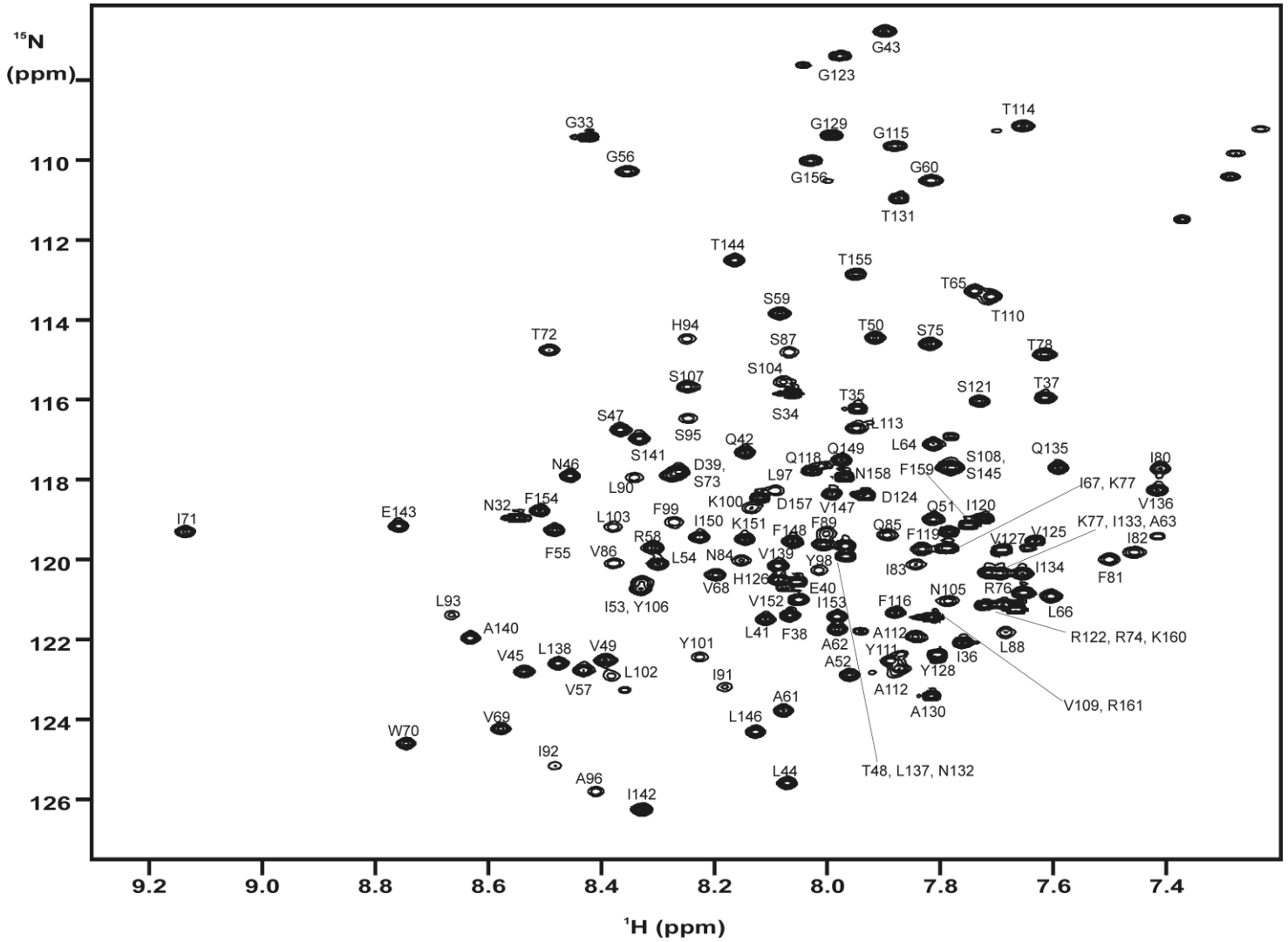

Figure 4-2 [^{1}HN] and [^{15}N] Backbone assignments of Ste2p TM1-TM3 (G31-R161) in 50%TFE:water at 600 MHz and 45°C.

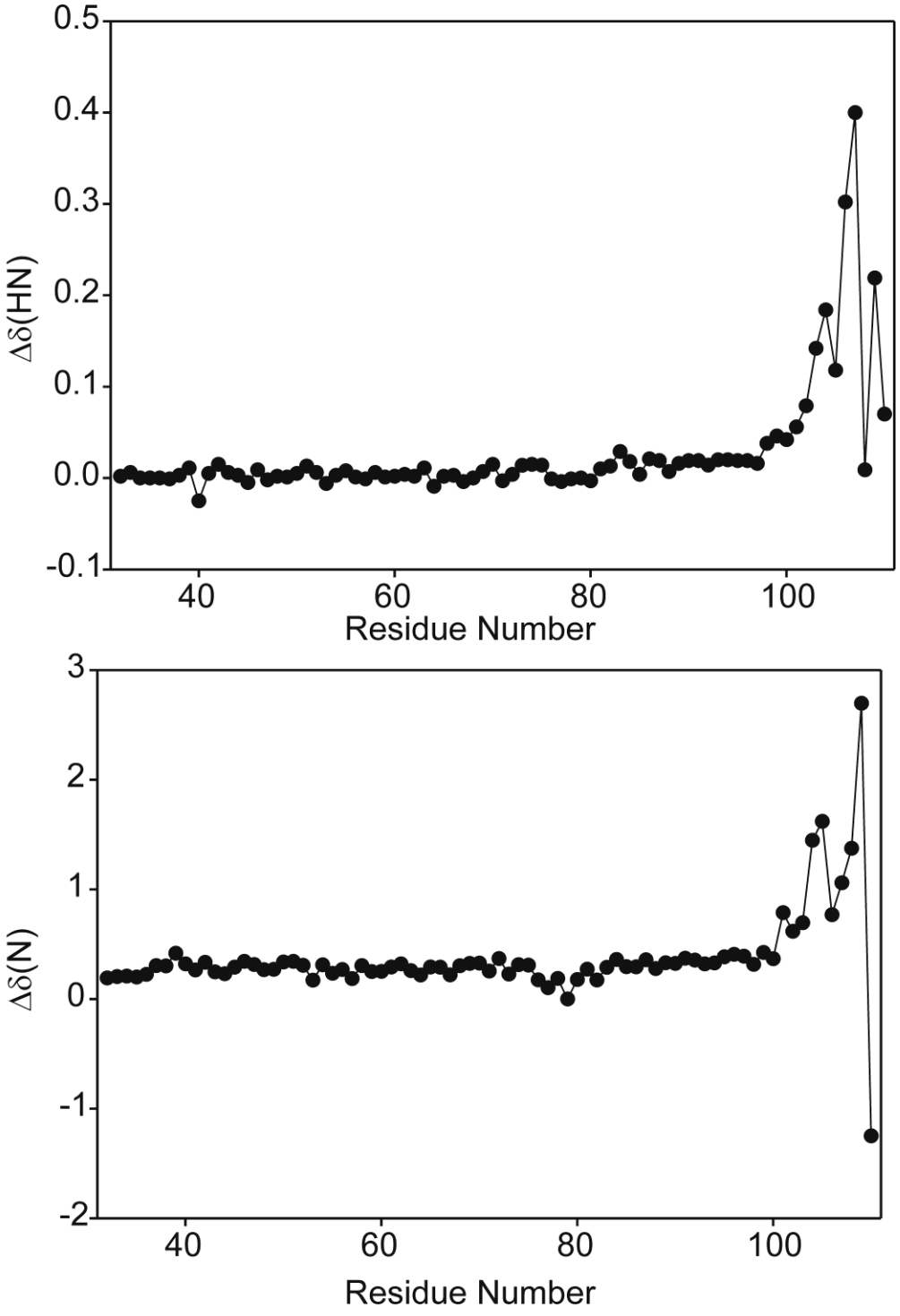

Figure 4-3: Comparison of chemical shift assignments for Ste2p TM1-TM2 (G31-T110) and TM1-TM3 (G31-R161) in 50% TFE:water. Chemical shift assignments for the backbone amide (HN (Top) and N (bottom)) atoms were subtracted. The largest deviations are observed at the C-terminus of TM2.

Side chain assignments were conducted using the HCCH-TOCSY, HCCC(CO)NH, (HM)CM(CGCBCA)NH, and (HM)CM(CBCA)NH experiments as described above. Data from all three experiments were combined to assign all [^{13}C,^{1}H]-HSQC crosspeaks. All side chain assignments were made except for the aromatic residues.

Qualitative Analysis of the Secondary Structure of Ste2p TM1-TM3(G31-R161) in 50% TFE:water by Chemical Shift Differencing

The assigned chemical shifts were used to analyze the secondary structure of the TM1-TM3 fragment using chemical shift difference analysis (*102*). CO, CA, and HA secondary chemical shifts versus the position in the sequence are shown in Figure 4-4. Positive values for CO and CA differencing and negative values for HA differencing indicate a helical secondary structure. The trends seen in the chemical shift difference data reveal the presence of long helical stretches for TM1-TM3 in the organic-aqueous solvent (see Figure 4-4). This is not surprising as TM1-TM3 is a fragment from a helical membrane protein and TFE is a known helix inducer. The CO, CA and HA chemical shift difference analyses show three major regions of helicity with distinct breaks. The breaks are consistent with the putative location of the loop regions (approximate boundaries: residues 31-48, 67-83, 101-130, 151-161)[71]. Analysis of the differencing (Figure 4-4) reveals reduced helicity in the TM1 region of the fragment centered around a GXXXG motif (residues 56-60). The breaks in helicity at these predicted flexible regions suggest that the observed helicity is a structural feature of the fragment as opposed to an artifact of the solvent system.

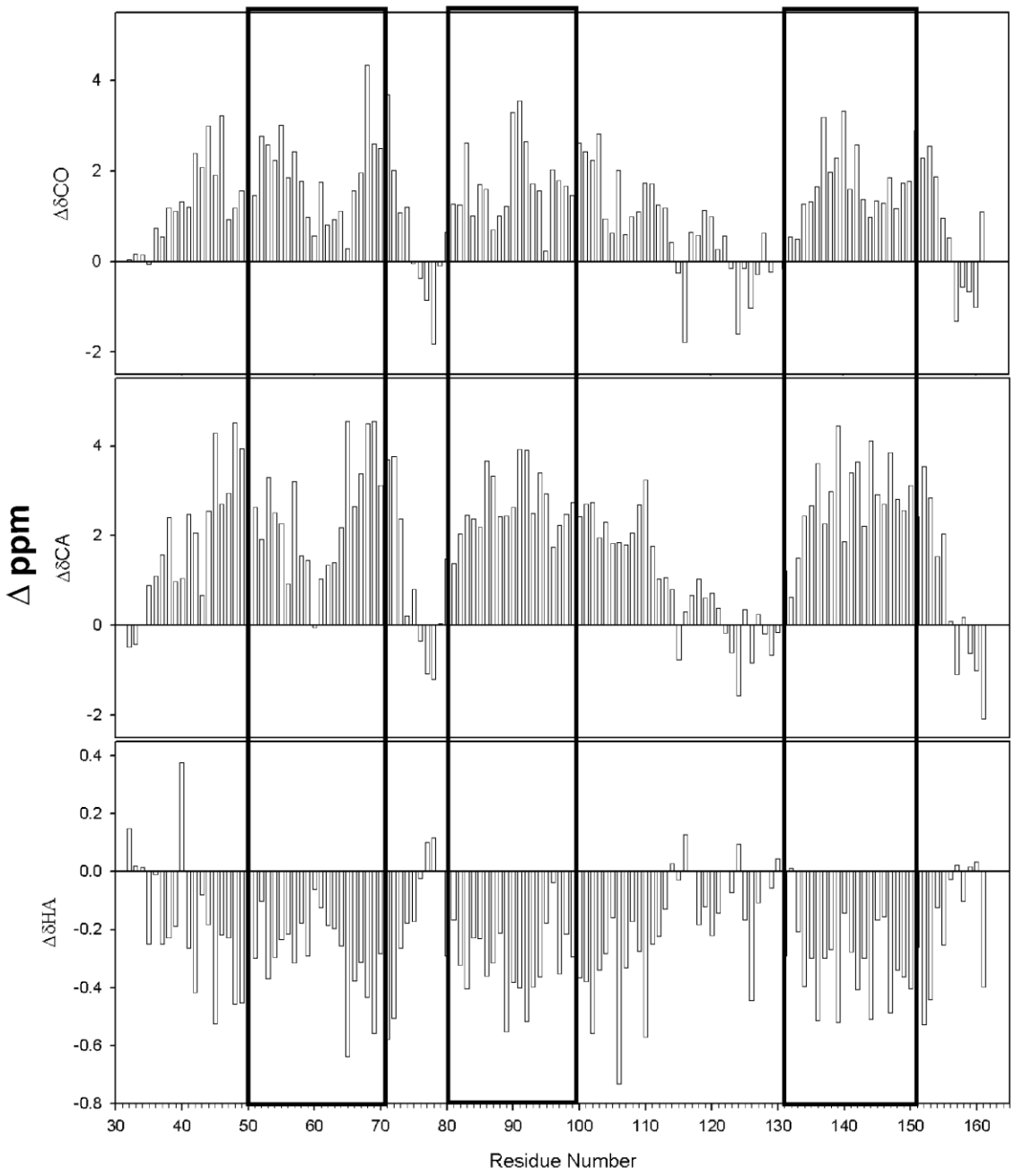

Figure 4-4: Secondary chemical shifts of CO, CA, and HA for Ste2p TM1-TM3 (G31–R161) in 50% TFE:water at 45°C. The difference values were calculated as described in the Methods. The boxed helical boundaries are those calculated by sequence-based TM prediction software (average of four programs).

Preliminary inspection of the first two regions of helicity revealed that the helical stretches are too long to be considered strictly TM helices which normally have approximately 20-24 residues depending on the tilt of the helix. In order to assess whether the observed extension of these helices in the chemical shift data was an artifact of the solvent system, boundaries for the TM helices of Ste2p were required. Putative TM helix boundaries for Ste2p were determined from the rhodopsin-templated computer model and from several sequence based computer programs (Table 4-2). The rhodopsin template model of Ste2p (*71*) predicts three TM helices encompassing residues 49-66, 84-100, and 131-150. As large portions representing the putative loop regions of the receptor are truncated in this model, these boundaries are most likely inaccurate. The transmembrane prediction programs TM Pred, TMHMM, Split, and SOSUI (*115-118*) all produced very similar results, with the beginning of TM1, TM2, and TM3 ranging from 47-52, 79-80, and 130-133 and the ends ranging from 71-73, 97-104, 148-155, respectively. As the results from all four programs were consistent and represented logical boundaries, the average of these results was calculated and used as the putative TM helix boundaries for TM1-TM3 of Ste2p. These calculated boundaries were 50-72, 80-100, and 132-152 for TM1, TM2, and TM3 respectively, and are represented by the boxed regions in Figure 4-4.

Table 4-2: Determination of helix boundaries for Ste2p TM1-TM3(G31-R161).

Method	TM1	TM2	TM3
Eiler's Model	49-66	84-100	131-150
TM Pred	52-73	80-97	132-148
TMHMM	51-73	80-102	133-155
Split	47-71	80-101	130-153
SOSUI	50-71	79-101	132-154

Comparison of the predicted helical boundaries to the chemical shift differencing results reveals that the NMR chemical shift analysis for the three TM helices results in stretches that are longer than those predicted by the various algorithms. TM1 is N-terminally extended by approximately fourteen residues and C-terminally extended by approximately 3-4 residues, while TM2 and TM3 are C-terminally extended, with TM2 extended by approximately eleven residues and TM3 extended by approximately five residues. Analysis of the CO differencing (Figure 4-4, top) data indicates smaller chemical shift differences for residues 48-51 which might be consistent with a reduced helicity beginning at residue 48, and a break in this helix. Based on this observation, the extended TM1 helix observed by chemical shift differencing most likely represents a coalescence of two helical regions, one in the N-terminal extracellular region of Ste2p and the second corresponding to TM1 of this receptor. This N-terminal helix was previously identified in the structure of TM1-TM2 in LPPG micelles (*57*), and encompasses residues 39-47. An N-terminal helix was also observed in a 2TM fragment of the human Y4 receptor (*54*), which is significant as this receptor is also a peptide-liganded GPCR. The observation of this NT helix in two GPCRs that bind peptides suggests that this structural feature may be a common feature of this subclass of GPCRs, and that it may be a functionally relevant feature. The CO chemical shift differencing analysis also reveals a dip in the region following the carboxyl end of the second predicted second transmembrane helix centered around residue 104. Biochemical data and molecular modeling indicate that EL1 contains structured regions, including a short helical region ranging from residues 106-114(*78*). The helical region in EL1 revealed by the NMR study could correspond to that helical domain. Based on these observations, we

conclude that the extended helical regions observed by chemical shift differencing are not an artifact of the solvent system and represent structural features of this receptor fragment. Helical regions have been found in EL2 in the crystal structures of several GPCRs, and these helices have been implicated in directing ligands to the binding pockets (*30*). This further supports the hypothesis that both the N-terminal and EL1 helices may be functionally relevant. Detailed analysis of the chemical shift differencing results reveals five helical regions (approximate boundaries: residues 38-49, 51-75, 80-103, 108-115, 132-158). These NMR results are therefore in good agreement with the TM helix predictions.

Assessment of the Secondary Structure of Ste2p TM1-TM3(G31-R161) by H-D Exchange and T_2 Relaxation Experiments

H-D exchange and T_2 relaxation experiments were conducted in order to gain additional insight into the secondary structure of the TM1-TM3 fragment in 50% TFE:water. The results of both of these experiments resembled the chemical shift differencing analysis in that three putative TM helices are observed with distinct breaks in the predicted flexible regions.

Analysis of the H-D exchange results (Figure 4-5) reveals that the residues in the predicted TM helices (boxed regions) had longer H-D exchange times than those in the loop regions, with the majority of the exchange times in the predicted helices ranging from approximately 10-200 h and from 0.02-1 h in the loop regions. Faster exchange is also observed in the middle of TM1, centered around the GXXXG region. The calculated TM helix boundaries seem to correlate very well with the data for TM1 and TM3, as slow exchange times are observed within the boxed regions, and fast exchange times are observed just outside the boxed regions. Except for residues in the very center of TM2,

the exchange data for this TM domain shows that many residues near the ends of the helix have exchange times between 1 and 10 h. This suggests that the helix is not very stable and deviates from the chemical shift differencing data. The putative TM3 region seems to have stronger hydrogen bonding at the N-terminus of the helix, however the difference in exchange times across this region are in general not as large as in the TM2 helix. The additional helices in the N-terminal region and in EL1 are not as obvious as in the chemical shift differencing, however, residues 40-48 in the N-terminal region and 110-114 in EL1 have longer exchange times than the rest of the residues in these flexible regions, suggesting a higher degree of hydrogen bonding in these regions.

T_2 relaxation rates (Figure 4-6) follow an essentially identical pattern as the chemical shift differencing data (compare boxed boundaries of Figures 4-4 and 4-6). The N and C-termini of the TM1-TM3 fragment have exchange times on the order of hundreds of milliseconds. Residues predicted to belong to the TMs have relaxation times between 25-50 ms, with increased relaxation times observed for the GXXXG region in TM1. TM1 again appears to be N-terminally extended, and this extension correlates to the position of the N-terminal helix. Residues in the putative loop regions have relaxation times above 50 ms, with relaxation times for residues in EL1 gradually increasing until residue 120. This gradual increase is consistent with the presence of a short helix in the beginning of EL1. Interestingly the relaxation rates in the TM2 region are the slowest of the 3 transmembrane regions. This is not consistent with the exchange data. Overall, the relaxation results seem to confirm the helix boundaries identified from chemical shift differencing.

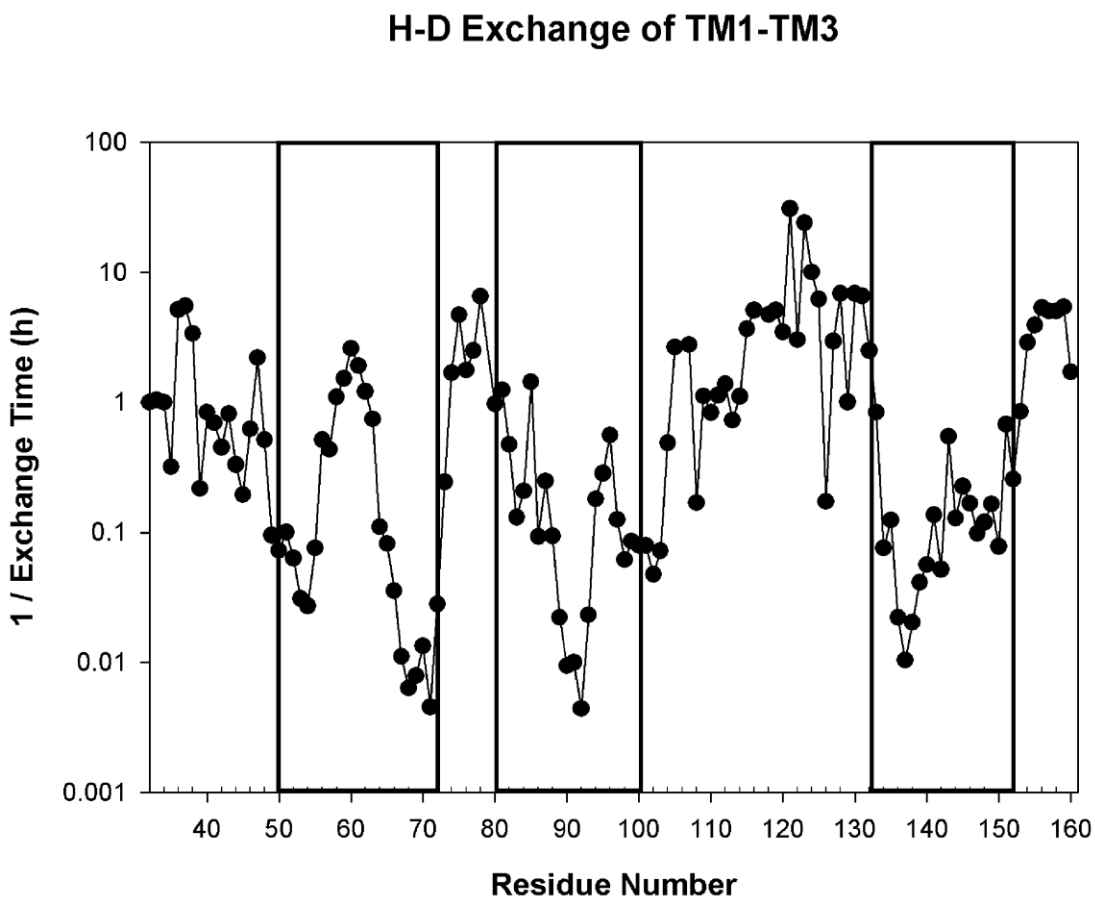

Figure 4-5: Evaluation of the secondary structure of Ste2p TM1-TM3 (G31-R161) by Hydrogen-Deuterium Exchange. Fully protonated [^{15}N]-TM1-TM3 was solubilized in TFE-d_3:D_2O with 0.1% TFA-d. Exchange times were calculated from a series of [^{15}N,^{1}H]-HSQC experiments conducted at 45°C, and plotted as a function of residue number with an inverse logarithmic scale. The boxed helical boundaries are those calculated by sequence-based TM prediction software (average of four programs). The putative TM helices have longer exchange times than the predicted flexible region.

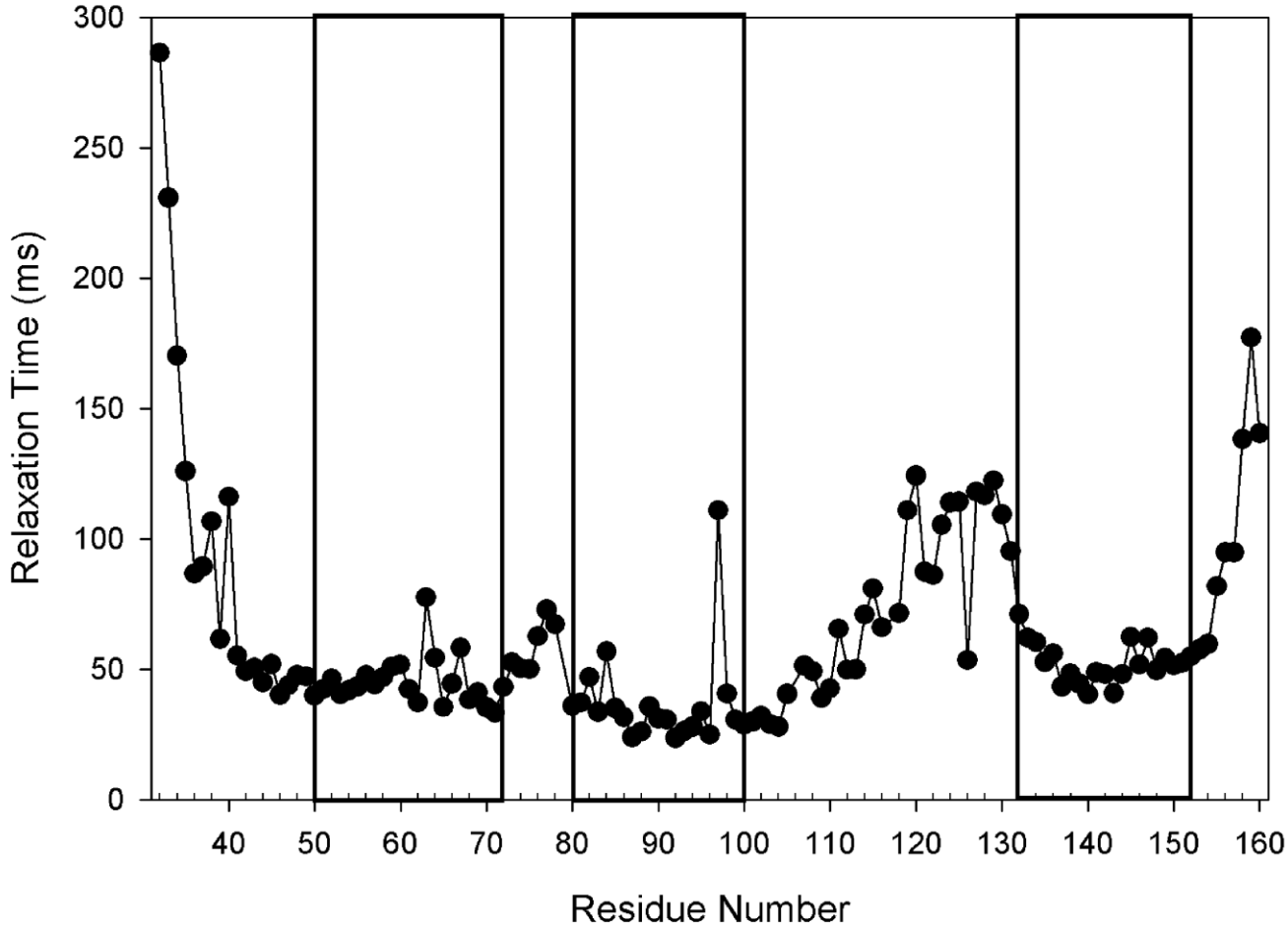

Figure 4-6: Evaluation of the secondary structure of Ste2p TM1-TM3 (G31-R161) by T_2 Relaxation. The T_2 relaxation rates for all residues in TM1-TM3 were measured in 50% TFE:water using the series of experiments conducted at 45°C, outlined in Table 4-1. The boxed helical boundaries are those calculated by sequence-based TM prediction software (average of four programs). The putative TM helices have shorter relaxation times than the predicted flexible regions.

Assessment of the Relative Flexibility of Ste2p TM1-TM3(G31-R161) by HNOE Analysis

Previous studies on the TM1-TM2 fragment of Ste2p revealed that the construct became more structured at low temperature, with long-range PRE contacts observed only when the NMR experiments were run at 25°C. Although backbone assignments were completed at 45°C due to the superior spectral resolution at this temperature, it was believed that very few tertiary contacts would form at this temperature. I therefore ran a series of HSQC experiments in 5°C increments at temperatures between 25 and 40°C in order to transfer the backbone amide assignments from 45°C to lower temperature (data not shown). It was found that 30°C was the lowest temperature for which chemical shift assignment adaptation could be performed. Therefore, I conducted my HNOE analysis and PRE experiments at temperatures no lower than 30°C.

As a first attempt at assessing the degree of tertiary structure of TM1-TM3, HNOE experiments were performed at both 45°C and 30°C. Although the HNOE experiment does not give direct information regarding tertiary fold, it can be used to qualitatively assess the relative mobility of the protein as a function of residue number. If tertiary fold in TFE:water is temperature dependent, then the fragment should become less mobile at 30°C.

HNOE data collected at 45°C (Figure 4-7, top) closely resembles the T_2 relaxation data. Both the N and C-termini have very low amplitude ratios, suggesting that these regions of the fragment are highly flexible. There are three major regions with amplitude ratios clustered at 0.5 and above (dashed line, Figure 4-7). These regions fall within the putative TM regions (boxed regions). There is a small dip in the amplitude ratios centered at the GXXXG region in TM1, and the helix again seems to be N-terminally

extended, consistent with the presence of the N-terminal helix. The putative loop regions show decreased mobility with respect to the termini, but increased mobility with respect to the TM regions. Moreover, the beginning of EL1 shows decreased mobility with respect to the carboxyl end of the loop. This difference in mobility is consistent with the presence of some degree of order in the N-terminal portion of this loop.

The first obvious observation upon comparison of the results obtained at 45°C and 30°C is that the data at 45°C is much smoother than the data at 30°C. This is most likely due to the sharper lines and lower degree of spectral overlap at 45°C. At both temperatures, the N and C-terminal residues are highly flexible, with amplitude ratios that deviate the most from one. At 30°C, the amplitude ratio for most regions of TM1-TM3 increases. Some of these increases may be artifacts (trends in TM2, for instance). We believe, however, that in general the data supports increased rigidity for the protein at 30°C. TM1 appears to become less mobile at decreased temperature (many ratios >0.7), with the trend toward increased mobility centered around the GXXXG region maintained. The C-terminal portion of TM3 also exhibits a decrease in mobility, with amplitude ratios approaching one. The most marked differences are observed for EL1. The N-terminal portion of this loop remains less mobile than the C-terminal portion, however the C-terminal portion exhibits an increase in amplitude ratio from 0 to approximately 0.25. This decrease in mobility would suggest that EL1 becomes more constrained at decreased temperature, implying that some degree of tertiary structure is being adopted. Although no solid conclusions can be drawn from this HNOE data, it is the first evidence that might suggest that TM1-TM3 adopts some degree of tertiary structure in TFE:water at 30°C.

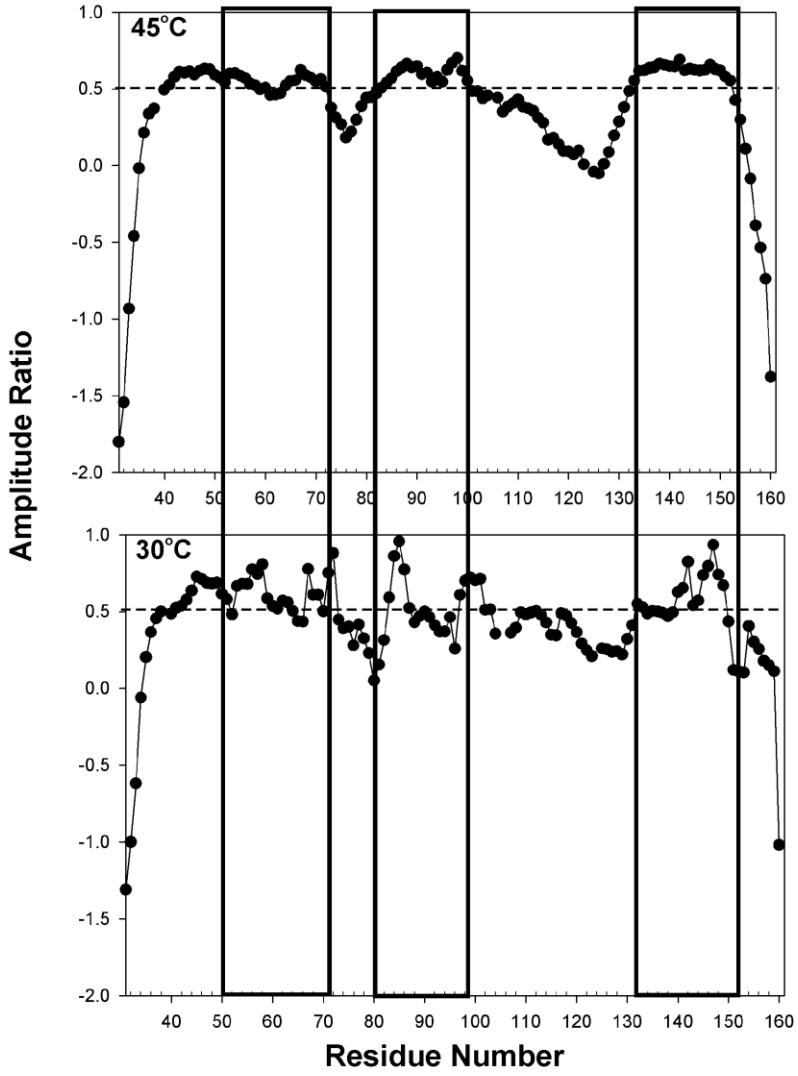

Figure 4-7: Evaluation of the relative mobility of the Ste2p TM1-TM3 (G31-R161) construct by HNOE. [^{15}N,^{1}H]-HSQC experiments were conducted at 45°C and 30°C with and without the addition of a pulse which generated an NOE between the amide nitrogen and proton. The ratio of the peak amplitude with the pulse to the amplitude without the pulse for every residue in the fragment was calculated and plotted as a function of residue number. The boxed helical boundaries are those calculated by sequence-based TM prediction software (average of four programs).

Paramagnetic Relaxation Enhancement Experiments for Ste2p TM1-TM3(G31-R161)

PRE experiments were conducted for directly expressed TM1-TM3S104C with the nitroxide radical spin label, MTSL, attached via disulfide linkage to the Cys at position 104. As mentioned in Chapter 2, approximately 5 mg of [^{15}N]-labeled TM1-TM3S104CDE was obtained from 2L of culture volume. Additional material has yet to be produced. As a result, two 2.5 mg samples were used for MTSL and Ac-MTSL attachment. Both reactions appeared to go to completion as judged by HPLC and MS analysis, but the spectral quality for the Ac-MTSL was very poor. As a result, [^{15}N]-TM1-TM3 was used for all PRE reference data.

[^{15}N,^{1}H]-HSQC experiments were conducted for TM1-TM3S104CDE-MTSL and TM1-TM3 at 45°C, 35°C, and 30°C. The 45°C spectrum was collected because this temperature gave the best resolution. The 30°C spectrum was collected because HNOE experiments conducted at this temperature gave the first indications of tertiary fold, and this was found to be the lowest temperature where spectral resolution was maintained. However, spectra collected at this temperature were severely overlapped and displayed broad linewidths with respect to the 45°C spectra. Due to the poor sample quality, the signal to noise ratio for the spin labeled spectra collected at all temperatures was very low, and the calculated amplitudes were found to be incorrect in some regions of the spectrum due to noise. The 45°C and 35°C spectra were used for PRE analysis because the signal to noise in these spectra was higher than in the 30°C spectrum. All overlapped peaks and peaks in areas of high noise for both temperatures were omitted from amplitude ratio calculations.

Upon inspection of the PRE data (Figure 4-8, top), it is evident that the main

effect of the spin label is seen for residues close in sequence to the MTSL attachment site. Residues at the end of TM2 were more affected by the close proximity of the spin label than those in the loop, with an upward trend in intensity ratios apparent for the end of the loop and the beginning of TM3. At both temperatures, amplitude ratios for the majority of residues in TM1 fall between 0.8 and 1. In both cases, residues at the extracellular side of the helix have lower intensity ratios than those at the intracellular end of the helix. This effect is more pronounced at 35°C, where intensity ratios for the extracellular side of TM1 average ~0.7. This suggests that residue 104 is sometimes close in space to residues at the extracellular end of TM1. A similar effect is observed for the N-terminal portion of TM3. This suggests that residue 104 is also close in space to residues in TM3 at certain times. Therefore, comparison of data collected at 45°C and 35°C suggests that the TM1-TM3 fragment is adopting some level of transient tertiary structure in TFE:water. Because the molecular motions are slowed at lower temperature, the structure is more pronounced at 35°C.

Although this data gives some indication of tertiary folding, the experimental design was flawed and it does not provide conclusive tertiary contacts for use as restraints in structure calculation. Due to the inability to produce large quantities of directly expressed mutant protein, the sample quality for the MTSL sample was poor and the control used was not the same construct. The control sample had a slightly different sequence, a higher concentration, and was higher quality than the MTSL-labeled sample. A margin of error in the intensity calculations was introduced by the poor signal to noise ratio in the MTSL spectra. Because the MTSL-labeled peptide has a His tag and an S to C mutation, some assignments may be incorrect. This experiment needs to be repeated.

The ideal way to repeat this experiment is to produce double-labeled mutant peptide and complete backbone assignments. After assignments have been made, they can be directly used rather than adapted for the MTSL-labeled sample. A high-concentration sample of MTSL-labeled protein must be used and a high signal to noise ratio must be obtained. Attempts are currently being made to accomplish this goal. Concrete conclusions regarding the folding of this fragment in TFE:water can only be drawn when this is completed.

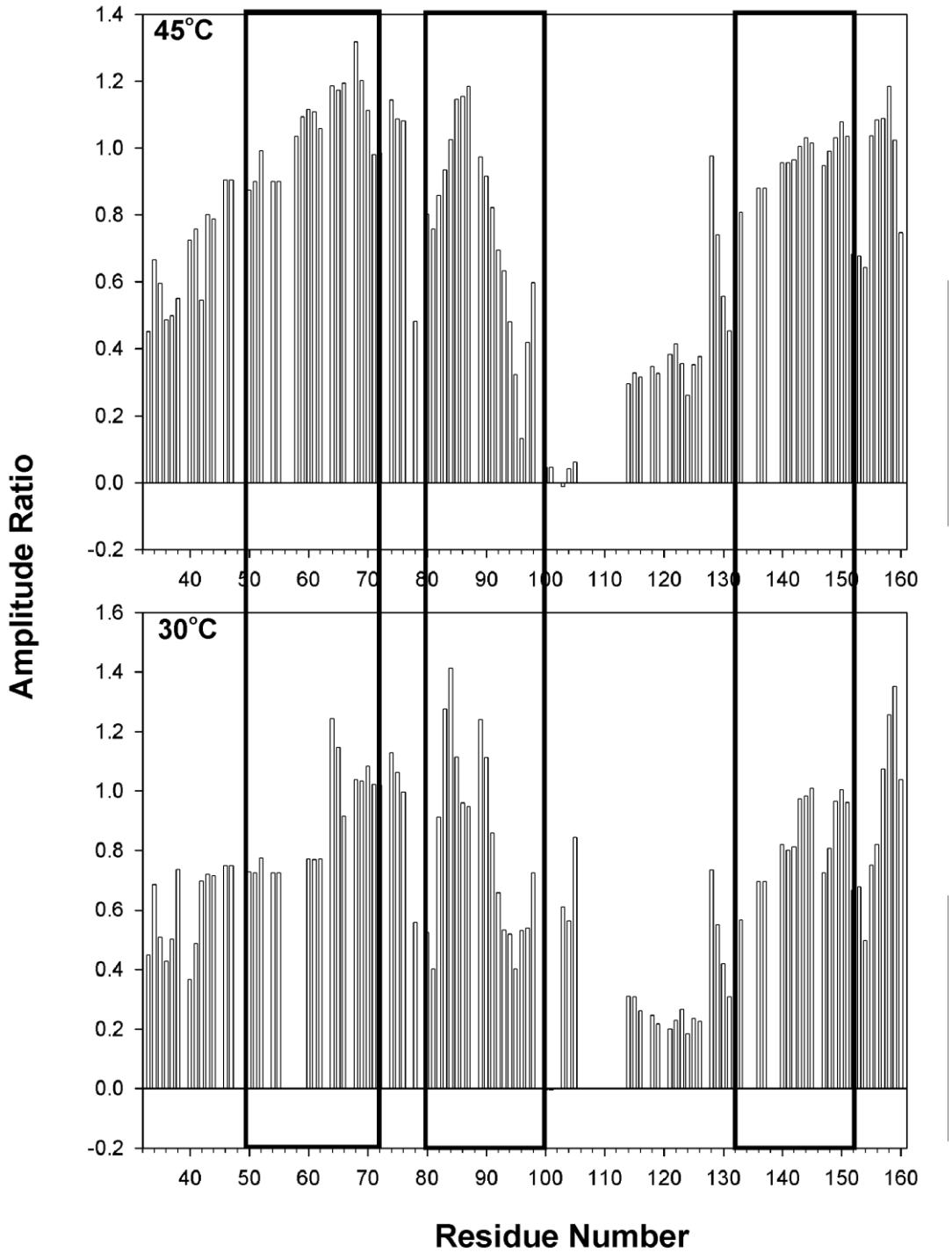

Figure 4-8: Evaluation of tertiary folding in 50% TFE:water for Ste2p TM1-TM3 (G31-R161) using PRE experiments. Peak amplitude ratios for [^{15}N,^{1}H]-HSQC experiments conducted at 45°C and 30°C in the presence and absence of a nitroxide radical spin label were used to evaluate the degree of tertiary folding of TM1-TM3. The boxed helical boundaries are those calculated by sequence-based TM prediction software (average of four programs).

NOESY Assignment and Structure Calculation for Ste2p TM1-TM3(G31-R161)

Automated [^{15}N]- and [^{13}C]-NOESY assignments were conducted using the ATNOS-CANDID component of the UNIO software suit, and five, seven-cycle CYANA structure calculations were performed with the resulting data. The twenty lowest energy structures of the last 80 structure calculation were used to evaluate the calculated structure. A total of 342 intraresidual ($|i-j|=0$), 531 sequential ($|i-j|=1$), 873 short-range ($|i-j| \leq 1$), 544 medium-range ($1<|i-j|>5$), and 5 long-range ($|i-j|\geq 5$) restraints were used in the calculation (Figure 4-9). The twenty structures had target function values ranging from 0.6 to 1.41 (mean target function 1.14 ± 0.18). A total of 10 violations were observed (nine distance restraint violations and one van der Waals violation). The calculated Ramachandran plot showed that 90.3 % of the sequence fell within the most favored regions, with the remaining 9.7% falling in the additionally allowed regions (Figure 4-10). The majority of the residues in the most favored region were in the right-handed α-helix portion of the plot, validating the results of the secondary structure evaluations described above.

MOLMOL calculation of the secondary structure from the PDB file identifies four α-helices consisting of residues 36-75, 83-111, 118-121, and 132-152. The boundaries of the first helix are consistent with the N- and C-terminal extension of the calculated boundaries of TM1. The N-terminal extension is most likely the fusion of TM1 and the observed helix in the N-terminal domain. The coalescence of these two putative helices is most likely the outcome of using TFE/water, a helix-inducing solvent system. We expect that it is likely that two distinct helices will be observed when the studies in TM1-TM3 in detergent micelles are completed.

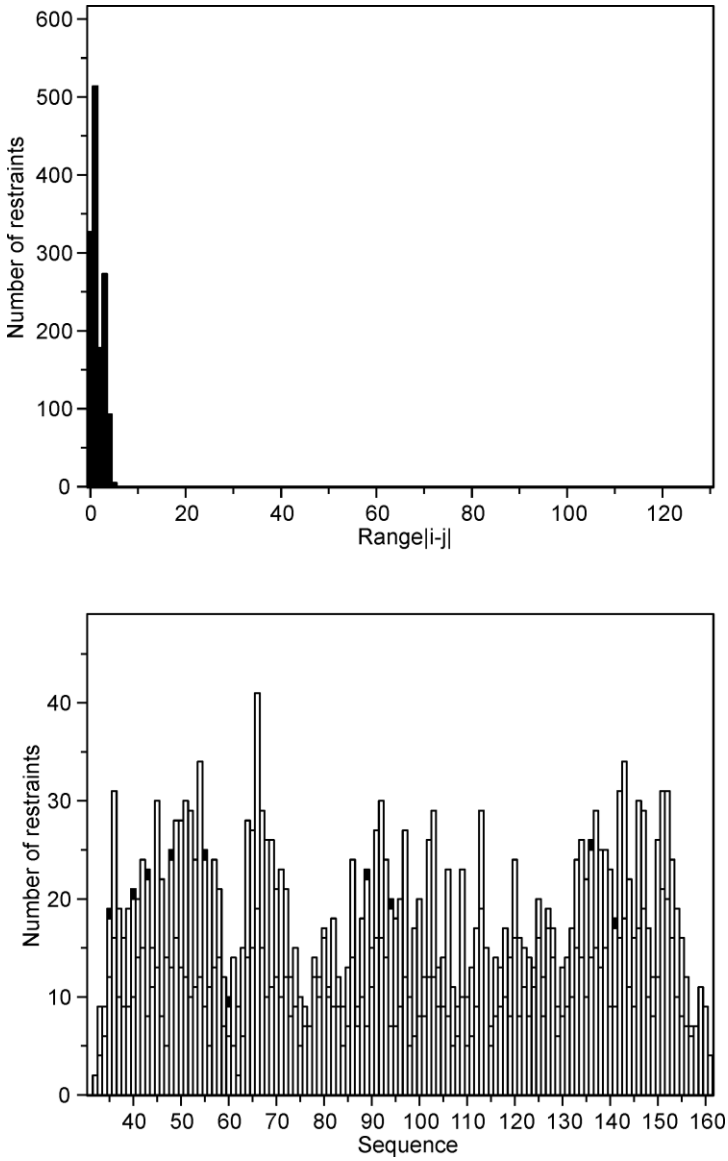

Figure 4-9: Restraints of CYANA structure calculation for Ste2p TM1-TM3 (G31-R161). The number of restraints vs. the distance between residues involved is plotted in (A). No connections were observed with distances greater than approximately 10Å. The number of restraints used for each residue is plotted in (B). White bars indicate intra-residue connectivities. Light gray bars indicate short-range connectivities. Dark gray bars indicate medium-range connectivities. Black bars indicate long-range connectivities.

Table 4-3: NMR Constraints and Structural Statistics for 20 Structures of TM1-TM3 in 50% TFE:water

Distance Constraints	1422
Short Range $\|i-j\|\leq 1$	873
Medium Range $1<\|i-j\|<5$	544
Long Range $\|i-j\|\geq 5$	5
Hydrogen Bond Constraints	80
Dihedral Angle Constraints	240

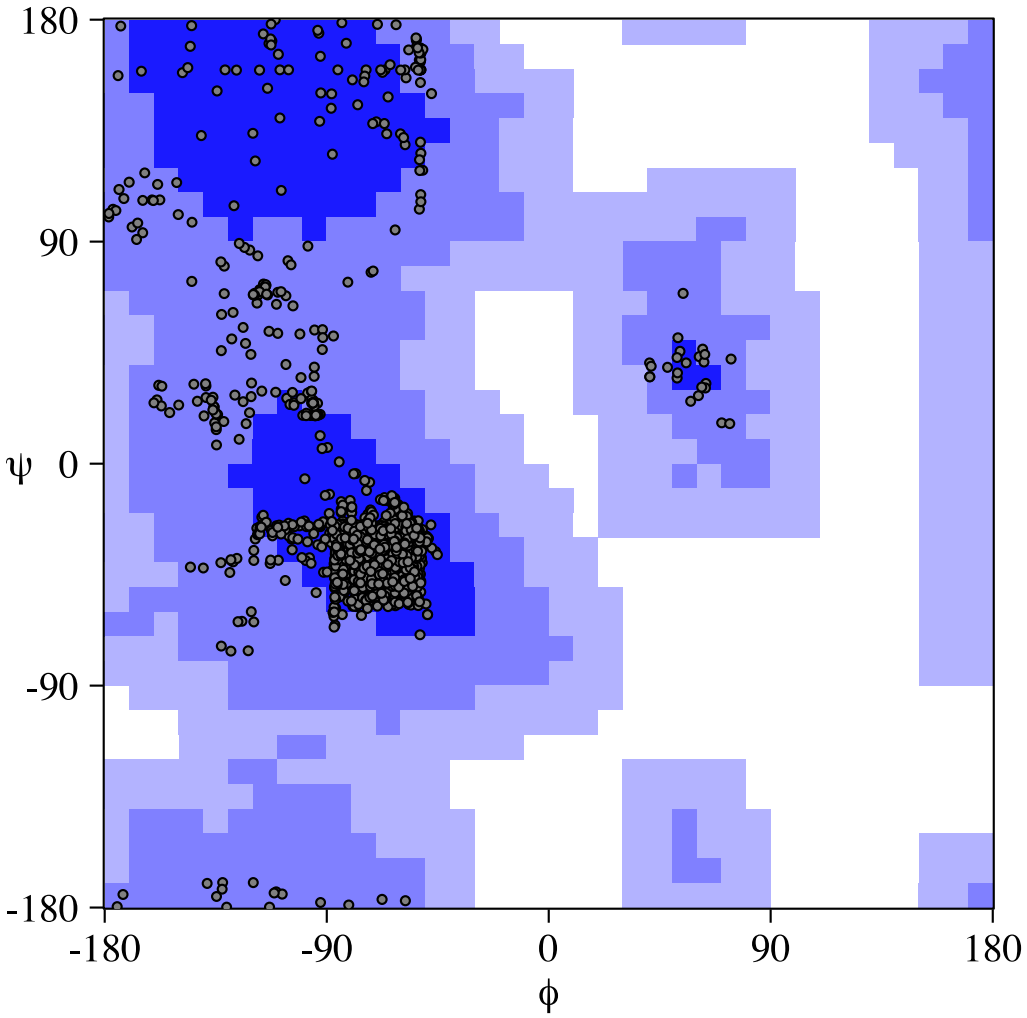

Figure 4-10: Ramachandran plot from CYANA structure calculation for Ste2p TM1-TM3 (G31-R161). 90.3% of all residues are found within the most favored regions of the plot, with the remaining 9.7% in the additionally allowed regions. The majority of the residues fall within the right-handed α-helix region of the plot.

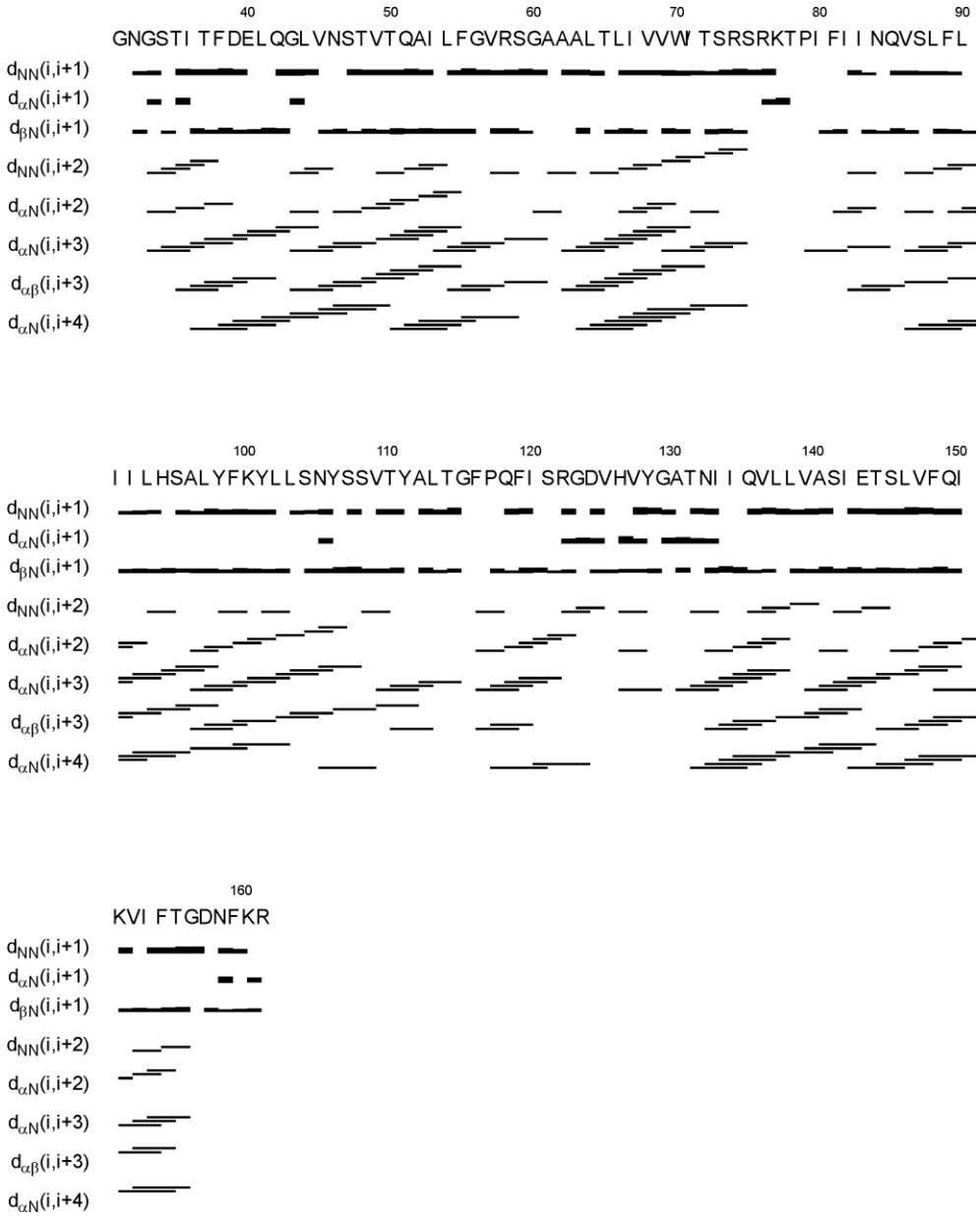

Figure 4-11: Inter-residue NOE connectivities for the CYANA structure calculation of Ste2p TM1-TM3 (G31-R161). The ATNOS-CANDID component of the UNIO software suite was used for automatic assignment of the [^{13}C]- and [^{15}N]-NOESY experiments for TM1-TM3 in 50% TFE:water. Connectivities are displayed as a function of residue. Regions of the sequence with a large number of i→i+3 and i→i+4 connectivities are considered helical.

Analysis of the NOESY connectivities identified in the N-terminus-TM1 region (Figure 4-11) reveals a large number of i→i+3 connectivities throughout this region and i→i+4 connectivities between residues 36-49, 51-59, and 63-74, with breaks in i→i+4 connectivities observed between residues 49-51 and 58-63. The reduction in i→i+4 connectivities between residues 49-51 suggests the weakening of the α-helix at this position. This is most likely the junction between the N-terminal helix and TM1. The decrease in i→i+4 connectivities in the middle of the putative TM1 (58-63) is centered around the GXXXG sequence and is therefore consistent with secondary shift, H-D exchange, and T_2 relaxation results.

The boundaries of the second calculated helix are consistent with a C-terminal extension of the calculated TM2 boundaries. This extension may also be the result of the helix-inducing solvent system, as it most likely represents a fusion of TM2 and the predicted EL1 helix (predicted boundaries for EL1 from secondary shifts and biochemical data: 108-115). Since a break in the i→i+4 connectivities is observed between residues 102-105, this is most likely the native break position between TM2 and the EL1 helix.

The small, four-residue helix observed between residues 118-121 was not predicted to exist based on secondary shifts, H-D exchange, or T_2 relaxation, and there is no biochemical data to support its existence. The boundaries of the fourth calculated helix are highly consistent with the calculated boundaries for TM3. The helix begins and ends in the exact positions predicted by the average of the sequence-based secondary structure prediction software results. There are strong i→i+4 connectivities throughout suggesting that this is a well-defined, strong α-helix.

MOLMOL was used to evaluate the RMSD values for the twenty analyzed structures. As very few long-range connectivities were identified using the NOESY data and no restraints were obtained from PRE experiments, the final calculated structures showed little convergence and a large RMSD when the entire sequence was included in the structure fitting calculation. To gain insight into the localized structural propensities of regions of TM1-TM3 we evaluated the RMSDs of portions of the complete sequence. Initial RMSD calculations were done based on the helix boundaries obtained from the structure calculation itself. In this evaluation backbone RMSD values no higher than 3 were obtained, with RMSD decreasing as a function of helix length (Table 4-3, Figure 4-12A) with the TM1/NT region having a RMSD of 3.02 for backbone atoms and TM3 having an RMSD of 0.99. To evaluate the influence of the helix boundaries on the accuracy of the structure we reduced the length of the helical regions using boundaries based on experimental NOEs or predictive algorithms. Accordingly the NT and EL1 helix boundaries deduced from the i→i+4 NOE connectivities were used for a second round of RMSD calculations (Table 4-3, Figure 4-12B) and backbone RMSD values were found to decrease for all TM domains, with TM1 exhibiting the highest RMSD of 2.36 ± 1.16, TM2 and TM3 having backbone RMSDs of 1.03 ± 0.41 and 1.49 ± 0.51, respectively and the short NT and EL1 helices having RMSDs of 0.43 ± 0.18 and 0.69 ± 0.28, respectively. The two RMSD calculations performed using the TM boundaries calculated using the sequence-based TM boundary prediction software (Table 4-3, Figure 4-12C) and those derived from the Eilers model (*71*) (Table 4-3, Figure 4-12D) and the secondary shift predicted NT and EL1 helix values gave the smallest RMSDs for all regions of TM1-TM3. This trend of decreasing RMSD is not surprising as the length of

the TM helices decreased in the successive calculations.

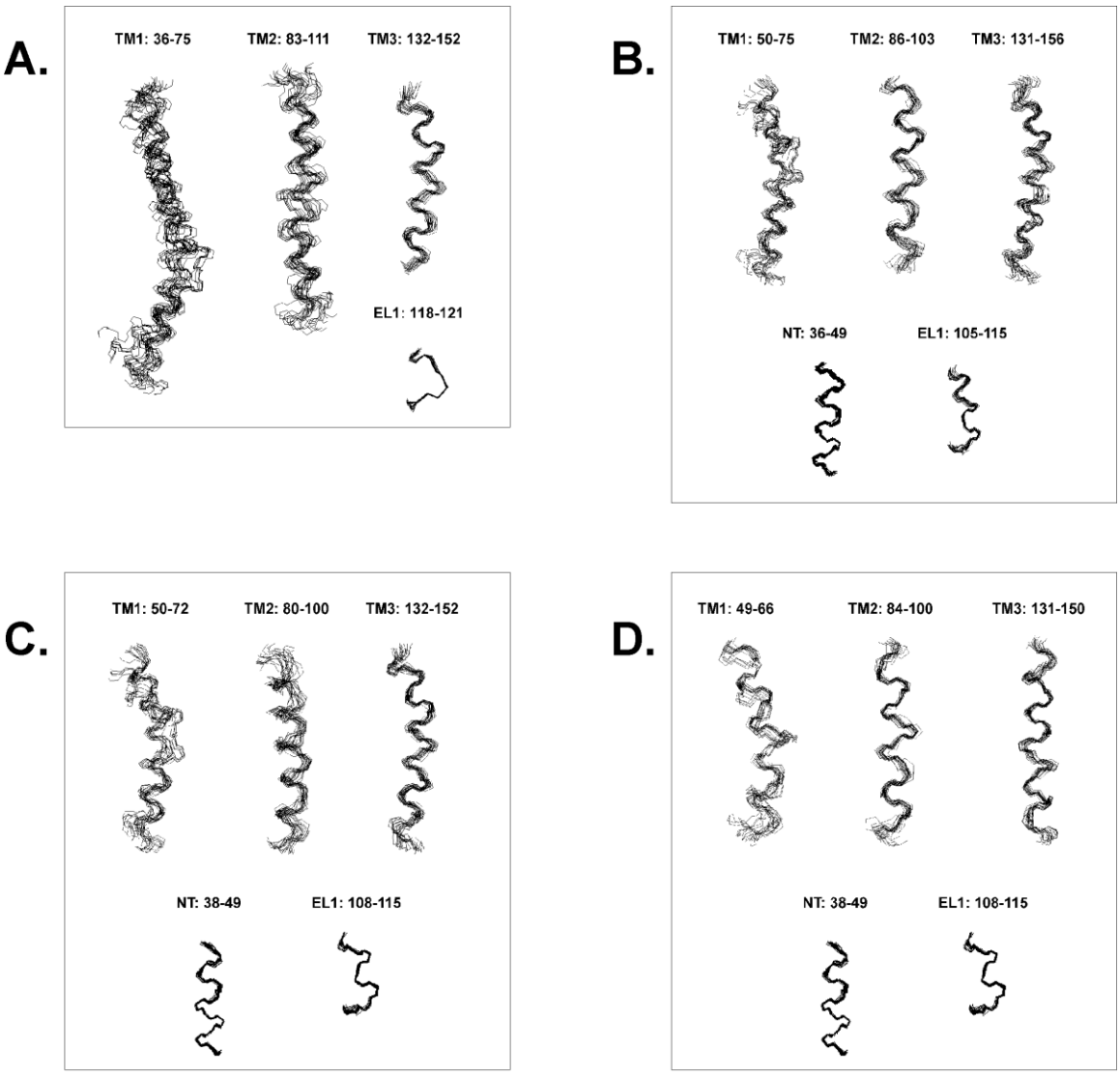

Figure 4-12: Convergence of the lowest 20 energy CYANA calculated structures for Ste2p TM1-TM3 (G31-R161). The helices were produced by fitting the helical boundaries calculated from the structure (A), those determined from inter-residue NOE connectivities (B), the calculated TM boundaries and the secondary shift boundaries for the NT and EL1 helices (C), and the template predicted TM boundaries[71] and the secondary shift boundaries for the NT and EL1 helices (D).

Table 4-4: Calculated RMSD Values for the TM1-TM3 Structure: Eilers TM Boundaries

Method	Putative Ste2p Region	Included Residues	Backbone RMSD	Heavy Atom RMSD
Structure[a]	TM1/NT	36-75	3.02 ± 1.41	3.81 ± 1.52
	TM2/EL1	83-111	2.09 ± 0.81	2.98 ± 0.84
	New EL1	118-121	0.24 ± 0.11	1.17 ± 0.31
	TM3	132-152	0.99 ± 0.29	1.64 ± 0.29
NOE[b]	TM1	50-75	2.36 ± 1.16	3.05 ± 1.13
	TM2	86-103	1.03 ± 0.41	1.98 ± 0.46
	TM3	131-156	1.49 ± 0.51	2.07 ± 0.44
	NT	36-49	0.43 ± 0.18	1.31 ± 0.21
	EL1	105-115	0.69 ± 0.28	1.53 ± 0.40
Sequence-based Calculation[c]	TM1	50-72	2.10 ± 1.09	2.77 ± 1.04
	TM2	80-100	1.47 ± 0.5	2.5 ± 0.56
	TM3	132-152	0.99 ± 0.31	1.64 ± 0.29
	NT	38-49	0.39 ± 0.18	1.33 ± 0.23
	EL1	108-115	0.43 ± 0.22	1.14 ± 0.41
Eilers[d]	TM1	49-66	1.54 ± 0.79	2.1 ± 0.71
	TM2	84-100	0.82 ± 0.31	1.8 ± 0.34
	TM3	131-150	0.85 ± 0.28	1.49 ± 0.27
	NT	38-49	0.39 ± 0.18	1.33 ± 0.23
	EL1	108-115	0.43 ± 0.22	1.14 ± 0.41

[a] Boundaries determined in MOLMOL from the calculated structure.
[b] Boundaries from the calculated structure modified based on i→i+4 NOE connectivites.
[c] Average boundaries from sequence-based TM prediction software.
[d] Boundaries from the rhodopsin-templated model of Ste2p (*71*)

All of the data used in the structure calculation was collected at 45°C. As PRE data suggests the lack of tertiary structure at 45°C, the fact that our final calculated TM structures have RMSD values no lower than 1 is not surprising. TM1 had the highest RMSD value in all four calculations, with an average backbone RMSD of 2.26 ± 1.05. This is logical as the high experimental temperature would lead to structural fluctuations around the GXXXG region. TM2 and TM3 showed smaller calculated RMSD values, with average RMSD values of 1.35 ± 0.5 and 1.08 ± 0.35, respectively, suggesting that

there are fewer structural fluctuations in these regions of the receptor. This would suggest that the T_2 relaxation data is a more accurate depiction of the secondary structure for TM2 as the H-D exchange data suggested that this helix was not well-defined. Although the observed RMSD values are not indicative of a high-resolution solution NMR structure, the fact that they do not exceed 3 in an organic-aqueous medium at high temperature would be consistent with an inherent proclivity toward structure formation for this GPCR fragment.

Conclusions

Extensive NMR characterization of the TM1-TM3 fragment of Ste2p has been conducted in 50% TFE:water. Complete backbone and sidechain assignments have been made for data collected at 45°C. The backbone assignments were used for qualitative secondary structure analysis via chemical shift differencing, H-D exchange, and T_2 relaxation and relative mobility analysis via HNOE. All analyses suggest the presence of three TM helices, with boundaries in close agreement with sequence-based calculations of TM boundaries, with additional helicity observed in the N-terminal region and the EL1. HNOE analysis suggests that certain areas of the fragment become less mobile at reduced temperatures. Reduction in the relative mobility of EL1 at 30°C would be consistent with the conclusion that the fragment adopts some degree of tertiary structure at reduced temperature. Initial PRE experiments suggest that the TM1-TM3 fragment adopts some level of transient tertiary structure at low temperature. Attempts are currently being made to repeat these experiments.

CYANA structure calculation using data collected at 45°C reveals the presence of 4 helices. Three of the four are in good agreement with the template predictions and the

secondary structure evaluations conducted using backbone assignments. Alignment of the calculated boundaries for the TM regions and the secondary shift boundaries for the NT and EL1 helices reveals acceptable convergence and relatively low RMSD values. PRE experiments will need to be repeated with a better-quality sample in order to identify specific long-range contacts. These contacts could then be used as restraints in a new structure calculation.

Our initial hypothesis was that increasing the number of TMs in a GPCR fragment would increase the probability for interhelical contacts and lead to a more defined tertiary structure in TFE:water. Comparison of the TM1-TM2 and TM1-TM3 structures in TFE:water would suggest that this hypothesis was incorrect as TM1-TM2 adopted a transient tertiary structure with observable tertiary contacts. However, all of the tertiary contacts in the TM1-TM2 structure were determined by PRE experiments. Until PRE experiments for TM1-TM3 can be repeated and accurate results can be obtained, concrete conclusions regarding the validity of the hypothesis cannot be drawn.

CHAPTER 5

Structural Characterization of Ste2p TM1 (G31-T78) and TM7CT40 (S267-S339) in LPPG:DPC Mixed Micelles and Assessment of Chemical Shift Transferability

Introduction

The motivation for the structural analysis of the Ste2p TM1-TM3 fragment was the hypothesis that an increased probability for interhelical contacts would lead to a better defined tertiary structure. Over the course of the study, it was observed that the chemical shifts for the TM1-TM2 fragment and the TM1-TM3 fragment in 50% TFE:water were very similar (Chapter 4, Figure 4-3). Based on this observation, it was hypothesized that chemical shift assignments for small fragments of GPCRs in a given membrane mimetic may aid in the assignment of larger fragments in the same system. As chemical shift assignments for the TM1-TM2, TM1-TM3, and TM127 fragments had been completed in detergent micelles, NMR analysis of small Ste2p fragments consisting of one TM were undertaken in order to test the transferability of chemical shifts. This project is a collaborative project with the Zerbe group at the University of Zurich and Güntert lab at the Frankfurt Institute of Advanced Studies of Goethe University, and has given me the opportunity to become proficient in NMR analysis of membrane protein fragments in detergent micelles. Two 1-TM fragments of Ste2p were analyzed in 4:1 LPPG:DPC micelles, as this was the membrane mimetic used for the characterization of both TM1-TM3 and TM127. Characterization of the TM1-TM2 fragment had been previously conducted in LPPG micelles (*57*), but as the mixed micelle system largely consisted of this detergent, it was decided that a direct comparison could be made.

The Ste2p TM1(G31-T78) fragment consisted of 30 residues of the N-terminal region and the first TM. This fragment was cloned as a direct expression construct with a 6-residue His tag and directly expressed in *E.coli* as described in Chapter 2. The final construct contained 55 residues, including the His tag. The TM7CT40 construct consisted of the seventh TM and 40 residues of the C-terminal tail. This construct had been previously expressed as a TrpΔLE fusion protein, and the isolated 73-residue peptide was subject to extensive NMR characterization in DPC micelles (*1, 58*). [^{15}N, ^{13}C]-labeled peptides were used for 3D NMR analysis for complete backbone and sidechain assignment in 4:1 LPPG:DPC micelles prepared using the Killian protocol (*90*). HNOE experiments were conducted in order to determine the dynamics of the fragments. Backbone and sidechain assignments for the TM1 fragment were used for a CYANA structure calculation for the 55 residue fragment. Finally, backbone chemical shift differences were calculated for comparison of the TM1 chemical shifts to those in TM1-TM2, TM127, and TM1-TM3 and the TM7 shifts to those in TM127. Chemical shift assignments for TM1, TM1-TM2, and TM127 are currently being analyzed by the Güntert lab at the Frankfurt Institute of Advanced Studies of Goethe University in order to determine the transferability of chemical shifts between fragments.

Materials and Methods

NMR Sample preparation for Ste2p TM1(G31-T78) and TM7CT40(S267-S339) in 4:1 LPPG:DPC Micelles

NMR samples for TM1 and TM7CT40 in 4:1 LPPG:DPC micelles were prepared using a modified version of the Killian protocol (*90*). Initial characterization of both fragments was conducted using [^{15}N]-labeled material. 0.8 mg of [^{15}N]-TM1 or 1 mg of [^{15}N]-TM7CT40 were solubilized in a mixture of 50μL hexafluoroisopropanol (HFIP)

and 25μL of water, and the sample was incubated at 37°C, 225 rpm for 15 minutes. A mixture of LPPG:DPC was made such that the final detergent concentrations in a 350 μL sample were 120 mM and 30 mM, respectively. The detergents were solubilized in 105 μL of 40 mM potassium phosphate buffer, pH 6.5, sonicated at room temperature at 30 W for 15 minutes, and then incubated at 37°C, 225 rpm for 15 minutes. The peptide solution was added to the detergent solution, and the sample was mixed well. The sample was placed in a 37°C water bath, and water was added in steps of 20 μL to bring the sample volume to 400 μL. Water was then added in steps of 100 μL to bring the final sample volume to 1 mL. The sample was then frozen on dry ice and lyophilized. After overnight lyophilization, the sample was removed, resuspended in 1 mL water by incubation at 37°C, 225 rpm, and lyophilized again. After the second lyophilization, 35 μL D_2O, 310 μL water, and 5 μL of 10 mM DSS were added, the sample was incubated at 37°C, 225 rpm for 30 minutes, and the clear sample was transferred to a Shigemi tube. [^{15}N]-HSQC and [^{15}N]-HSQC-TROSY experiments were conducted at 45°C, 600 MHz.

Three-Dimensional Heteronuclear NMR Experiments for Backbone and Sidechain Assignments of Ste2p TM1(G31–T78) and TM7 (S267-S339) in 4:1 LPPG:DPC micelles at 45 ℃

Triple-resonance NMR experiments were conducted for backbone and sidechain assignment of the TM1 and TM7CT40 Ste2p fragments in 4:1 LPPG:DPC micelles. All samples were prepared as described above. A 0.9 mM sample of [^{15}N,^{13}C]-TM1 or a 0.7 mM sample of [^{15}N,^{13}C]-TM7CT40 was used, respectively, for all 3D backbone experiments. [^{15}N]-HSQC, HNCO, HNCA, HN(CA)CO and HNCACB experiments were conducted for backbone assignment on a three-channel Varian NMR-S 600 MHz NMR spectrometer with a z-axis pulsed-field-gradient and a Varian 5mm [^{1}H, ^{15}N, ^{13}C,

^{2}D] cryo-probe at the College of Staten Island. [^{13}C]-HSQC, [^{1}H,^{1}H]-TOCSY, [^{15}N]- and [^{13}C]-NOESY, and HCCC(CO)NH experiments for side chain assignment were performed on a three-channel Bruker AV-700 700 MHz NMR spectrometer equipped with a CRYO TXI inverse triple resonance cryoprobe at the University of Zurich. A 0.6 mM [^{15}N,^{13}C]-TM1 sample was prepared in deuterated detergent for these experiments in order to reduce background signal from the detergent. As no additional double labeled material was available for TM7CT40, the [^{15}N,^{13}C]-TM7CT40 prepared in protonated detergent was used for the assignment of this peptide. Backbone and sidechain assignments for TM1 were conducted following the procedures described in Chapter 4 for the TM1-TM3 fragment of Ste2p. Initial attempts at backbone assignments for TM7 using NMRView5 were unsuccessful with many of the resonances unassigned. The Stripscope function of CARA was then used to automatically find connected residues. A 3D spectrum was opened in Stripscope. One spin system was chosen, and the show all predecessors/successor functions were used to find sequential connections. This was repeated using all experiments collected for all spin systems until no further assignments could be made.

Assessment of the Relative Flexibility of Ste2p TM1 (G31-T78) and TM7CT40(S267-S339) in 4:1 LPPG:DPC Mixed Micelles by HNOE Analysis

The [^{15}N]-TM1 and [^{15}N]-TM7CT40 in 4:1 LPPG:DPC micelles samples described above were subject to HNOE analysis at 45°C. The experiments were run on a three-channel Bruker Advance 600 MHz NMR spectrometer equipped with a CRYO TXI inverse triple resonance cryoprobe at the University of Zurich as described in Chapter 4. The amplitude and volume for each crosspeak in each data set was calculated using CARA. All data points for overlapping peaks were removed, and the resulting data was

plotted as a function of residue number.

NOESY Assignment and Structure Calculation for Ste2p TM1 (G31-T78)

Structure calculation for TM1 was performed as described in Chapter 4. Briefly, TALOS+ torsion angles were calculated from the backbone assignments. All chemical shifts, TALOS+ angles, and the [^{15}N]- and [^{13}C]-NOESY spectra were provided to the UNIO software suite for automated NOESY assignment by the ATNOS-CANDID component and structure calculation by the CYANA component. Seven 80 structure CYANA iterations were performed to calculate the structure of TM1 in 4:1 LPPG:DPC mixed micelles. The 20 lowest energy structures from the last CYANA cycle were used for analysis.

Chemical Shift Comparison for Ste2p TM1 (G31-T78) and TM7CT40 (S267-S339)

Backbone chemical shifts for TM1, TM1TM2, TM1-TM3, TM127, and TM7 in micelles were compiled. Chemical shift differences were calculated by subtracting the HN, N, CO, CA, and CB chemical shift for each residue in the smaller fragment from the same residue in the larger fragment. The results were plotted as a function of residue number.

Results

Backbone and Sidechain Chemical Shift Assignments for Ste2p TM1(G31-T78) and TM7CT40(S267-S339)

Three-dimensional heteronuclear experiments conducted on [^{15}N, ^{13}C]-labeled TM1 and TM7CT40 were used for backbone and sidechain assignment of these Ste2p fragments. Backbone assignments for TM1 were made using NMRView5. The wide chemical shift dispersion and limited spectral overlap for this fragment allowed for assignment of resonances for all residues except the His tag (Figure 5-1). Side chain

assignments were completed using CARA. All side chain carbons and aliphatic protons have been assigned.

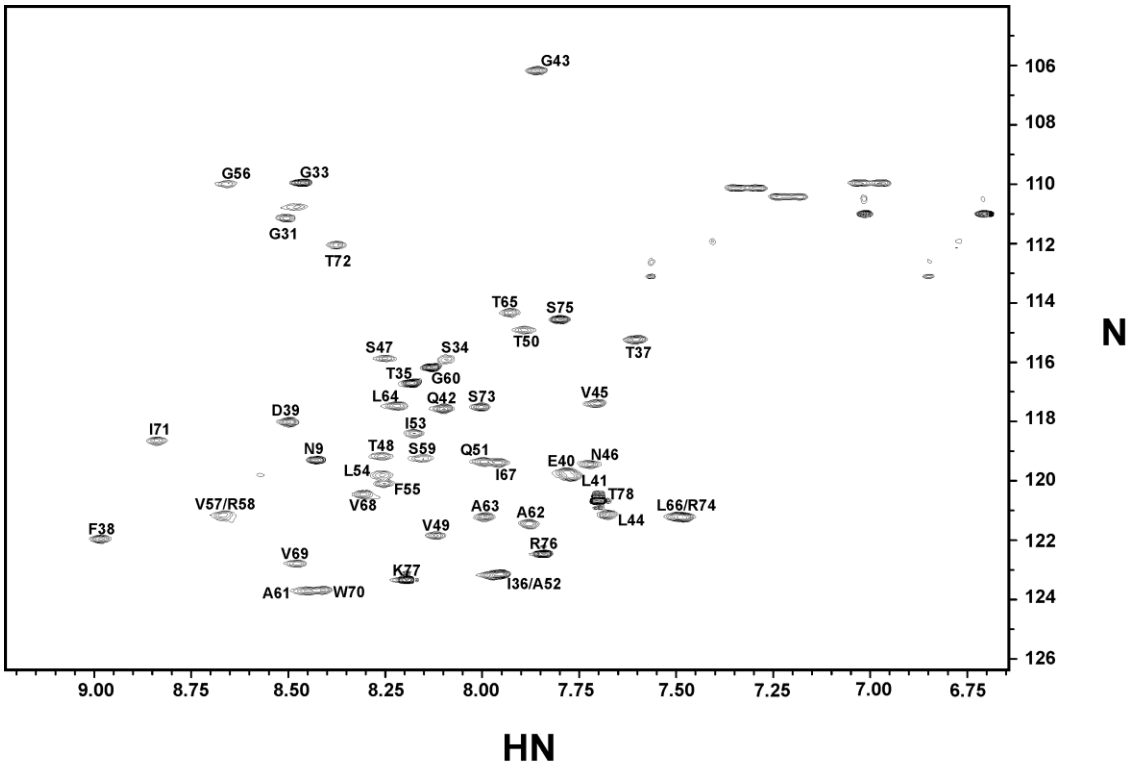

Figure 5- 1 [^{1}HN] and [^{15}N] Backbone assignments of Ste2p TM1 (G31-T78) in 4:1 LPPG:DPC micelles at 600 MHz and 45°C. The concentration of the peptide was 0.9 mM; the overall concentration of detergent was 150 mM (120 mM LPPG: 30 mM DPC). In the assignment of this spectrum 10% $D^{2}O$ was used.

Backbone assignments for TM7CT40 were complicated by the large degree of spectral overlap. Initial attempts at backbone assignments using NMRView5 were unsuccessful as only 30 of the 73 residues could be assigned using this method. CARA was then used for backbone assignments, and possible residue connections were automatically determined by the program. Correct resonances were selected from the provided possibilities, and the chemical shifts were compared to those obtained for TM7CT40 in DPC micelles. This process was repeated until no further assignments could be made. 63 of the 73 residues were successfully assigned (Figure 5-2). Of the ten unassigned residues, two also were not assigned in DPC micelles (*58*). As there was only

enough double-labeled material for one sample, experiments for side chain assignments were performed in protonated detergent. The aliphatic protons from the detergent caused a large degree of spectral overlap in the [^{13}C, ^{1}H]-HSQC spectrum. Side chain assignments have therefore not been made. Additional double-labeled material will need to be expressed and purified, and the experiments repeated in deuterated detergent.

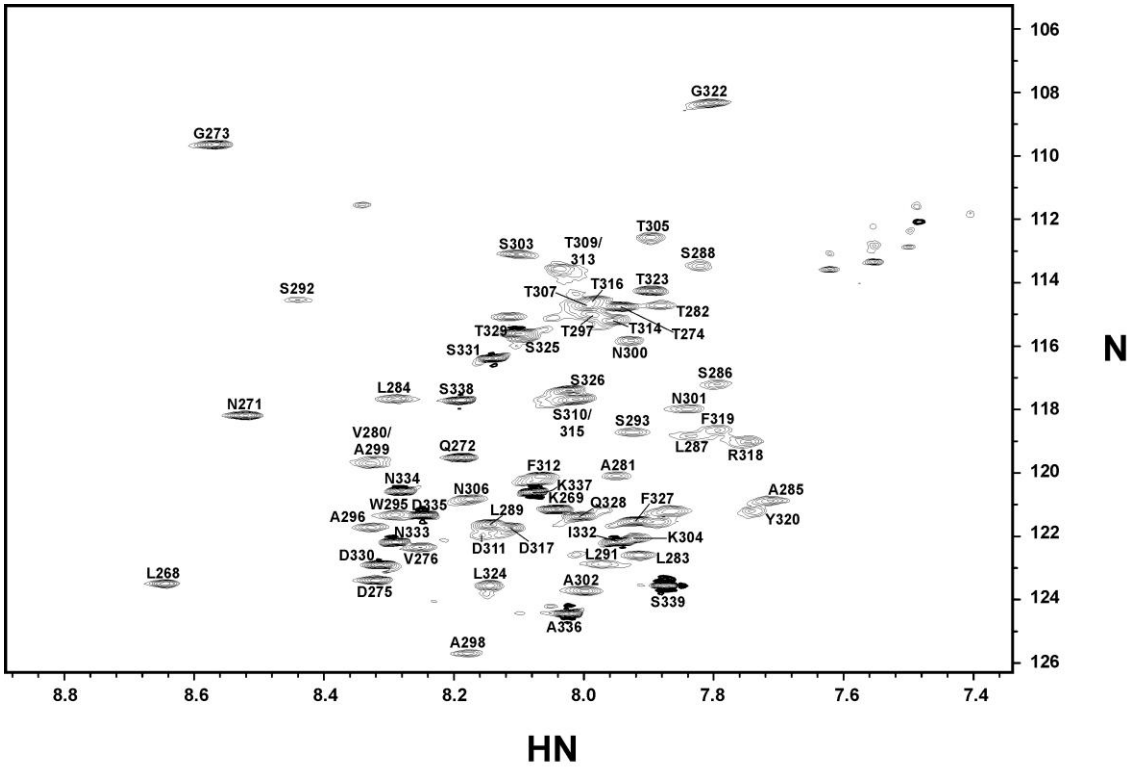

Figure 5-2 [^{1}HN] and [^{15}N] Backbone assignments of Ste2p TM7CT40 (S267-S339) in 4:1 LPPG:DPC micelles at 600 MHz and 45°C. The concentration of the peptide was 0.7 mM; the overall concentration of detergent was 150 mM (120 mM LPPG: 30 mM DPC).

Assessment of the Relative Flexibility of Ste2p TM1 (G31-T78) and TM7CT40(S267-S339) in 4:1 LPPG:DPC Mixed Micelles by HNOE Analysis

HNOE analysis was performed at 45°C to assess the relative mobility of the TM1 and TM7CT40 fragments in 4:1 LPPG:DPC micelles. Analysis of the amplitude ratios as a function of residue number (Figure 5-3) reveals that overall the TM1 fragment is less mobile than the TM7CT40 fragment, with amplitude ratios larger than 0.5 (dashed line,

Figure 5-3) for the majority of the TM1 sequence and ratios centered around or below 0.5 for the majority of the TM7CT40 sequence. This is consistent with the limited ^{1}H chemical shift dispersion in the [^{15}N,^{1}H]- HSQC spectrum for the TM7CT40 fragment. Mobility for the predicted TM regions (boxed regions, Figure 5-3) of the TM1 and TM7CT40 fragments is lower than the putative IL1 and EL3 loop regions. The N and C-terminal tail regions also show increased mobility with respect to the putative TM regions, but there is evidence that a portion of both tail regions is structured. Residues 36-45 of the TM1 fragment show decreased mobility with respect to the remainder of the N-terminal tail, suggesting that this region adopts some level of structure. This is consistent with the results for the TM1-TM2 and TM1-TM3 fragments in both TFE:water and detergent micelles, with an N-terminal helix observed in the LPPG NMR structure of TM1-TM2 between residues 39-47 (*57*). Residues 301-320 of the TM7CT40 fragment also show decreased mobility with respect to the remainder of the C-terminal tail. This could suggest the presence of helix 8, which was observed between residues 309-324 in the DPC structure of TM7CT40 (*58*). This eighth helix has also been observed in the majority of the solved GPCR crystal structures and is thought to be functionally relevant. Taken together, the HNOE results for both TM1 and TM7CT40 suggest that both fragments adopt a helical structure in the putative TM regions with additional helicity observed in the N- and C-terminal tail regions. The position of the TM regions and short NT and CT helices are in accord with those reported in previous studies (*57, 58*) and this reproducibility suggests that these structural features are present in the native receptor.

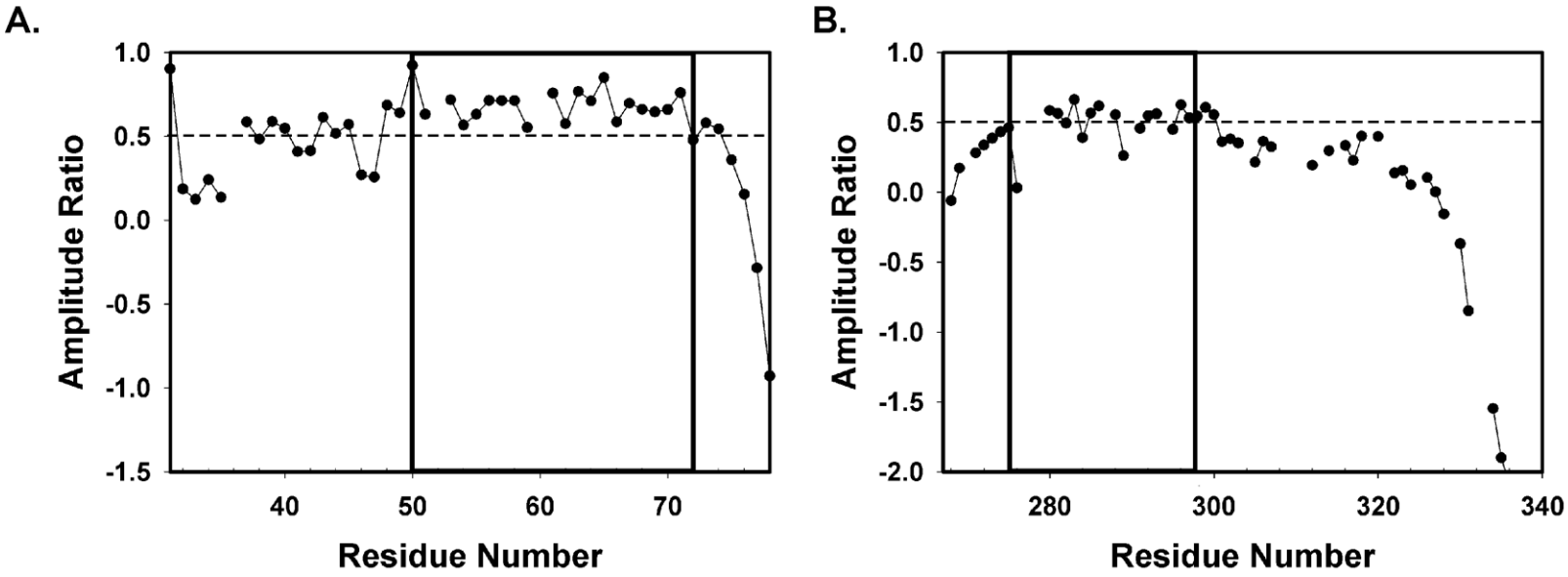

Figure 5-3: Evaluation of the relative mobility of the Ste2p TM1 (G31-T78) and TM7CT40 (S267-S339) constructs by HNOE. [^{15}N,^{1}H]-HSQC experiments were conducted with and without the addition of a pulse which generated an NOE between the amide nitrogen and proton at 45°C. The ratio of the peak amplitude with the pulse to the amplitude without the pulse for every residue in the fragments was calculated and plotted as a function of residue number. The boxed helical boundaries are the average calculated boundaries from TM prediction software.

NOESY Assignment and Structure Calculation for Ste2p TM1 (G31-T78)

Chemical shift assignments were used for a structure calculation for the TM1 (G31-T78) fragment of Ste2p. A CYANA structure calculation was performed using TALOS+ angles and ATNOS-CANDID NOESY assignments. Seven 80 structure UNIO cycles were conducted in order to come to the lowest energy structures. A total of 113 intraresidue restraints, 76 sequential restraints, 189 short-range restraints, 48 medium-range restraints, and 3 long-range restraints were used in the calculation. The only violation in the structure calculation was a single angle violation. Analysis of the lowest 20 energy structures in Molmol reveals low convergence of the full length structures (Figure 5-4, Table 5-1). A backbone RMSD of 3.06 ± 0.46 was obtained for the overall structure (Figure 5-4, Table 5-1). Analysis of the bundle reveals that the fragment adopts a highly kinked structure. There is a distinct bend in the structure centered around the GXXXG region (residues 56-60). When Molmol is used to calculate the secondary structure of the structures, three distinct helices are identified encompassing residues 42-45, 55-62, and 64-73. When these three regions of the structures are annealed, RMSD values below 0.5 are obtained for all of the helices (Table 5-1). The boundaries of the first helix are consistent with the helix in the N-terminal region observed in the structure of the TM1-TM2 fragment in LPPG micelles (*57*). The second helix is the first half of the putative TM domain and is terminated two residues after the last G of the GXXXG region. The third helix is the remainder of the putative TM domain and begins four residues after the last G of the GXXXG region. Structural fluctuations around the GXXXG region were not as pronounced in the LPPG structure of TM1-TM2 as the backbone RMSD for TM1 was found to be 0.40 ± 0.13 (*57*). This suggests that the

presence of a second TM restricts structural fluctuations about the GXXXG region.

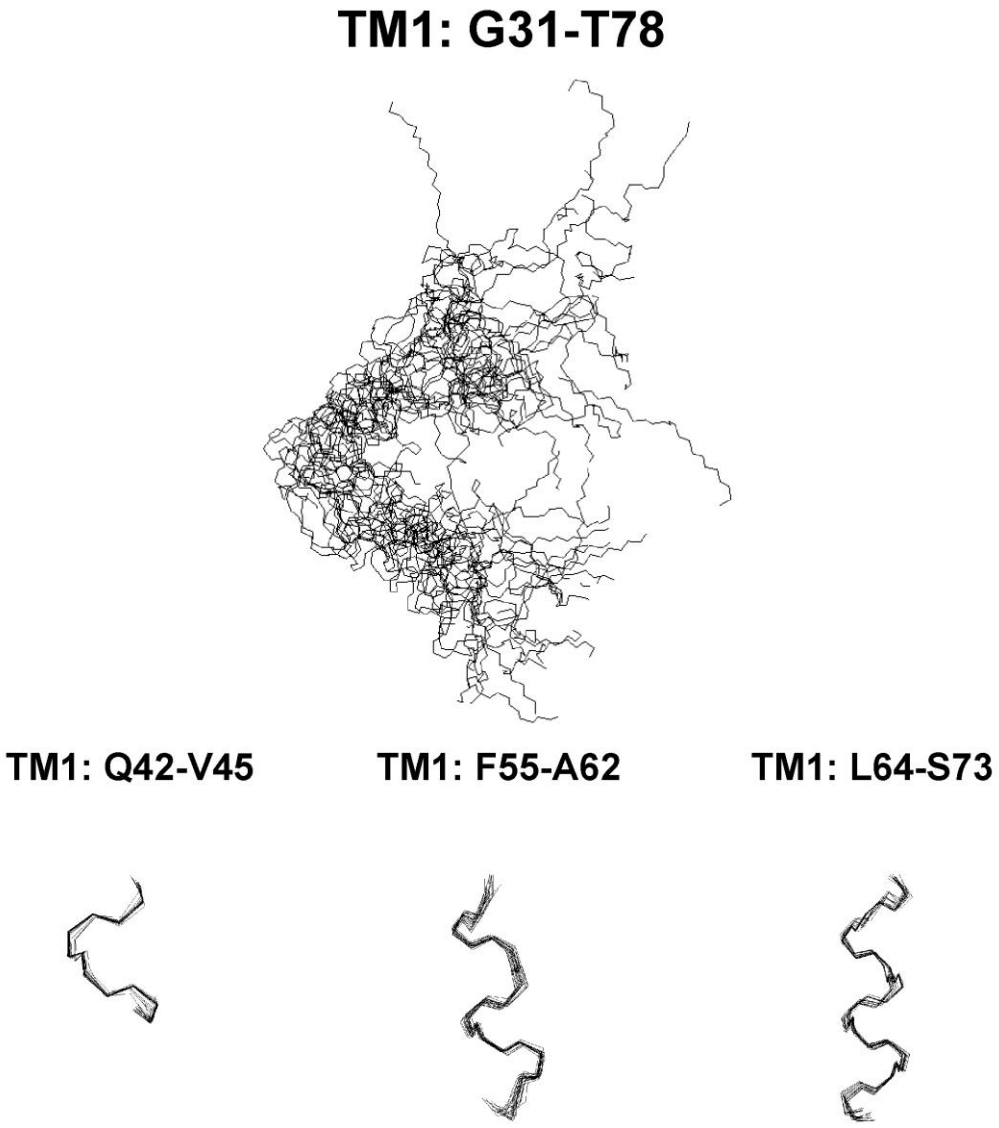

Figure 5-4: Structure of the Ste2p TM1 (G31-T78) fragment in 4:1 LPPG:DPC micelles. The overall structure (top) has low convergence and is severely kinked in the GXXXG region. Three distinct helical domains identified by MolMol are consistent with the N-terminal helix (bottom, left), the N-terminus of TM1 (bottom, middle), and the bottom of TM1 (bottom right). These distinct helices show very good convergence and low RMSD values.

Table 5-1 NMR Constraints and Structural Statistics for 20 Structures of TM1 in 4:1 LPPG:DPC micelles

Distance Restraints	240
Short Range $\|i-j\|\leq 1$	189
Medium Range $1<\|i-j\|<5$	48
Long Range $\|i-j\|\geq 5$	3
Dihedral Angle Constraints	110

Table 5-2: Calculated RMSD Values for the TM1 (G31-T78) Structure

Included Residues	Backbone RMSD	Heavy Atom RMSD
31-78	3.06 ± 0.46	4.21 ± 0.41
42-45	0.19 ± 0.12	1.18 ± 0.35
55-62	0.43 ± 0.15	1.42 ± 0.27
64-73	0.46 ± 0.25	1.33 ± 0.31

Chemical Shift Comparison for Ste2p TM1 (G31-T78) and TM7CT40 (S267-S339)

In addition to gaining experience in NMR analysis of membrane peptides in detergents and determining the structure of TM1 of Ste2p in micelles an important goal of the analysis of TM1 and TM7CT40 in LPPG:DPC micelles was to assess the transferability of chemical shifts from short fragments to long fragments. The Güntert lab is currently developing software for this purpose. They are analyzing the chemical shifts for TM1, TM7 and TM127 in 4:1 LPPG:DPC and TM1-TM2 in LPPG in order to determine whether the chemical shifts can be considered transferable. A simple assessment of the assigned chemical shifts was conducted in order to determine whether the efforts of the Güntert lab are reasonable. Differences in chemical shifts were calculated for TM1, TM7, TM1-TM2, TM1-TM3, and TM127. In all comparisons, the chemical shift for a given atom in the shorter fragment was subtracted from that of the longer fragment, and the results were plotted as a function of residue number. Upon inspection of the CO, CA, CB, and N chemical shifts, it is obvious that the chemical shifts may be calibrated differently (data not shown). This is reasonable as all of these nuclei are calibrated indirectly. As a result, no concrete conclusions can be drawn from the analysis of these nuclei and such conclusions must await the more detailed analysis of the Güntert laboratory. The only nucleus calibrated using a standard was the amide proton and therefore I was able to compare these chemical shifts without the

complication of referencing differences. Amide chemical shift difference comparisons between the TM1, TM1-TM2, and TM1-TM3 fragments reveals an interesting pattern (Figure 5-5). Comparison of the chemical shift differences for TM1 to TM1-TM2 and TM1-TM3 (Figure 5-5 A and B, respectively) reveals that all residues have differences no larger than 0.2 ppm except for residues in the GXXXG region which exhibit differences of approximately 0.4. This pattern is not observed when the chemical shifts for TM1-TM2 and TM1-TM3 are compared (Figure 5-5C). This suggests that the GXXXG region adopts a drastically different conformation in the TM1 fragment than in the TM1-TM2 and TM1-TM3 fragments. Analysis of the chemical shift differences for the TM1, TM7, and TM127 fragments reveals a similar pattern (Figure 5-6). The majority of the chemical shift differences for both TM1 and TM7CT40 compared to TM127 are around 0.2. The largest differences in TM1 shifts are in the GXXXG region, with difference values of 0.4 (Figure 5-6A). There is one outlier with a difference of 0.4 in the TM7 comparison plot (Figure 5-6B), which may be an indication of assignment error. Analysis of the TM1TM2 to TM127 comparison plot reveals that the chemical shifts are nearly identical throughout the overlapping sequence, suggesting that these two fragments adopt a very similar structure with respect to the GXXXG region and that the presence of additional TM helices restricts motion about this flexible region. The chemical shift differences surrounding the GXXXG region in all cases is consistent with the large degree of flexibility observed in the calculated TM1 structure described above and suggests that chemical shift transfer from small to large fragments may be possible in structured regions, but will be difficult for flexible regions.

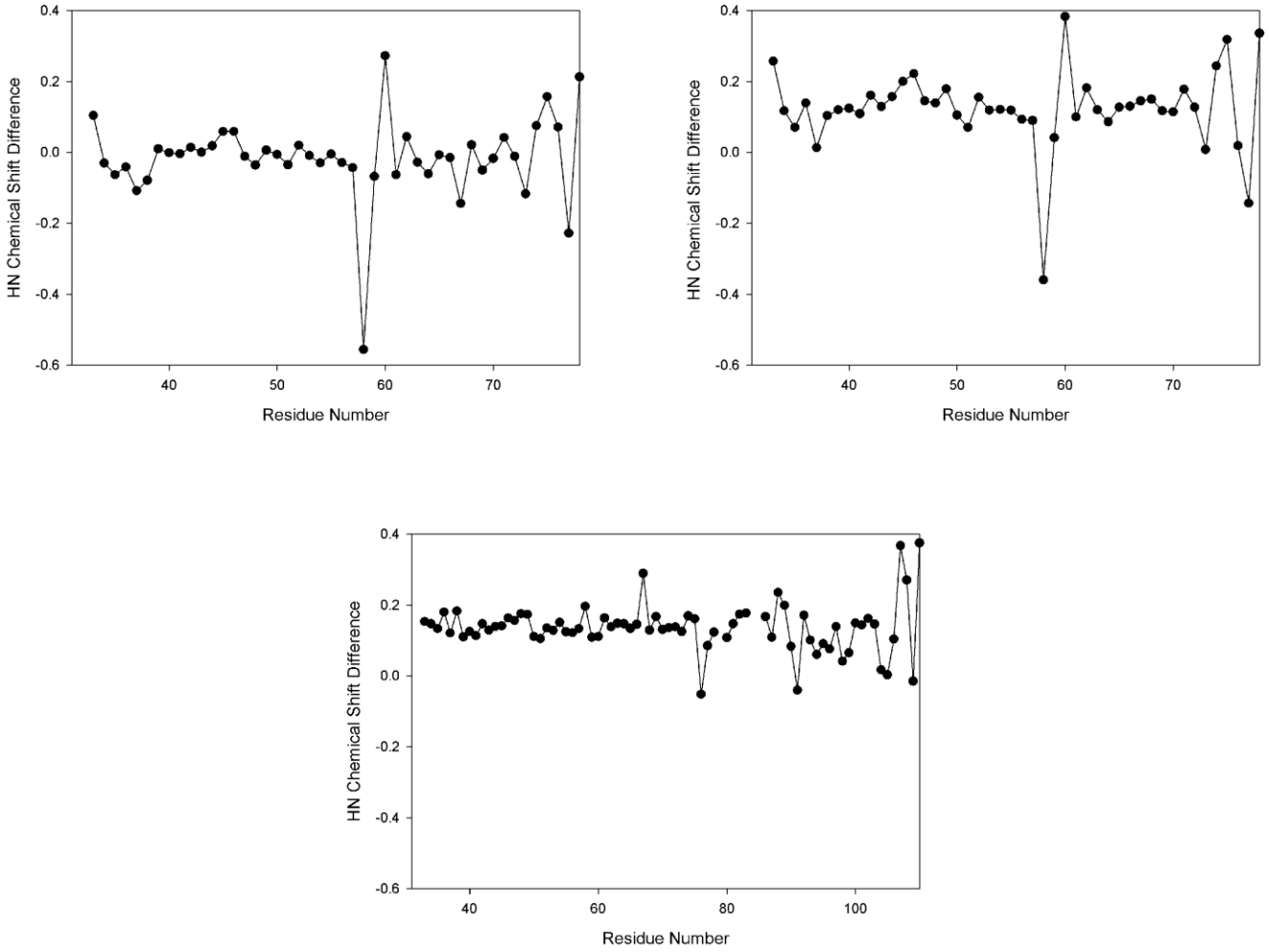

Figure 5-5: Comparison of amide proton chemical shifts for TM1 and TM1-TM2 (A), TM1 and TM1-TM3 (B), and TM1-TM2 and TM1-TM3 (C). Chemical shift differences were calculated by subtracting the value for the atom in the smaller fragment from the value of the same atom in the larger fragment.

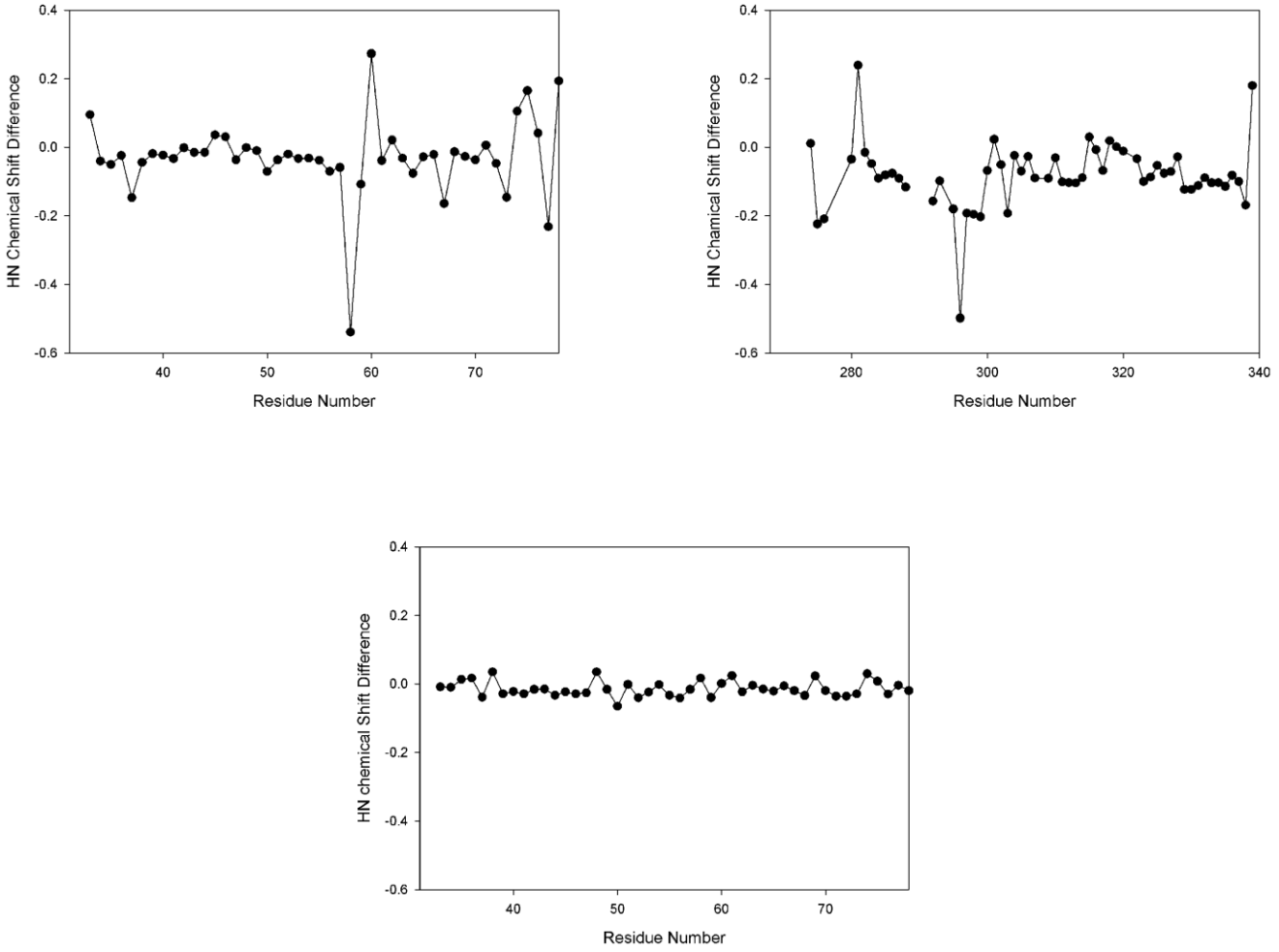

Figure 5-6: Comparison of amide proton chemical shifts for TM1 and TM127 (A), TM7 and TM127 (B), and TM1-TM2 and TM127 (C). Chemical shift differences were calculated by subtracting the value for the atom in the smaller fragment from the value of the same atom in the larger fragment

Conclusions

Analysis of the TM1 and TM7CT40 fragments of Ste2p in 4:1 LPPG:DPC micelles was undertaken to gain experience with NMR analysis in detergent micelles and to assess the chemical shift transferability between GPCR fragments of increasing size. The backbone assignments and HNOE analysis for both TM1 and TM7CT40 and side chain assignment and structure calculation for TM1 in detergent micelles were successfully completed and indicate that the TM1 fragment of Ste2p adopts a helical structure in the transmembrane region with possible helical regions in the N-terminal tail. Overall the helix in this single TM fragment is less stable than those in the same region of 2TM and 3TM fragments containing the identical sequence. Comparison of the amide proton chemical shifts for TM1, TM7CT40, TM1-TM2, TM1-TM3, and TM127 suggests that these assignments may be transferrable among fragments of increasing size for all structured regions. Chemical shifts in flexible regions, such as the GXXXG region of TM1, exhibit large differences and may therefore not be transferrable. This means that chemical shift assignments for small fragments may indeed aid in the assignment of larger fragments, but complete transferability may not be possible. Assessment of the chemical shifts by the Güntert lab will reveal whether this pattern is true for all nuclei.

CHAPTER 6

Conclusions and Future Directions

Solution-state NMR studies are desirable for GPCRs because they will provide valuable dynamic and structural information. However, such studies are hindered by the large size of the receptor-membrane mimetic complex, spectral overlap and the tendency of these proteins to aggregate. Our lab has been studying fragments of the GPCR Ste2p in order to determine whether small fragments of GPCRs can fold into native-like tertiary structures independently and whether the structure of a full-length receptor can be built using fragments. The two main goals of this dissertation research were to determine (1) whether GPCR fragments containing three TM domains form more defined tertiary structures than smaller fragments containing 2 TM domains and (2) whether chemical shifts from small fragments can be adapted to assign large fragments. To this end, two 3TM containing fragments (TM1-TM3 and TM127) and one 1TM containing fragment (TM1) of Ste2p were cloned, expressed, and purified. The 3TM fragments were subjected to extensive biophysical analysis culminating with a detailed structural analysis of the TM1-TM3 fragment in 50% TFE:water. The expressed TM1 fragment and the previously expressed TM7CT40 construct were subjected to NMR analysis in 4:1 LPPG:DPC micelles. A structure was determined for the TM1 fragment. Chemical shift assignments for all Ste2p fragments are currently being analyzed to determine the transferability of shifts from small fragments to larger fragments.

Biosynthesis of TM1-TM3, TM127, and TM1

Expression of the TM1-TM3, TM127, and TM1 constructs was accomplished using a combination of *E.coli* expression as TrpΔLE fusion proteins and direct

expression. Both the 3TM fragments were best expressed as TrpΔLE fusion proteins, with yields of uniformly [^{15}N]-labeled purified CNBr cleavage products of 15-20 mg/L and up to 11 mg/L of culture volume for the TM1-TM3 and TM127 fragments, respectively. These are very high yields for membrane protein fragments of larger than 125 residues. Similar studies conducted by other groups on GPCR fragments of up to 80 residues resulted in yields of no more than 3 mg/L of culture volume (*70, 84, 85*). The fact that I obtained yields of up to five times those of most previous studies is a significant contribution to the field. Such yields were obtained because I spent significant time optimizing the expression conditions and expression strains used to generate the fusion proteins or directly obtained material. Recently I have found that direct expression from the first natural amino acid of Ste2p (position 1) was much higher than for the same fragment starting at residue 30. The success in these studies should allow other laboratories working with membrane protein fragments to generate the tens of milligram quantities that are needed for NMR condition optimization or solid-state NMR studies. This should hopefully increase the rate of progress on structural and dynamic analysis of this understudied class of proteins.

Expression of multi-milligram quantities of large GPCR fragments in multiple labeling schemes for NMR analysis has never been accomplished for fragments of this size. The 80-residue TM1-TM2 fragment analyzed by our lab was previously the largest fragment labeled with ^{15}N, ^{13}C, and ^{2}H obtained in high yield. The 131-residue TM1-TM3 fragment, which is approximately 50 residues longer than the TM1-TM2 fragment, was obtained in similar quantities. This suggests that our expression strategy is an

effective method for the expression of Ste2p fragments of varying lengths, and that it may be used to express even longer fragments in the future.

Direct expression of the TM1-TM3 fragment resulted in the isolation of approximately 5 mg of purified peptide for PRE experiments. The quality of this isolated product was not as high as that obtained from CNBr cleavage. Direct expression and purification of TM1-TM3 was not reproducible, and additional material has yet to be obtained. However, as I indicated above, in a recent collaborative study with Dr. Leah Cohen in our research group I have been able to greatly increase direct expression yields by constructing a fragment initiating from the natural codon of the first residue of Ste2p. Our biology counterpart at the University of Tennessee has also found that direct expression of the entire Ste2p from amino acid one is also much more efficient than expressing an N-terminal truncated variant starting at residue 30 (Professor Jeffery M. Becker, personal communication). In contrast, direct expression of the TM1 construct which also initiated at G31 was highly successful and reproducible, with yields of [^{15}N,^{13}C]-labeled peptide of approximately 11 mg/L of culture volume. The difference in expression profiles for all of these fragments suggests that direct expression must be carefully optimized for each fragment of interest, and in particular the point of initiation may be a critical parameter in the protein optimization. In previous studies on GPCR fragments, to my knowledge, attention has not been paid to this variable.

Biophysical Analysis of TM1-TM3 and TM127

Biophysical characterization of the cleaved and purified TM1-TM3 and TM127 constructs was conducted using CD and NMR spectroscopy in both TFE:water mixtures

and detergent micelles. CD analysis was used as both a screen for solubility and a qualitative secondary structure analysis.

Both constructs were found to be highly helical in TFE:water mixtures, however the TM127 fragment exhibited limited solubility in this solvent system. The trend of increasing helicity with increasing TFE content was expected, as TFE is a strong helix-inducing solvent. 50% TFE:water was chosen as a suitable solvent system for first attempts at NMR structural analysis of the TM1-TM3 fragment. TM127 was not extensively analyzed in TFE:water due to the need for very high TFE concentrations for this construct. The use of TFE/water as a membrane mimetic is controversial. Our group has argued previously that valuable information on membrane protein structure and insights into structural tendencies can nevertheless be gained form such studies (*55, 57*). This assumption has been accepted in my choice to conduct studies in this organic aqueous medium.

CD investigations in detergent micelles revealed an interesting trend. Both the TM1-TM3 and TM127 constructs were found to be helical in a number of detergent micelles, but detergents with carbon chains of at least 14 carbon atoms and a glycerol head group were preferred. Fractional helicities (40-50%) predicted for the native-like state (*71*) were obtained in the most favored detergents, and LMPG and LSPG micelles were chosen as suitable detergents for first-attempt NMR investigations of TM1-TM3 and TM127, respectively, in micelles.

Initial [^{15}N,^{1}H]-HSQC analysis of TM1-TM3 in 50% TFE:water produced high-quality spectra with relatively wide peak dispersion in the [^{1}H]-dimension and the correct number (129) of visible crosspeaks. Optimal spectral resolution was obtained at 45°C,

and all subsequent analyses in this solvent system were conducted at this temperature. Complete backbone assignments allowed for the qualitative analysis of the secondary structure of this construct via NMR. Regions of helicity with distinct breaks were observed using the chemical shift differencing approach. The helical regions were consistent with the three TM regions of this fragment. Extended helicity observed in the NT-region and the EL1 were expected based on previous structural analysis of the TM1-TM2 fragment in both TFE/water and detergent micelles(*55, 57*) and Cys accessibility studies on the EL1 (*78*). These results were confirmed with both H-D exchange and T_2-relaxation experiments. The presence of breaks in helicity at loop positions and certain critical residues in the transmembrane regions would be consistent with the conclusion that the TM1-TM3 fragment is adopting a native-like structure in 50% TFE:water, and the helix-inducing solvent system is not creating artifacts.

[^{15}N,^{1}H]- HSQC analysis of TM1-TM3 in LMPG micelles revealed diminished spectral resolution compared to spectra collected in TFE:water. However, the peak dispersion and separation was determined to be suitable for structural determination. Extensive [^{15}N,^{1}H]- HSQC analysis was performed for TM1-TM3 in LMPG micelles in order to determine the optimal sample conditions. This was necessary because the sample lifetime was found to be unsatisfactory for 3D NMR experiments under the original sample conditions. After screening both pH and buffer composition, LMPG micelles prepared in HEPES buffer at pH 5.6 were determined to be the optimal sample conditions. These preliminary studies on TM1-TM3 and on TM127 in our laboratory were communicated to our collaborator, Professor Oliver Zerbe at the University of Zurich along with material that I prepared in various labeling patterns.

Initial NMR characterization of TM127 in LSPG micelles was discouraging. Sample preparation was not reproducible. Peak dispersion was limited and there was a large degree of spectral overlap. A large majority of the expected peaks were not visible in the spectra. Further optimization of sample composition and preparation conducted by Martin Poms at the University of Zurich led to the establishment of 4:1 LPPG:DPC mixed micelles prepared using the Killian protocol as the optimal sample conditions for structure determination for both TM1-TM3 and TM127. Martin is currently working on structural characterization of both fragments under these conditions. He has run into serious difficulties with the TM1-TM3 fragment and has yet to complete backbone assignments for this fragment. Backbone assignments for the TM127 construct have been completed and sidechain assignment is underway. His results are being used in the examination of the transfer of NMR assignments between homologous fragments of different lengths (see below).

Structural Analysis of TM1-TM3 and TM1

Detailed NMR structural analysis was completed for the TM1-TM3 construct in 50% TFE:water. A series of three-dimensional experiments were combined in order to complete sidechain assignments. HNOE analysis at 45°C and 30°C reveals three structured regions of the fragment, with the degree of flexibility decreasing at lower temperature. No tertiary contacts were observed in NOESY experiments. This is not surprising as all NOESY experiments were conducted at 45°C. PRE experiments conducted at 45°C and 35°C provide a preliminary indication of tertiary contacts between the MTSL label in the EL1 with residues in both TM1 and TM3. However, the spectral resolution obtained during these experiments were poor and attempts are being made to

repeat them. The final calculated structure has a low target function and very few violations. Although no tertiary contacts are present in the structure, all TM helices are apparent and resolved to RMSD values between 0.99 ± 0.31 and 3.02 ± 1.41 depending on the length of the helix and the boundaries used in the calculated fit. Attempts will continue to be made to improve this structure with tertiary contacts, however it may be that the TM1-TM3 fragment is only able to adopt a transient structure in TFE:water. My feeling is that the structure in LPPG:DPC micelles will be more defined due to the lateral pressure from the detergent micelles. Our original hypothesis was that increasing the size of the Ste2p fragment to contain a third transmembrane domain would raise the probability for tertiary contacts in the 3TM fragment and that this would lead to a more defined NMR structure than the 2TM fragment in TFE:water. This hypothesis has not yet been validated due to lack of PRE contacts. It may be that the long EL1 loop between TM2 and TM3 in the TM1-TM3 provides a large entropic advantage that leads to high flexibility. In the native receptor contacts between the three extracellular loops and the tail might provide enthalpic contributions that decrease the flexibility and favor TM-TM contacts. Given my findings it is my belief that is prudent to carry out further structural analysis on GPCR fragments in detergent micelles.

Structural analysis of the 55-residue TM1 fragment was undertaken in order to gain experience with structure determination in detergent micelles and to explore the chemical shift transferability between fragments. Complete backbone and sidechain assignments were made for this fragment in 4:1 LPPG:DPC micelles. HNOE analysis revealed that the most structured region of this fragment was the putative TM region. The final calculated structure had no violations and a low target function, but showed

little overall convergence. This is due to a large kink in the middle of the TM involving the GXXXG region. Three distinct helical regions were identified that corresponded to the NT-helix, TM1 before the GXXXG and TM1 after the GXXXG. These distinct helices all have RMSD values lower than 0.5. It is my belief that interactions with a second TM stabilize TM1 by allowing the GXXXG region to pack against interacting residues. This causes TM1 to become less flexible allowing the determination of a better converged structure for this region of the receptor. In Ste2p the GXXXG sequence in TM1 was shown to be a dimerization locus indicating that it does pack against other regions of the receptor (*75, 119*). This supports our working hypothesis that fragments of GPCRs containing 2TM helices would both stabilize the individual helices and be the smallest unit that can provide detailed structural information pertaining to the 3D shape of a region of the receptor. The Deber group has also shown that two TM constructs appear to fold into tertiary structures using gel analysis and various spectroscopic techniques (*120-126*).

Chemical Shift Transferability

NMR assignments for the TM1 and TM7CT40 fragments were made in 4:1 LPPG:DPC micelles in order to determine the chemical shift transferability of smaller fragments to larger fragments in the same membrane environment. Although backbone and sidechain assignments for the TM1 fragment were completed, complete assignments for the TM7CT40 fragment have not been obtained. The spectra for this fragment showed a large degree of overlap, and backbone assignments for ten residues are missing. Sidechain assignments have yet to be attempted due to the lack of double labeled

material. I will continue to work on this, and hope to complete sidechain assignments in the near future.

Although some assignments are still missing for the TM7CT40 construct, first attempts at the determination of chemical shift transferability are being made. A simple comparison of amide proton chemical shifts reveals that the assignments across all of the fragments considered appear to be very similar except for the GXXXG region. The assignments in this region differ greatly when compared to the 2TM fragment and the 3TM fragments containing TM1. Upon inspection of the comparison of the 2TM assignments to the two 3TM fragments, these differences are not observed. This suggests that the GXXXG region is in a very different conformation in the 1TM fragment and that the smallest possible fragment for GPCR structural analysis would contain 2TM domains. Detailed computational analysis of the chemical shift assignments is currently underway by the Güntert group, and we hope to have promising results sometime in the near future.

Future Directions

Our initial motivation for this project was to obtain a high resolution structure of a large GPCR fragment in TFE:water. Working in organic:aqueous membrane mimetics is favored over detergent micelles from the viewpoint of spectral resolution and sample stability. Linewidths and spectral dispersion are superior in TFE:water, and therefore assignments are easier to make. We believed that we would have more success with the 3TM fragment than the 2TM fragment due to a higher probability of tertiary contacts. Our structure determination seems to contradict that hypothesis as a transient, low-resolution structure was obtained. Our lab has now shifted our focus away from fragment

studies in TFE:water and toward the NMR analysis of a GPCR fragment in the context of the full receptor.

A high-resolution structure for the TM1-TM2 fragment of Ste2p has been determined in LPPG micelles. Our new question is whether that structure is the native structure of this region of the receptor. In order to answer that question, we have set out to create a segmentally labeled receptor. We have developed a protocol to link the TM1-TM2 fragment to the last 5TMs of Ste2p through a disulfide bond in in the EL1 portion of the receptor. By expressing the 2TM and 5TM fragments separately, we can isotopically label the 2TM fragment such that it will be NMR visible while keeping the 5TM fragment unlabeled and thus NMR invisible. Moreover, in theory it might be possible to do a binding assay with the reconstituted receptor thereby demonstrating its biological relevance. This will allow for the NMR characterization of the 2TM fragment within the context of the entire, hopefully biologically active, receptor. If structure determination is successful, we will be able to compare the isolated structure to this structure and assess the validity of structure determination of isolated GPCR fragments.

The designed ligation strategy is based on a Cys activation approach. We activate one of the Cys residues for disulfide bond formation by attachment of 5-5'-dithiobis(2-nitrobenzoic acid) (Ellman's reagent), 2,2'-dithiobis(5-nitropyridine), or 4,4'-dipyridyl disulfide. We also included oppositely charged residues surrounding the two Cys residues. By activating one Cys and creating an ionic attraction, we believed that we would increase the probability of herterodimer formation over homodimer formation, and have dubbed this approach guided reconstitution. Leah Cohen has been the lead investigator on this project and I have helped in the expression of certain fragments.

Initial attempts at reconstitution with small peptides and two TM containing polypeptides have been very encouraging. After I finish my dissertation defense I will concentrate on preparing the 5TM construct needed for the reconstitution of the complete receptor.

APPENDIX A

pKC01 Sequence

Below is the sequence of the expressed region of pKC01. The region is bold corresponds to the TrpΔLE. The lowercase, underlined, and italicized region corresponds to the Ste2p TM1-TM3 region (G31-R161).

ATGCATCACCATCACCATCACCATCACCATAAAGCAATTTTCGTACTGAAAGGTTCACTG
GACAGAGATCTCGACAGCCGTATTGAACTGGAACTGCGTACCGATCATAAAGAGCTGTC
TGAACATCTGCTGCTGGTTGATCTCGCCCGTAATGATCTGGCACGCATTGCTACCCCCG
GCAGCCGCTACGTCGCCGATCTCACCAAAGTTGACCGTTATTCCTATGTGCTGCACCTC
GTCTCTCGCGTAGTCGGCGAACTGCGTCACGATCTTGACGCCCTGCACGCTTATCGCGC
CGCTCTGAATCTGGGGACGTTAAGCGGTGCGCCGAAAGTACGCGCTAAGCTTTGG*atgggg*
aatggatctaccatcactttcgatgagttgcaaggtttagttaacagtactgttactcaggccattttgtttggtgtcagatctggtgcagctgctttgactttg
attgtcgtttggatcacatcgagaagcagaaaaacgccgattttcattatcaaccaagtttcattgtttttaatcattttgcattctgcactctattttaaatatttt
actgtctaattactcttcagtgacttacgctctcaccggatttcctcagttcatcagtagaggtgacgttcatgtttatggtgctacaaatataattcaagtcctt
tcttgtggcttctattgagacttcactggtgtttcagataaaagttattttcacaggcgacaacttcaaaaggtga

pKC02 Sequence

Below is the sequence of the expressed region of pKC02. The region is bold corresponds to the Histidine tag. The lowercase, underlined, and italicized region corresponds to the Ste2p TM1-TM3S104C region (G31-R161, S104C).

ATG**CACCATCACCATCACCAT**GGGA*atggatctaccatcactttcgatgagttgcaaggtttagttaacagtactgttactcaggccattttgtttggtgtcagatctggtgc*
agctgctttgactttgattgtcgtttggatcacatcgagaagcagaaaaacgccgattttcattatcaaccaagtttcattgtttttaatcattttgcattctgcactctattttaaatatttactgtgtaatta
ctcttcagtgacttacgctctcaccggatttcctcagttcatcagtagaggtgacgttcatgtttatggtgctacaaatataattcaagtccttcttgtggcttctattgagacttcactggtgtttcagata
aaagttattttcacaggcgacaacttcaaaaggtga

pKC03 Sequence

Below is the sequence of the expressed region of pKC03. The region is bold corresponds to the Histidine tag. The lowercase, underlined, and italicized region corresponds to the Ste2p TM1 region (G31-T78).

ATG**CATCACCATCACCATCACCATCACCAT***aaagcaattttcgtactgaaaggttcactggacagagatctcgacagccgtattgaactggaactgcgtaccgat*
cataaagagctgtctgaacatctgctgctggttgatctcgcccgtaatgatctggcacgcattgctacccccggcagccgctacgtcgccgatctcaccaaagttgaccgttattcctatgtgctgc
acctcgtctctcgcgtagtcggcgaactgcgtcacgatcttgacgccctgcacgcttatcgcgccgctctgaatctggggacgttaagcggtgcgccgaaagtacgcgctaagctttggtgatga

pSW127 Sequence

Below is the sequence of the expressed region of pSW127. The region is bold corresponds to the TrpΔLE. The lowercase, underlined, and italicized region corresponds to the Ste2p TM127 region (G31-T114,T274-L340).

ATGCATCACCATCACCATCACCATCACCATAAAGCAATTTTCGTACTGAAAGGTTCACTGGACAGAGATCTCGAC
AGCCGTATTGAACTGGAACTGCGTACCGATCATAAAGAGCTGTCTGAACATCTGCTGCTGGTTGATCTCGCCCG
TAATGATCTGGCACGCATTGCTACCCCCGGCAGCCGCTACGTCGCCGATCTCACCAAAGTTGACCGTTATTCCT
ATGTGCTGCACCTCGTCTCTCGCGTAGTCGGCGAACTGCGTCACGATCTTGACGCCCTGCACGCTTATCGCGCC
GCTCTGAATCTGGGGACGTTAAGCGGTGCGCCGAAAGTACGCGCTAAGCTTTGG*atggggaatggatctaccatcactttcgatgagt*
tgcaaggtttagttaacagtactgttactcaggccattctgtttggtgtcagatctggtgcagctgctttgactttgattgtcgtgtggatcacatcgagaagcagaaaaacgccgattttcattatca
accaagtttcattgtttttaatcattttgcattctgcactctattttaaatatttactgtctaattactcttcagtgacttacgctctcaccacagatgtcttgactactgttgcaacattacttgctgtattgtctt
taccattatcatcactgtgggccacggctgctaataatgcatccaaaacaaacacaattacttcagactttacaacatccacagataggttttatccaggcacgctgtctagctttcaaactgata
gtatcaacaacgatgctaaaagcagtctctga

APPENDIX B

Chemical Shift Assignments for Ste2p TM1-TM3 (G31-R161) in 50% TFE:water at 45°C.

Residue #	Amino Acid	δ (HN)	δ (N)	δ (CO)	δ (CA)	δ (HA)	δ (CB)	δ (HB)	Other
31	G			179.38					
32	N	8.56	118.98	175.13	52.74	4.81	38.27	2.86	
33	G	8.44	109.49	174.06	45.15	3.98			
34	S	8.06	115.84	174.64	58.46	4.49	63.22	3.93	
35	T	7.96	116.29	174.64	62.72	4.28	69.00	4.28	δ CG 20.74, δ HG 1.22,
36	I	7.77	122.17	176.31	62.34	4.15	37.84	1.88	δ CG 27.46, 16.19, δ HG 1.56, 1.23, 0.87, δ CD 11.64, δ HD 0.86
37	T	7.62	115.99	175.25	63.90	4.08	69.04	4.23	δ CG 20.39, δ HG 1.20
38	F	8.10	121.43	176.78	60.51	4.30	38.23	3.16	
39	D	8.28	117.80	177.43	55.28	4.40	36.56	3.02, 2.86	
40	E	8.07	120.58	177.76	57.25	4.09	27.46	2.09	δ CG 32.38, δ HG 2.42, 2.58
41	L	8.11	121.50	178.09	57.40	4.08	41.23	1.70	δ CG 26.19, δ HG 1.68, δ CD 22.61, δ HD 0.87
42	Q	8.15	117.31	178.29	58.49	3.83	27.78	1.97	δ CG 33.35, δ HG 2.13, 2.21
43	G	7.89	106.76	175.98	46.36	3.88			
44	L	8.07	125.55	179.88	57.66	4.17	41.29	1.86, 1.37	δ CG 26.05 δ HG 1.43, δ CD 22.48, 22.90, δ HD 0.85, 0.87
45	V	8.54	122.80	177.56	66.55	3.59	30.86	2.12	δ CG 19.79, 21.14, δ HG 0.92, 1.00
46	N	8.45	117.90	178.29	56.12	4.44	38.05	2.93, 2.78	
47	S	8.36	116.73	175.41	61.52	4.03	62.45	3.59	
48	T	8.00	119.62	175.89	66.37	3.98	68.44	4.37	δ CG 19.81, δ HG 1.25
49	V	8.39	122.51	177.23	66.25	3.69	31.17	2.12	δ CG 19.90, 21.46, δ HG 0.96, 1.03
50	T	7.91	114.44	175.69	66.93	3.84	68.24	4.28	δ CG 20.10 δ HG 1.21
51	Q	7.81	119.00	177.37	59.03	3.96	28.01	2.19	δ CG 33.77, δ HG 2.44
52	A	7.96	122.85	180.44	54.90	4.16	16.90	1.57	
53	I	8.32	120.70	178.16	64.52	3.78	37.32	2.03	δ CG 28.28, 15.69 δ HG 1.16, 1.76, 0.91 δ CD 11.66, δ HD 0.81
54	L	8.29	120.07	179.13	57.50	4.08	40.93	1.85	δ CG 26.15, δ HG 1.44, δ CD 21.56, 23.89, δ HD 0.83

55	F	8.47	119.21	178.60	60.26	4.31	38.22	3.24	
56	G	8.35	110.24	175.74	46.58	3.74			
57	V	8.43	122.71	178.10	65.45	3.80	31.32	2.21	δ CG 19.94, 20.74, δ HG 0.94, 1.06
58	R	8.30	119.75	177.78	58.17	4.07	29.51	1.89	δ CG 26.47, δ HG 1.70, δ CD 42.59, δ HD 3.13
59	S	8.07	113.77	175.47	59.97	4.18	62.90	3.61, 3.82	
60	G	7.81	110.47	174.45	45.62	4.02			
61	A	8.08	123.77	179.43	54.06	4.14	17.13	1.43	
62	A	7.98	121.82	178.48	54.37	4.08	16.99	1.44	
63	A	7.69	120.31	178.59	54.36	4.06	17.13	1.46	
64	L	7.80	117.06	178.01	57.10	4.10	41.28	1.79	δ CG 26.12, δ HG 1.71, δ CD 22.57, 24.32, δ HD 0.88, 0.92
65	T	7.73	113.30	174.99	66.43	3.80	68.14	4.29	δ CG 19.96, δ HG 1.25
66	L	7.60	120.89	178.45	57.78	3.98	40.92	1.80	δ CG 26.09, δ HG 1.55, δ CD 21.98, 23.03, δ HD 0.80, 0.85
67	I	7.82	119.70	177.52	64.90	3.84	37.34	2.10	δ CG 28.32, 16.26, δ HG 1.08, 1.00 δ CD 11.34 δ HD 0.82
68	V	8.19	120.37	180.00	66.80	3.69	30.91	2.28	δ CG 20.02, 21.57, δ HG 0.94, 1.07
69	V	8.58	124.21	178.26	66.84	3.55	30.80	2.30	δ CG 19.81, 21.67, δ HG 0.98, 1.07
70	W	8.74	124.55	178.65	60.99	4.27	28.26	3.44, 3.59	
71	I	9.13	119.26	179.26	64.95	3.57	37.57	1.88	δ CG 28.14, 16.01, δ HG 1.08, 0.94 δ CD 12.07 δ HD 0.89
72	T	8.48	114.65	176.71	65.66	3.94	68.53	4.28	δ CG 20.25 δ HG 1.25
73	S	8.28	117.76	175.58	60.87	4.56	62.68	3.93	
74	R	7.69	121.10	177.23	56.64	4.06	29.30	1.73	δ CG 26.07, δ HG 1.36, δ CD 42.15, δ HD 2.64
75	S	7.82	114.57	174.44	59.28	4.30	63.08	3.89	
76	R	7.68	120.81	175.65	56.16	4.22	29.04	1.80	δ CG 26.53, δ HG 1.59, δ CD 42.69 δ HD 3.13
77	K	7.82	119.66	175.49	55.40	4.36	32.22	1.73	δ CG 24.09, δ HG 1.39, δ CD 28.22, δ CE 41.65, δ HE 2.94
78	T	7.62	114.88	172.88	60.57	4.57	68.90	4.21	
79	P			176.80	63.75		30.68	1.81	δ CG 26.74, δ HG 2.00, δ CD 50.02, δ HD 3.82, 3.67
80	I	7.43	117.79	176.22	62.52	3.86	37.20	1.81	δ CG 27.49, 16.19, δ HG 1.22, 0.85, δ CD 11.44, δ HD 0.87

81	F	7.50	120.00	176.87	59.58	4.38	38.35	3.19	
82	I	7.44	119.79	176.82	63.18	3.83	37.25	1.90	δ CG 27.86, 16.35, δ HG 1.26 δ CD 11.24, δ HD 0.86, 0.86
83	I	7.83	120.15	178.18	63.68	3.75	37.15	1.91	δ CG 27.86, 15.84, δ HG 1.64 δ CD 11.56, δ HD 0.81
84	N	8.15	120.00	176.09	55.61	4.43	37.98	2.78	
85	Q	7.90	119.36	177.60	58.57	4.02	27.95	2.12	δ CG 33.89, δ HG 2.53
86	V	8.37	120.05	177.25	65.94	3.75	30.96	2.15	δ CG 19.86, 21.26, δ HG 0.91, 1.02
87	S	8.06	114.81	175.19	61.99	4.18	62.58		
88	L	7.68	121.80	177.90	57.35	4.07	40.87	1.74	δ CG 26.11, δ HG 1.43, δ CD 22.46, 23.01 δ HD 0.74, 0.85
89	F	8.00	119.33	176.80	60.47	4.25	38.04	3.30	
90	L	8.34	117.92	180.19	57.71	3.98	40.91	1.38	δ CG 26.22, δ HG 2.06, δ CD 21.44, 24.01, δ HD 0.91
91	I	8.18	123.14	179.12	64.94	3.75	37.38	2.15	δ CG 28.05, 15.77, δ HG 1.15, 0.90 δ CD 11.58, δ HD 0.83
92	I	8.48	125.11	178.22	65.18	3.63	36.90	2.00	δ CG 27.91, 15.20, δ HG 1.08, 0.87, δ CD 11.48, δ HD 0.84
93	L	8.66	121.37	178.61	57.62	3.96	40.97	1.37	δ CG 25.97, δ HG 1.46, δ CD 22.31, 22.92, δ HD 0.76, 0.69
94	H	8.25	114.46	176.39	59.30	4.16	27.18	3.33	
95	S	8.24	116.44	174.73	61.54	4.17	62.43	3.93	
96	A	8.41	125.77	179.71	54.84	4.21	17.01	1.53	
97	L	8.09	118.30	178.68	57.15	4.08	40.97	1.49	δ CG 26.06, δ HG 1.54, δ CD 22.60, 23.70, δ HD 0.86, 0.83
98	Y	8.01	120.29	177.07	60.54	4.30	37.15	3.25	
99	F	8.27	119.06	177.04	60.78	4.26	37.81	3.21	
100	K	8.14	118.74	178.95	59.26	3.90	31.08	1.98	δ CG 24.69, δ HG 1.61,1.38, δ CD 28.65, δ CE 41.59, δ HE 2.90
101	Y	8.22	122.40	177.81	60.43	4.14	37.16	3.10, 3.21	
102	L	8.38	122.87	179.12	57.86	3.79	41.05	1.71	δ CG 25.85, δ HG 1.46, δ CD 22.26, 23.77, δ HD 0.80, 0.74
103	L	8.37	119.14	179.72	56.96	3.98	40.88	1.54	δ CG 26.23, δ HG 1.58, δ CD 22.22, 23.93, δ HD 0.76, 0.83
104	S	8.06	115.42	175.44	60.69	4.19	62.54	3.96	

105	N	7.77	120.92	175.71	55.17	4.50	38.55	2.47, 2.51	
106	Y	8.32	120.70	177.40	59.82	4.29	37.44	3.08	
107	S	8.23	115.72	175.08	60.45	4.27	62.58	3.96	
108	S	7.79	117.67	175.48	60.53	4.30	62.47	4.01	
109	V	7.85	121.92	176.75	65.16	3.83	31.22	2.10	δ CG 19.93, 20.55, δ HG 0.94, 0.98
110	T	7.71	113.43	176.44	64.89		68.19	4.06	δ CG 20.41, δ HG 1.28
111	Y	7.88	122.51	177.11	59.99	4.29	37.27	3.09	
112	A	7.87	122.71	178.92	54.12	4.04	17.31	1.48	
113	L	7.94	116.72	178.08	56.11	4.23	41.83	1.80, 1.55	δ CG 26.16, δ HG 1.81, δ CD 21.53, 23.88, δ HD 0.80, 0.79
114	T	7.65	109.20	175.13	62.58	4.24	69.61	4.23	δ CG 20.19, δ HG 1.21
115	G	7.88	109.65	173.65	44.89	3.93			
116	F	7.88	121.28	173.81	58.56	4.67	37.72	3.10	
117	P			177.54	64.30		30.47	1.72	δ CG 27.06, δ HG 1.90, δ CD 49.63, δ HD 3.72, 3.50
118	Q	8.05	117.91	176.48	58.01	4.06	27.46	2.33	δ CG 33.17, δ HG 2.29
119	F	7.77	119.21	176.73	58.75	4.43	38.34	3.14	
120	I	7.70	119.06	176.57	61.92	3.93	37.45	1.78	δ CG 16.53, 26.96, δ HG 1.37, 1.12, 0.82, δ CD 11.49, δ HD 0.78
121	S	7.74	116.16	174.76	58.86	4.33	63.08	3.95	
122	R	7.75	121.21	176.58	56.41	4.25	29.56	1.81	δ CG 26.81, δ HG 1.81, δ CD 42.69, δ HD 3.14
123	G	7.99	107.47	173.74	44.99	3.87			
124	D	7.94	118.42	174.72	52.67	4.68	37.19	2.83, 2.91	
125	V	7.64	119.56	175.52	62.41	3.96	31.67	2.05	δ CG 19.66, 19.75, δ HG 0.86, 0.83
126	H	8.30	120.58	173.80	54.99	4.66	28.08	3.18, 3.12	
127	V	7.70	119.81	175.38	62.31	4.01	32.10	1.98	δ CG 19.57, 20.02, δ HG 0.86, 0.84
128	Y	7.82	122.44	176.02	57.82	4.52	38.00	3.05, 2.93	
129	G	8.00	109.45	173.66	45.01	3.93			
130	A	7.83	123.44	177.67	52.83	4.30	18.02	1.41	
131	T	7.89	110.99	174.54	63.34	4.21	68.73	4.21	δ CG 20.54, δ HG 1.26
132	N	7.96	119.89	175.63	53.88	4.67	38.21	2.83	

133	I	7.72	120.34	176.06	62.54	3.94	37.38	1.91	δ CG 27.77, 16.21, δ HG 1.53, 0.91, δ CD 11.58, δ HD 0.86
134	I	7.66	120.35	176.84	63.66	3.75	36.66	1.92	δ CG 28.11, 16.01 δ HG 1.59, 0.92, δ CD 10.83, δ HD 0.83
135	Q	7.59	117.70	177.22	59.11	3.96	28.02	2.14	δ CG 33.84, δ HG 2.33
136	V	7.41	118.25	177.30	65.92	3.60	30.96	2.25	δ CG 19.79, 21.08, δ HG 0.93, 1.02
137	L	7.97	119.66	180.08	57.20	4.07	40.97	1.91	δ CG 25.81, δ HG 1.53, δ CD 21.32, 23.84, δ HD 0.80, 0.84
138	L	8.48	122.61	178.86	58.09	4.09	41.18	1.84	δ CG 26.24, δ HG 1.70, δ CD 22.37, 22.92, δ HD 0.84, 0.71
139	V	8.09	120.17	177.96	66.76	3.59	30.80	2.19	δ CG 19.85, 21.67, δ HG 0.93, 1.08
140	A	8.63	121.95	181.00	54.89	4.11	16.82	1.49	
141	S	8.33	116.94	176.08	62.12	4.20		3.95	
142	I	8.32	126.22	178.16	64.67	3.74	37.69	2.01	δ CG 28.24, 15.69, δ HG 1.30, 0.89 δ CD 11.72, δ HD 0.82
143	E	8.76	119.16	177.81	59.26	3.98	27.35	2.22, 2.11	δ CG 32.40, δ HG 2.37, 2.63
144	T	8.16	112.48	175.68	66.04	3.93	68.51	4.29	δ CG 20.17, δ HG 1.24
145	S	7.77	117.66	175.83	61.52	4.13	62.60	4.02	
146	L	8.13	124.30	178.18	57.74	4.20	41.19	1.89	δ CG 26.02, δ HG 1.66, δ CD 22.19, 23.26, δ HD 0.85, 0.87
147	V	7.99	118.37	177.52	66.19	3.63	30.86	2.18	δ CG 19.97, 21.46, δ HG 0.92, 1.06
148	F	8.06	119.53	176.76	60.92	4.19	38.33	3.24	
149	Q	7.97	117.49	177.63	58.87	3.89	28.35	2.33	δ CG 34.03, δ HG 2.64
150	I	8.22	119.44	177.35	64.40	3.74	37.22	1.99	δ CG 28.00, 15.40, δ HG 1.09, 0.66, δ CD 11.40, δ HD 0.81
151	K	8.15	119.48	179.24	59.21	3.99	31.46	1.97, 1.85	δ CG 24.51, δ HG 1.38, 1.57, δ CD 28.69, δ HD 1.63,δ CE 41.56, δ HE 2.90
152	V	8.05	120.96	177.95	65.88	3.59	30.82	2.06	δ CG 20.05, 21.08, δ HG 0.81, 1.02
153	I	7.98	121.40	178.12	64.14	3.70	37.14	1.87	δ CG 28.05, 15.84, δ HG 1.84, 0.66, δ CD 11.51, δ HD 0.82
154	F	8.50	118.73	177.46	59.69	4.42	38.23	3.17	
155	T	7.95	112.79	175.65	62.90	4.16	69.05	4.32	δ CG 20.84, δ HG 1.25
156	G	8.02	110.07	174.42	45.64	3.94			
157	D	8.12	118.47	175.00	53.16	4.62	37.03	2.86	
158	N	7.98	117.96	174.52	53.62	4.56	38.14	2.61, 2.58	

159	F	7.75	119.15	174.92	57.48	4.56	38.44	3.04, 3.16	
160	K	7.68	121.35	175.33	55.64	4.29	32.24	1.85	δ CG 23.87, δ HG, 1.39, δ CD 28.42, δ HD 1.73, δ CE 41.70, δ HE 2.98
161	R	7.83	121.68	177.13	54.40	4.37	30.04	1.93, 1.77	δ CG 26.41, δ HG 1.64, δ CD 42.69, δ HD 3.17

APPENDIX C

Chemical Shift Assignments for Ste2p TM1 (G31-T78) in 4:1 LPPG:DPC micelles at 45°C.

Residue #	Amino Acid	δ (HN)	δ (N)	δ (CO)	δ (CA)	δ (HA)	δ (CB)	δ (HB)	Other
31	G	8.39	110.96	171.43	43.22	3.89	27.27		
32	N	8.35	119.29	173.45	51.07	4.69	36.74	2.76, 2.79	
33	G	8.37	110.03	171.71	43.22	3.89			
34	S	8.05	116.16	172.08	56.17	4.47	61.81	3.80	
35	T	8.09	116.81	171.73	59.85	4.31	67.56	4.11	δ CG 19.70, δ HG 1.09
36	I	7.89	123.14	172.37	58.65	4.27	36.96	1.72	δ CG 15.13, 25.53, δ HG 0.76, 1.47,1.09, δ HD 0.75, δ CD 11.22
37	T	7.79	115.99	173.04	57.79	4.54	68.95		δ CG 19.63, δ HG 1.17
38	F	8.91	121.97	175.04	58.95	4.26	36.48	3.13	
39	D	8.39	118.53	176.29	54.89	4.29	38.26	2.58	
40	E	7.74	119.97	175.97	52.84	3.98	31.27	2.01	δ CG 28.91, δ HG 2.48, 2.44
41	L	7.77	120.45	175.58	54.98	4.04	39.58	1.56, 1.69	δ CG 24.77, δ CD 21.83, 22.91, δ HD 0.80, 0.83,
42	Q	8.06	117.90	175.59	56.88	3.78	26.15	1.96, 2.13	δ CG 31.91, δ HG 2.25, 2.17,
43	G	7.81	106.50	173.08	44.14	3.83			
44	L	7.63	121.62	175.86	54.54	4.21	40.42	1.85, 1.59	δ CG 24.84, δ HG 1.77, δ CD 21.64, 22.94, δ HD 0.83, 0.87,
45	V	7.70	117.87	173.12	61.95	3.78	29.58	2.12	δ CG 19.46, δ HG 0.88
46	N	7.76	119.63	173.93	51.41	4.69	36.94	2.86, 2.95	δ CG 19.83, δ HG 1.20
47	S	8.17	115.87	173.49	58.19	4.38	61.38	4.05	
48	T	8.17	119.38	174.17	63.95	4.00	66.43	4.20	
49	V	8.09	121.98	174.92	64.28	3.65	29.53	2.18	δ CG 20.71, 19.23, δ HG 0.92, 1.05
50	T	7.84	115.29	174.08	64.55	3.63	65.89	4.23	δ CG 20.30, δ HG 1.19
51	Q	7.96	119.51	176.13	56.76	3.97	26.27	1.95, 2.13	δ CG 31.95, δ HG 2.45, 2.24
52	A	7.94	123.38	177.09	53.22	4.18	16.17	1.54	
53	I	8.11	118.57	175.54	63.20	3.67	35.47	2.02	δ CG 15.13, 27.10, δ HG 0.86, 2.06, 1.88 δ CD 11.31, δ HD 0.81

54	L	8.19	120.09	176.79	56.14	3.99	39.06	1.83, 1.53	δ CG 24.84, δ HG 1.77, δ CD 20.87, 22.78, δ HD 0.83
55	F	8.21	120.33	176.60	59.01	4.26	36.66	3.27	
56	G	8.60	110.14	172.48	45.38	3.70			
57	V	8.59	121.27	171.67	64.42	3.64	29.48	2.17	δ CG 20.67, 19.37, δ HG 1.05, 0.92
58	R	8.59	121.16	175.18	53.19	3.99	27.40		δ CG 25.64, δ HG 1.61, 1.79, δ CD 41.41, δ HD 3.13
59	S	8.09	119.51	177.00	56.68	4.31	61.78	3.83	
60	G	8.04	116.24	173.71	43.23	3.92			
61	A	8.39	123.97	177.47	53.33	3.99	15.88	1.46	
62	A	7.82	121.73	177.74	52.68	4.12	15.86	1.50	
63	A	7.95	121.38	176.46	52.67	4.05	16.20	1.46	
64	L	8.15	117.63	176.18	55.92	3.99	39.42	1.77, 1.68	δ CG 24.91, δ HG 1.76, δ CD 22.17, δ HD 0.84
65	T	7.87	114.77	173.32	65.05	3.78	66.47	4.25	δ CG 19.34, δ HG 1.21
66	L	7.45	121.38	176.94	56.16	3.99	39.57	1.56, 1.87	δ CG 24.61, δ HG 1.69, δ CD 21.64, 22.94, δ HD 0.83, 0.87
67	I	7.89	119.50	175.22	62.99	3.80	35.79	2.08	δ CG 15.23, 27.38, δ HG 0.92, 2.00, 2.17 δ CD 11.46, δ HD 0.80
68	V	8.23	120.68	177.01	65.30	3.60	29.43	2.17	δ CG 21.14, 19.43, δ HG 1.06, 0.94
69	V	8.42	123.02	176.11	64.61	3.62	29.27	2.17	δ CG 20.72, 18.96, δ HG 1.07, 0.94
70	W	8.34	123.80	176.82	59.83	4.21	26.54	3.37	
71	I	8.74	118.70	176.19	62.91	3.44	36.11	2.03	δ CG 15.46, 27.03 δ HG 0.99, 1.26, 1.10, δ CD 11.81, δ HD 0.87
72	T	8.29	112.34	174.47	62.96	4.01	66.98	4.30	δ CG 19.64, δ HG 1.27
73	S	7.93	117.76	173.17	58.25	4.30	61.16	3.93	
74	R	7.42	121.39	174.42	54.50	4.06	27.79	1.66, 1.52	δ CG 24.31, δ HG 1.33, 1.24, δ CD 40.33, δ HD 2.61, 2.51
75	S	7.73	114.82	172.00	56.72	4.31	61.64	3.84	

76	R	7.78	122.68	173.68	53.80	4.32	28.42	1.84, 1.73	δ CG 24.94, δ HG 1.63, 1.57, δ CD 41.20, δ HD 3.12
77	K	8.13	123.61	173.71	54.13	4.33	30.68	1.84, 1.72	δ CG 22.39, δ HG 1.42, 1.39, δ CD 26.70, δ HD 1.63 δ CE 39.96, δ HE 2.95
78	T	7.63	121.14		60.81	4.09	68.44	4.17	δ CG 19.38, δ HG 1.13

APPENDIX D: Accuracy of H-D Exchange and T2 Relaxation Experiments

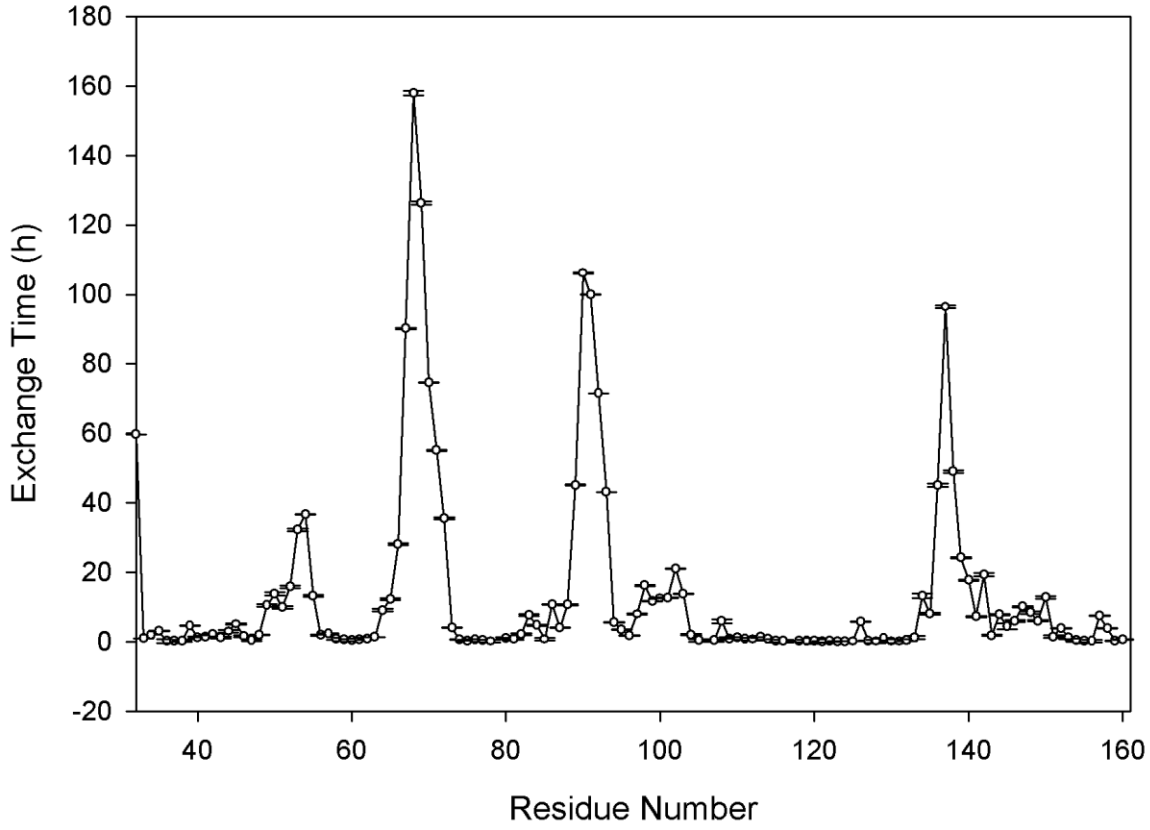

Figure 1: H-D exchange results expressed with standard deviation. The raw H-D exchange rates were plotted as a function of residue number. The standard deviation of the fit calculation is expressed as an error bar.

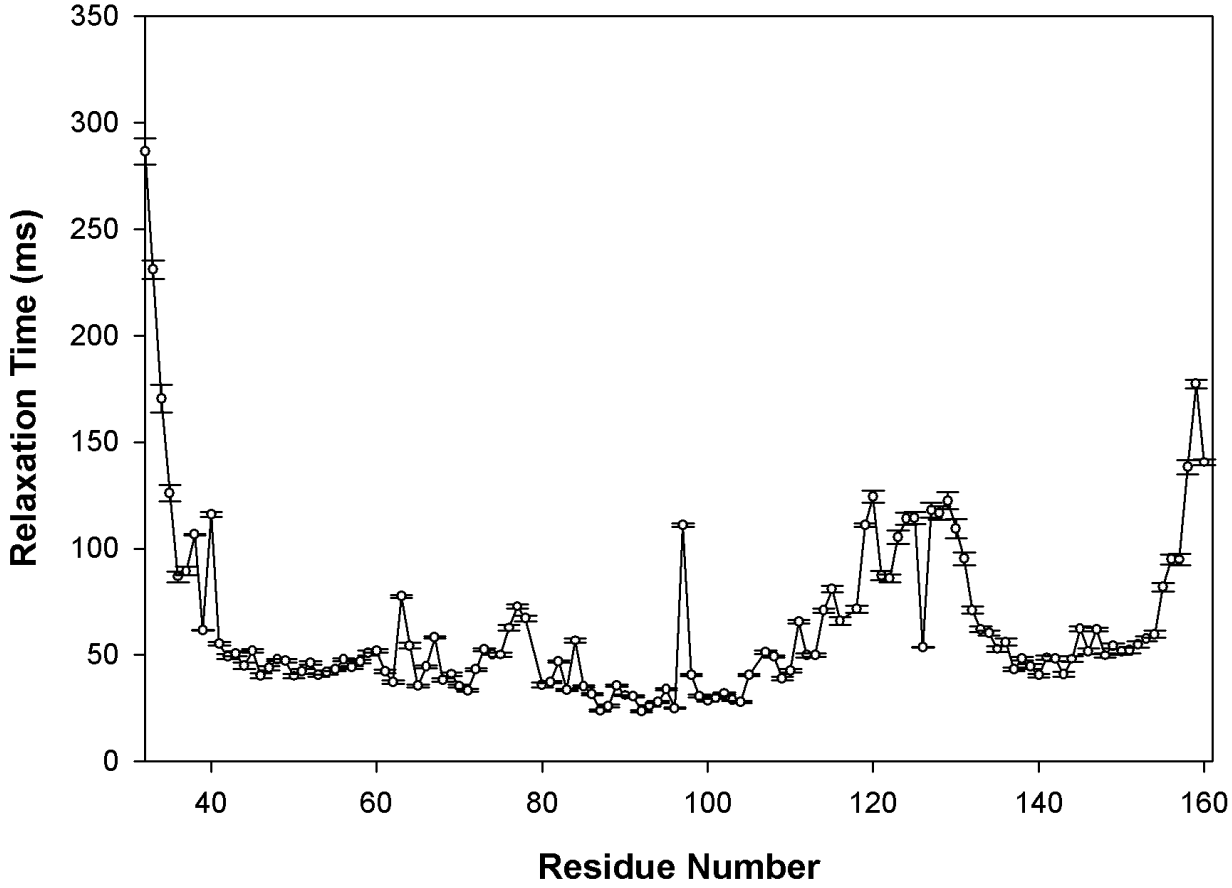

Figure 2: H-D exchange results expressed with standard deviation. The raw H-D exchange rates were plotted as a function of residue number. The standard deviation of the fit calculation is expressed as an error bar.

REFERENCES

1. 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.
2. Mierke, D. F., and Giragossian, C. (2001) Peptide hormone binding to G-protein-coupled receptors: structural characterization via NMR techniques, *Medicinal research reviews 21*, 450-471.
3. Foord, S. M., Bonner, T. I., Neubig, R. R., Rosser, E. M., Pin, J. P., Davenport, A. P., Spedding, M., and Harmar, A. J. (2005) International Union of Pharmacology. XLVI. G protein-coupled receptor list, *Pharmacol Rev 57*, 279-288.
4. Bridges, T. M., and Lindsley, C. W. (2008) G-protein-coupled receptors: from classical modes of modulation to allosteric mechanisms, *ACS chemical biology 3*, 530-541.
5. Hamm, H. E. (1998) The many faces of G protein signaling, *J Biol Chem 273*, 669-672.
6. Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many drug targets are there?, *Nat Rev Drug Discov 5*, 993-996.
7. McCusker, E. C., Bane, S. E., O'Malley, M. A., and Robinson, A. S. (2007) Heterologous GPCR expression: a bottleneck to obtaining crystal structures, *Biotechnol Prog 23*, 540-547.
8. Arinaminpathy, Y., Khurana, E., Engelman, D. M., and Gerstein, M. B. (2009) Computational analysis of membrane proteins: the largest class of drug targets, *Drug Discov Today 14*, 1130-1135.
9. 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.
10. 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 beta2-adrenergic G protein-coupled receptor, *Science 318*, 1258-1265.
11. 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 beta2 adrenergic G-protein-coupled receptor, *Nature 450*, 383-387.
12. 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*, 1211-1217.
13. 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.

14. 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.
15. 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.
16. Chien, E. Y., Liu, W., Zhao, Q., Katritch, V., Han, G. W., Hanson, M. A., Shi, L., Newman, A. H., Javitch, J. A., Cherezov, V., and Stevens, R. C. (2010) Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist, *Science 330*, 1091-1095.
17. Wu, B., Chien, E. Y., Mol, C. D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P., Bi, F. C., Hamel, D. J., Kuhn, P., Handel, T. M., Cherezov, V., and Stevens, R. C. (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists, *Science 330*, 1066-1071.
18. Dore, A. S., Robertson, N., Errey, J. C., Ng, I., Hollenstein, K., Tehan, B., Hurrell, E., Bennett, K., Congreve, M., Magnani, F., Tate, C. G., Weir, M., and Marshall, F. H. (2011) Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine, *Structure 19*, 1283-1293.
19. Lebon, G., Warne, T., Edwards, P. C., Bennett, K., Langmead, C. J., Leslie, A. G., and Tate, C. G. (2011) Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation, *Nature 474*, 521-525.
20. Rasmussen, S. G., Choi, H. J., Fung, J. J., Pardon, E., Casarosa, P., Chae, P. S., Devree, B. T., Rosenbaum, D. M., Thian, F. S., Kobilka, T. S., Schnapp, A., Konetzki, I., Sunahara, R. K., Gellman, S. H., Pautsch, A., Steyaert, J., Weis, W. I., and Kobilka, B. K. (2011) Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor, *Nature 469*, 175-180.
21. Rasmussen, S. G., DeVree, B. T., Zou, Y., Kruse, A. C., Chung, K. Y., Kobilka, T. S., Thian, F. S., Chae, P. S., Pardon, E., Calinski, D., Mathiesen, J. M., Shah, S. T., Lyons, J. A., Caffrey, M., Gellman, S. H., Steyaert, J., Skiniotis, G., Weis, W. I., Sunahara, R. K., and Kobilka, B. K. (2011) Crystal structure of the beta2 adrenergic receptor-Gs protein complex, *Nature 477*, 549-555.
22. Shimamura, T., Shiroishi, M., Weyand, S., Tsujimoto, H., Winter, G., Katritch, V., Abagyan, R., Cherezov, V., Liu, W., Han, G. W., Kobayashi, T., Stevens, R. C., and Iwata, S. (2011) Structure of the human histamine H1 receptor complex with doxepin, *Nature 475*, 65-70.
23. Granier, S., Manglik, A., Kruse, A. C., Kobilka, T. S., Thian, F. S., Weis, W. I., and Kobilka, B. K. (2012) Structure of the delta-opioid receptor bound to naltrindole, *Nature 485*, 400-404.
24. Haga, K., Kruse, A. C., Asada, H., Yurugi-Kobayashi, T., Shiroishi, M., Zhang, C., Weis, W. I., Okada, T., Kobilka, B. K., Haga, T., and Kobayashi, T. (2012) Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist, *Nature 482*, 547-551.
25. Hanson, M. A., Roth, C. B., Jo, E., Griffith, M. T., Scott, F. L., Reinhart, G., Desale, H., Clemons, B., Cahalan, S. M., Schuerer, S. C., Sanna, M. G., Han, G. W., Kuhn, P., Rosen, H., and Stevens, R. C. (2012) Crystal structure of a lipid G protein-coupled receptor, *Science 335*, 851-855.

26. Kruse, A. C., Hu, J., Pan, A. C., Arlow, D. H., Rosenbaum, D. M., Rosemond, E., Green, H. F., Liu, T., Chae, P. S., Dror, R. O., Shaw, D. E., Weis, W. I., Wess, J., and Kobilka, B. K. (2012) Structure and dynamics of the M3 muscarinic acetylcholine receptor, *Nature 482*, 552-556.
27. Manglik, A., Kruse, A. C., Kobilka, T. S., Thian, F. S., Mathiesen, J. M., Sunahara, R. K., Pardo, L., Weis, W. I., Kobilka, B. K., and Granier, S. (2012) Crystal structure of the micro-opioid receptor bound to a morphinan antagonist, *Nature 485*, 321-326.
28. Thompson, A. A., Liu, W., Chun, E., Katritch, V., Wu, H., Vardy, E., Huang, X. P., Trapella, C., Guerrini, R., Calo, G., Roth, B. L., Cherezov, V., and Stevens, R. C. (2012) Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic, *Nature 485*, 395-399.
29. Wu, H., Wacker, D., Mileni, M., Katritch, V., Han, G. W., Vardy, E., Liu, W., Thompson, A. A., Huang, X. P., Carroll, F. I., Mascarella, S. W., Westkaemper, R. B., Mosier, P. D., Roth, B. L., Cherezov, V., and Stevens, R. C. (2012) Structure of the human kappa-opioid receptor in complex with JDTic, *Nature 485*, 327-332.
30. Audet, M., and Bouvier, M. (2012) Restructuring G-protein- coupled receptor activation, *Cell 151*, 14-23.
31. 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, *J Am Chem Soc 127*, 8010-8011.
32. 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 128*, 5300.
33. Etzkorn, M., Martell, S., Andronesi, O. C., Seidel, K., Engelhard, M., and Baldus, M. (2007) Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy, *Angew Chem Int Ed Engl 46*, 459-462.
34. 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, *Angew Chem Int Ed Engl 47*, 7297-7300.
35. Reckel, S., Gottstein, D., Stehle, J., Lohr, F., Verhoefen, M. K., Takeda, M., Silvers, R., Kainosho, M., Glaubitz, C., Wachtveitl, J., Bernhard, F., Schwalbe, H., Guntert, P., and Dotsch, V. (2011) Solution NMR structure of proteorhodopsin, *Angew Chem Int Ed Engl 50*, 11942-11946.
36. Stehle, J., Scholz, F., Lohr, F., Reckel, S., Roos, C., Blum, M., Braun, M., Glaubitz, C., Dotsch, V., Wachtveitl, J., and Schwalbe, H. (2012) Characterization of the ground state dynamics of proteorhodopsin by NMR and optical spectroscopies, *J Biomol NMR 54*, 401-413.
37. 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.
38. Park, S. H., Das, B. B., Casagrande, F., Tian, Y., Nothnagel, H. J., Chu, M., Kiefer, H., Maier, K., De Angelis, A. A., Marassi, F. M., and Opella, S. J. (2012) Structure of the chemokine receptor CXCR1 in phospholipid bilayers, *Nature*.

39. Kontaxis, G., Delaglio, F., and Bax, A. (2005) Molecular fragment replacement approach to protein structure determination by chemical shift and dipolar homology database mining, *Methods Enzymol 394*, 42-78.
40. Ridge, K. D., Lu, Z., Liu, X., and Khorana, H. G. (1995) Structure and function in rhodopsin. Separation and characterization of the correctly folded and misfolded opsins produced on expression of an opsin mutant gene containing only the native intradiscal cysteine codons, *Biochemistry 34*, 3261-3267.
41. 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.
42. 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.
43. Maggio, R., Vogel, Z., and Wess, J. (1993) Reconstitution of functional muscarinic receptors by co-expression of amino- and carboxyl-terminal receptor fragments, *FEBS Lett 319*, 195-200.
44. Scarselli, M., Armogida, M., Chiacchio, S., DeMontis, M. G., Colzi, A., Corsini, G. U., and Maggio, R. (2000) Reconstitution of functional dopamine D(2s) receptor by co-expression of amino- and carboxyl-terminal receptor fragments, *Eur J Pharmacol 397*, 291-296.
45. Sigrist, H., Wenger, R. H., Kislig, E., and Wuthrich, M. (1988) Refolding of bacteriorhodopsin. Protease V8 fragmentation and chromophore reconstitution from proteolytic V8 fragments, *Eur J Biochem 177*, 125-133.
46. Ozawa, S., Hayashi, R., Masuda, A., Iio, T., and Takahashi, S. (1997) Reconstitution of bacteriorhodopsin from a mixture of a proteinase V8 fragment and two synthetic peptides, *Biochim Biophys Acta 1323*, 145-153.
47. Marti, T. (1998) Refolding of bacteriorhodopsin from expressed polypeptide fragments, *J Biol Chem 273*, 9312-9322.
48. Tiburu, E. K., Bowman, A. L., Struppe, J. O., Janero, D. R., Avraham, H. K., and Makriyannis, A. (2009) Solid-state NMR and molecular dynamics characterization of cannabinoid receptor-1 (CB1) helix 7 conformational plasticity in model membranes, *Biochim Biophys Acta 1788*, 1159-1167.
49. Tiburu, E. K., Tyukhtenko, S., Deshmukh, L., Vinogradova, O., Janero, D. R., and Makriyannis, A. (2009) Structural biology of human cannabinoid receptor-2 helix 6 in membrane-mimetic environments, *Biochem Biophys Res Commun 384*, 243-248.
50. 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.
51. Ruan, K. H., So, S. P., Wu, J., Li, D., Huang, A., and Kung, J. (2001) Solution structure of the second extracellular loop of human thromboxane A2 receptor, *Biochemistry 40*, 275-280.
52. Wu, J., Feng, M., and Ruan, K. H. (2008) Assembling NMR structures for the intracellular loops of the human thromboxane A2 receptor: implication of the G protein-coupling pocket, *Arch Biochem Biophys 470*, 73-82.

53. Geng, L., Wu, J., So, S. P., Huang, G., and Ruan, K. H. (2004) Structural and functional characterization of the first intracellular loop of human thromboxane A2 receptor, *Arch Biochem Biophys 423*, 253-265.
54. Zou, C., Naider, F., and Zerbe, O. (2008) Biosynthesis and NMR-studies of a double transmembrane domain from the Y4 receptor, a human GPCR, *J Biomol NMR 42*, 257-269.
55. Cohen, L. S., Arshava, B., Neumoin, A., Becker, J. M., Guntert, P., Zerbe, O., and Naider, F. (2011) Comparative NMR analysis of an 80-residue G protein-coupled receptor fragment in two membrane mimetic environments, *Biochim Biophys Acta 1808*, 2674-2684.
56. 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.
57. 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.
58. 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.
59. 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.
60. 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.
61. 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.
62. Cohen, L. S., Becker, J. M., and Naider, F. (2010) Biosynthesis of peptide fragments of eukaryotic GPCRs in Escherichia coli by directing expression into inclusion bodies, *J Pept Sci 16*, 213-218.
63. Elion, E. A. (2000) Pheromone response, mating and cell biology, *Current opinion in microbiology 3*, 573-581.
64. Dohlman, H. G. (2002) G proteins and pheromone signaling, *Annual review of physiology 64*, 129-152.
65. 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.
66. Caroccia, K. E., Estephan, R., Cohen, L. S., Arshava, B., Hauser, M., Zerbe, O., Becker, J. M., and Naider, F. (2011) Expression and biophysical analysis of a triple-transmembrane domain-containing fragment from a yeast G protein-coupled receptor, *Biopolymers*.

67. 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.
68. Chopra, A., Yeagle, P. L., Alderfer, J. L., and Albert, A. D. (2000) Solution structure of the sixth transmembrane helix of the G-protein-coupled receptor, rhodopsin, *Biochimica et Biophysica Acta 1463*, 1-5.
69. Katragadda, M., Chopra, A., Bennett, M., Alderfer, J. L., Yeagle, P. L., and Albert, A. D. (2001) Structures of the transmembrane helices of the G-protein coupled receptor, rhodopsin, *J Pept Res 58*, 79-89.
70. 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.
71. 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.
72. 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.
73. 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.
74. 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.
75. Kim, H., Lee, B. K., Naider, F., and Becker, J. M. (2009) Identification of specific transmembrane residues and ligand-induced interface changes involved in homo-dimer formation of a yeast G protein-coupled receptor, *Biochemistry 48*, 10976-10987.
76. 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.
77. Tantry, S., Ding, F. X., Dumont, M., Becker, J. M., and Naider, F. (2010) Binding of fluorinated phenylalanine alpha-factor analogues to Ste2p: evidence for a cation-pi binding interaction between a peptide ligand and its cognate G protein-coupled receptor, *Biochemistry 49*, 5007-5015.
78. 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.
79. 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.

80. 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.
81. 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.
82. Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins, *J Biomol NMR 20*, 71-75.
83. 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, *J Am Chem Soc 125*, 13868-13878.
84. Kerman, A., and Ananthanarayanan, V. S. (2005) Expression and spectroscopic characterization of a large fragment of the mu-opioid receptor, *Biochim Biophys Acta 1747*, 133-140.
85. 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.
86. 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.
87. Greenfield, N., and Fasman, G. D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation, *Biochemistry 8*, 4108-4116.
88. Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *J Am Chem Soc 114*, 10663-10665.
89. Kelly, A. E., Ou, H. D., Withers, R., and Dotsch, V. (2002) Low-conductivity buffers for high-sensitivity NMR measurements, *J Am Chem Soc 124*, 12013-12019.
90. Killian, J. A., Trouard, T. P., Greathouse, D. V., Chupin, V., and Lindblom, G. (1994) A general method for the preparation of mixed micelles of hydrophobic peptides and sodium dodecyl sulphate, *FEBS Lett 348*, 161-165.
91. Bodenhausen, G., and Ruben, D. J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy, *Chemical Physics Letters 69*, 185-189.
92. 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.
93. Grzesiek, S., and Bax, A. (1992) Improved 3D Triple-Resonance NMR Techniques Applied to a 31-kDa Protein., *J Magn Reson 96*, 432-440.
94. 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.
95. 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.

96. 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.
97. 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.
98. 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.
99. Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1992) A 4D HCCH-TOCSY experiment for assigning the side chain 1H and 13C resonances of proteins, *J Biomol NMR 2*, 655-659.
100. Montelione, G. T., Lyons , B. A., Emerson, S. D., and Tashiro, M. (1992) An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins, *Journal of the American Chemical Society 114*, 10974-10975.
101. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, *J Biomol NMR 44*, 213-223.
102. 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, *J Biomol NMR 4*, 171-180.
103. Noggle, R. C., and Schrimer, R. E. (1971) *The Nuclear Overhauser Effect: Chemical Applications*, Adacemic Press, New York, NY.
104. Thrippleton, M. J., and Keeler, J. (2003) Elimination of Zero-Quantum Interference in Two-Dimensional NMR Spectra *Angewandte Chemie International Edition 42*, 3938-3941.
105. Kumar, A., Ernst, R. R., and Wuthrich, K. (1980) A two-dimensional nuclear Overhauser enhancement (2DNOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules., *Biochem Biophys Res Commun 95*, 1-6.
106. Macura, S., and Ernst, R. R. (1980) Elucidation of cross relaxation in liquids by two-dimensional N.M.R. spectroscopy, *Molecular Physics 41*, 95-117.
107. 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.
108. Liang, B., Bushweller, J. H., and Tamm, L. K. (2006) Site-directed parallel spin-labeling and paramagnetic relaxation enhancement in structure determination of membrane proteins by solution NMR spectroscopy, *J Am Chem Soc 128*, 4389-4397.
109. 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.

110. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS, *J Biomol NMR 24*, 171-189.
111. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA, *Journal of molecular biology 319*, 209-227.
112. Guntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods in molecular biology (Clifton, N.J 278*, 353-378.
113. Fiorito, F., Herrmann, T., Damberger, F. F., and Wuthrich, K. (2008) Automated amino acid side-chain NMR assignment of proteins using (13)C- and (15)N-resolved 3D [(1)H, (1)H]-NOESY, *J Biomol NMR 42*, 23-33.
114. Volk, J., Herrmann, T., and Wuthrich, K. (2008) Automated sequence-specific protein NMR assignment using the memetic algorithm MATCH, *J Biomol NMR 41*, 127-138.
115. Hofmann, K., and Stoffel, W. (1993) TMbase-A database of membrane spanning protein segments, *Biol. Chem. 374*.
116. Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins, *Bioinformatics 14*, 378-379.
117. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, *Journal of molecular biology 305*, 567-580.
118. Juretic, D., Zoranic, L., and Zucic, D. (2002) Basic charge clusters and predictions of membrane protein topology, *J Chem Inf Comput Sci 42*, 620-632.
119. Gehret, A. U., Bajaj, A., Naider, F., and Dumont, M. E. (2006) Oligomerization of the yeast alpha-factor receptor: implications for dominant negative effects of mutant receptors, *J Biol Chem 281*, 20698-20714.
120. Therien, A. G., Grant, F. E., and Deber, C. M. (2001) Interhelical hydrogen bonds in the CFTR membrane domain, *Nat Struct Biol 8*, 597-601.
121. 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.
122. 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.
123. 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.
124. Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G., and Deber, C. M. (2009) Detergent binding explains anomalous SDS-PAGE migration of membrane proteins, *Proc Natl Acad Sci U S A 106*, 1760-1765.
125. Nadeau, V. G., Rath, A., and Deber, C. M. (2012) Sequence hydropathy dominates membrane protein response to detergent solubilization, *Biochemistry 51*, 6228-6237.

126. Tulumello, D. V., and Deber, C. M. (2012) Efficiency of detergents at maintaining membrane protein structures in their biologically relevant forms, *Biochim Biophys Acta 1818*, 1351-1358.