Biosynthesis and Biophysical Analysis of Domains of the Alpha-factor Receptor (Ste2p): A *Saccharomyces cerevisiae* G Protein-Coupled Receptor

By

RACHA ESTEPHAN

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

2005

UMI Number: 3187449

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

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. Lesley Davenport Executive Officer

Dr. Ruth E. Stark _____

Dr. Robert Bittman _____

Dr. Tatyana Polenova _____

Dr. Richard E. Davis _____

Supervisory Committee

The City University of New York

Abstract

BIOSYNTHESIS AND BIOPHYSICAL ANALYSIS OF DOMAINS OF THE ALPHA-FACTOR RECEPTOR (STE2P): A *SACCHAROMYCES CEREVISIAE* G PROTEIN-COUPLED RECEPTOR

By

Racha Estephan

Mentor: Professor Fred Naider

The yeast *Saccharomyces cerevisiae* α-factor pheromone receptor (Ste2p) belongs to the G protein-coupled family of receptors (GPCRs). To analyze the structure of Ste2p a 33-residue fragment comprising the sixth transmembrane domain (M6) spanning residues 238-270 of Ste2p and a 73-mer multi-domain fragment spanning residues 267-339 containing the third extracellular loop, the seventh transmembrane domain, and 40 residues of the cytosolic tail (E3-M7-24-T40) have been biosynthesized as fusion proteins. The M6 fusion protein (M6FP), a variant M6FP(P258L), and the multidomain M7 fusion protein (M7FP) were purified to near homogeneity as judged by HPLC, and their molecular weights verified by gel electrophoresis and mass spectrometry. The 33- and the 73-residue peptides were released from their fusion proteins by CNBr and isolated via HPLC. These peptides were obtained in ^{14}N, ^{15}N, and $^{13}C/^{15}N$ forms in approximately 10 mg quantities.

Circular dichroism analysis of M6, M6(P258L), and E3-M7-24-T40 was performed in trifluoroethanol/water mixtures, and in the presence of detergent micelles and lipid bilayers. The M6 and M6(P258L) peptides could not be studied at concentrations higher than 100 μM in the presence of detergent micelles and did not

integrate into bilayers. In contrast, the E3-M7-24-T40 peptide integrated into detergent micelles at concentrations (200-500 μM) suitable for NMR investigations and into bilayers.

HSQC experiments performed in organic-aqueous solvents and detergent micelles on the [^{15}N]-labeled E3-M7-24-T40 peptide showed a clear dispersion of the nitrogen-amide proton correlation crosspeaks indicative of a pure, uniformly labeled molecule that assumed a partially ordered structure. 2D and 3D NMR experiments were performed in trifluoroethanol/water (1:1) and chloroform/methanol/water (4:4:1) to obtain a high-resolution structure of the 73-residue multiple domain peptide. Chemical shift indices suggested that in both aqueous-organic media helical subdomains existed in both the transmembrane and cytoslic tail of the multi-domain peptide. NMR structural modeling of the [^{15}N]-labeled E3-M7-24-T40 peptide in both aqueous-organic media reveals a peptide with three helical parts interrupted by three proline residues. This cytosolic tail participates in down-regulation of Ste2p and the helical tendency in the cytosolic tail of the protein may play a role in protein-protein interactions leading to endocytosis and desensitization of this GPCR.

This thesis is dedicated to my dear parents Edouard and Hanan, my siblings Rima, Amir and Rita, my beautiful son Marvin, and my wonderful and loving husband Simon

Acknowledgments

I would like to thank my mentor Dr. Fred Naider for his intellectual support, encouragement, and enthusiasm. I want to thank him for giving me the opportunity to join his research laboratory and be a member of a wonderful research family. Working on his proposed thesis project I was able to learn numerous techniques and scientific information that allowed me to be a better thinker and scientist.

I also would like to thank my thesis committee members, Dr. Tatyana Polenova, Dr. Robert Bittman, and Dr. Richard Davis, for their time and their helpful comments on this research project. I want to especially thank Dr. Ruth Stark for her caring personality and for all her helpful advises not only in this research project but also in everything that came along the way.

I want to thank all members of Dr. Naider's laboratory, and especially Jacqueline Englander, for all her help on the biosynthesis and purification part of my thesis work and for her friendship, and Dr. Boris Arshava, for his enormous help on the processing and interpretation of NMR data. I also would like to thank Dr. Hsin Wang, the NMR manager of the College of Staten Island, for his help in setting up some NMR experiments.

I am thankful to the New York City Louis Strokes Alliance for Minority Participation (NYC LS-AMP) in Research from the National Science Foundation for the financial support during the first two years of my thesis. I am also thankful to the National Institutes of Health, grants GM22086-27S1/29S1, for financial support throughout my thesis work.

I am very grateful to my family for their constant support and love. My eternal gratitude goes to my parents, Edouard and Hanan, for encouraging me to be the best that I can be. A special thanks to my mother for being the best mother on earth, for always not only taking care of me but also of my husband and son. I am thankful to my sister, Rima, for being my friend, for deciding to get her Ph.D. degree and walk this long journey with me, and for looking after me. I want to also thank my brother, Amir, and my sister, Rita, for always cheering me up when I needed it.

My eternal gratitude goes to the man of my dreams, the love of my life, my soul mate, my husband Simon for his endless support and eternal love that helped me overcome this long journey. Finally, my biggest thanks go to my son Marvin who is the center and the reason of my life. He is the cutest and the most wonderful thing that ever happened to me and I love him dearly.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AngII / AT_2	angiotensin II
BPTI	bovine pancreatic trypsin inhibitor
BSA	bovine serum albumin
CD	circular dichroism
$CDCl_3$	chloroform
CD_3OH	methanol
CNBr	cyanogen bromide
CSI	chemical shift index
Da	dalton
DHPC	1,2-dihexanoyl-*sn*-glycero-3-phosphocholine
DMPC	1,2-dimyristoyl-*sn*-glycero-3-phosphocholine
DMPG	1,2-dimyristoyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)]
DPC	dodecylphosphocholine
E3-M7-24-T40	multidomain fragment containing the third extracellular loop, the seventh transmembrane domain, and 40 residues of the cytosolic tail of Ste2p
EDTA	ethylenediaminetetraacetic acid
ESI-MS	electrospray ionization mass spectrometry
GPCRs	G protein-coupled receptors
G protein	heterotrimeric guanine nucleotide-binding protein
GuHCl	guanidinium hydrochloride

HPLC	high-performance liquid chromatography
HSQC	heteronuclear single quantum correlation
Hz	hertz
IPTG	isopropyl-β-D-thiogalactopyranoside
$^{3}J_{HN\alpha}$ / $J_{NH\text{-}CH\alpha}$	NH-CHα coupling constants
LB	Luria Broth media
M6 / TM6	transmembrane domain six
M6(P258L)	transmembrane domain six with a Pro258Leu mutation
M7 / TM7	transmembrane domain seven
M6FP / M6FP(WT)	M6 wild-type fusion protein
M6FP(P258L)	M6 fusion protein with a Pro258Leu mutation
M7FP	M7 or E3-M7-24-T40 fusion protein
MHz	megahertz
m/*z*	mass-to-charge ratio
Nle	norleucine
NMR	nuclear magnetic resonance spectroscopy
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser enhanced spectroscopy
OD	optical density
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride

PPG	1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)]
ppm	parts per million
[θ]	mean residue ellipticity
Φ	phi dihedral angle
Raf-1	serine-threonine protein kinase
RMSD	root-mean-square deviation
RP-HPLC	reverse-phase high-performance liquid chromatography
rpm	revolutions per minute
SCAM	substituted cysteine accessibility method
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
Ste2p	α-factor pheromone receptor
Ste3p	**a**-factor receptor
TBS	tris-buffered saline
TBST	tris-buffered saline + Tween-20
TEMED	N,N,N',N'-tetramethylethylenediamine
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
TM	transmembrane
TOCSY	total correlation spectroscopy
TROSY	transverse relaxation optimized spectroscopy

TrpΔLE	portion of the histidine-tagged TrpΔLE1413 polypeptide.
UV	ultraviolet

LIST OF TABLES

Table Page

LIST OF FIGURES

Chapter I: Introduction

Membrane proteins

Membrane proteins currently constitute approximately 30% of the proteins encoded by the human genome (Wallin and von Heijne, 1998). Many of these reside in the plasma membrane and mediate flow of information or substances between the cytosol and extracellular space. Despite their large number and physiological significance relatively little is known about the three dimensional structures of membrane proteins. To date less than 1% of this group of biomolecules, mostly membrane channels and transporters, have been successfully studied at the atomic level (Lundstrom, 2004). Even less detailed structural information is available about membrane receptors. The structure of the extracellular portions of some membrane receptors has been determined, but knowledge of the structure of entire membrane receptors is very limited. Of the thousands of proteins in the plasma membrane of eukaryotic cells, only rhodopsin (Palczewski et al., 2000), aquaporin (Sui et al., 2001), and prostaglandin H_2 synthase-1 (Picot et al., 1994; Garavito et al., 1995) structures are known.

The primary reason for the relatively small number of membrane protein structures is that it is difficult to produce and solubilize multi-milligram quantities of stable membrane proteins for crystallization, and for NMR studies at relatively high concentrations (Grisshammer and Tate, 1995). Because of their hydrophobic nature membrane proteins are prone to aggregation when outside their native membrane, and are unstable in aqueous solution. It is also possible that some membrane proteins fail

to crystallize not because of poor stability, but because of extensive conformational fluctuations that prevent formation of an ordered array (Kiefer, 2003).

In the cell membrane proteins adopt their native conformation by inserting cotranslationally into the membrane with the help of a protein complex called translocon (Chin et al., 2002). The physical properties of the lipid bilayer such as bilayer thickness and lateral pressure may be crucial for stability and have been shown to effect the rate of refolding of bacteriorhodopsin (Curran et al., 1999) and other membrane proteins. Membrane proteins often become inactive when completely removed from lipids. In a number of cases interactions with specific lipids can be replaced by other lipids as long as certain structural features are present (Marsh, 1990; Lee, 1998; Lund et al., 1989; Levi et al., 2000; Callaghan et al., 1997), whereas in some other cases only certain lipids are required (Burger et al., 2000).

In 1990, Popot and Engelman proposed a two-stage model for the folding of integral membrane proteins where first, independently stable transmembrane helices are inserted into the membrane and then the helices associate with one another to form the native 3 dimensional structure (Popot and Engelman, 1990). The two-stage model was first tested experimentally using peptides corresponding to the seven transmembrane domains of bacteriorhodopsin in a lipid bilayer (Hunt et al., 1997a). Five of the seven peptides formed stable α helices, while the others showed low degrees of membrane partitioning and form non-native secondary structures. Similar studies of the yeast Ste2p receptor showed that five of the seven transmembrane domains form stable helices in micelles or vesicles, while the other two formed β sheets and/or aggregates (Xie et al., 2000). Both studies hypothesize that the folding

of these helices in the context of the full-length protein might occur by more complex processes and involve some external constraints or interactions with other helices (Hunt et al., 1997b; Popot and Engelman, 2000). A comprehensive analysis of helical propensities for a human GPCR adenosine A_2a receptor supports models for membrane-protein folding in which interactions between transmembrane domains are required for proper insertion and folding of some transmembrane helix domains (Lazarova et al., 2004).

Based on the two-stage model (Popot and Engelman, 1990) and the evident stability of single transmembrane helices inserted in the lipid bilayer one can study peptides corresponding to individual transmembrane domains, even when long range connectivity and forces involved in transmembrane helix association are absent (Hunt et al., 1997a; Reddy et al., 1994; Wigley et al., 1998; Deber et al., 1999). The relevance of this approach has been verified by the ability of bacteriorhodopsin and rhodopsin fragments to assemble into active forms (Marti, 1998; Katragadda et al., 2001a; Ridge et al., 1995; Ridge et al., 1996; Ridge et al., 1999). Since the insertion of transmembrane helices into the lipid bilayer is driven by hydrophobic forces (White and Wimley, 1999), studies of fragments of receptors in membrane mimetic solvents might be a good starting point for obtaining the biologically relevant conformation of regions of the receptor.

Although the features required for the subsequent association of inserted helices are not well understood (Stevens and Arkin, 1999), it has been suggested that polar interactions can mediate transmembrane helix-helix association (Bargmann et al., 1986; Machamer et al., 1993; Manolios et al., 1990; Zhou et al., 2000). Additional

studies have shown that polar side chains such as Asn (Choma et al., 2000; Gratkowski et al., 2001), small residues such as Gly and Ala (Eilers et al., 2000; Javadpour et al., 1999), and GxxxG motifs in transmembrane domains of membrane proteins, such as glycophorin A (Russ and Engelman, 2000; Senes et al., 2000), may play an important role in the association of some transmembrane helix pairs. Buried polar side chains occur less frequently but may be important for function, conformational specificity, and thermodynamic stability (for a review, see DeGrado et al., 2003). These studies provide some understanding of the forces that drive helix-helix interactions in the membrane, which might allow the prediction of the tertiary structures of transmembrane regions of integral membrane proteins.

A high frequency (~60%) of transmembrane helix kinks or deformations occurs at proline residues in membrane protein structures. Proline side chains usually prevent normal α-helix geometry, possibly generating these kinks, because of steric conflicts with the preceding residue and the loss of backbone hydrogen bonding (Reiersen and Rees, 2001). These proline-induced distortions in membrane proteins may be tolerated because of the high stability of helices in a membrane environment (Chamberlain et al., 2003; White and Wimley, 1999) and may be functionally important. As in rhodopsin where six of seven helices contain significant distortions (Palczewski et al., 2000), the majority of transmembrane helices contain significant irregularities from normal helix geometry (Chamberlain et al., 2003).

Not all distortions and kinks in helices occur at proline; ~40% of all transmembrane helix distortions are at nonproline residues. Although the causes of nonproline distortions are not well understood, it has been proposed that during

evolution nonproline kinks were caused by the introduction of prolines, which were later removed leaving nonproline kinks stabilized by tertiary interactions (Yohannan et al., 2004). Three prolines in bacteriorhodopsin transmembrane helices were changed to alanines with little structural consequences, and nonproline kinks in membrane proteins of known structure were found to be prolines in homologous sequences (Yohannan et al., 2004). Based on the different kink patterns predicted in different G protein-coupled receptors (GPCRs), GPCRs are expected to have different structures.

G protein-coupled receptors

Among membrane receptors the largest family is that of G protein-coupled receptors (Dohlman et al., 1991; Fraser et al., 1994), and GPCRs are the target of about two-thirds of drugs on the market (Archer et al., 2003; Gurrath, 2001). Currently, it is estimated that there are about 1000 GPCRs encoded by the human genome, yet only one high resolution structure, that of the GPCR of the visual system, bovine rhodopsin, elucidated at 2.8 Å resolution, has appeared in the literature (Palczewski et al., 2000). The structure of the proton pump bacteriorhodopsin determined at 1.55 Å (Luecke et al., 1999), although not a GPCR, has been used as a model for the GPCRs because of its homology to rhodopsin. Models of these proteins based on the X-ray structures and on calculated structures using hydropathy analysis and the assumption that the transmembrane domains are helical superimpose fairly well (Baldwin et al., 1997; Herzyk et al., 1998; Bourne and Meng, 2000). This finding resulted in working models for all GPCRs in which there

are seven transmembrane domains connected by intracellular and extracellular loops (Dohlman et al., 1991; Cramer et al., 1992; Baldwin, 1993; Baldwin et al., 1997; Herzyk and Hubbard, 1995; Herzyk and Hubbard, 1998; Luecke et al., 1999; Palczewski et al., 2000) and the amino and carboxy-terminal are located extracellularly and intracellularly, respectively (Cartwright and Tipper, 1991).

GPCRs have been found in a wide range of organisms and include receptors for peptide hormones and neurotransmitters, chemokines, prostanoids, proteinases, taste and odorant molecules, and photons and calcium ions (Watson and Arkinstall, 1994). GPCRs play an important role in intercellular signaling and regulation. Binding of a peptide hormone seems to stabilize the allosteric state of the GPCR (Marshall, 2001) that mediate cellular responses to a wide variety of physical and chemical extracellular signals (Baldwin, 1993) by activating heterotrimeric G proteins. These in turn initiate signal transduction pathways and generate second messenger cascades ultimately leading to transcriptional activation of specific genes (Dohlman et al., 1991; Dohlman and Thorner, 2001). In many systems, it has been suggested that the peptide ligands bind at the interface between the transmembrane segments and the extracellular loops, while nonpeptide antagonists bind within the transmembrane segments (Marshall, 2001).

Attenuation of cellular responses mediated by GPCR signaling can occur by uncoupling of activated receptors from G proteins, and/or agonist-induced receptor endocytosis contributing to both desensitization and resensitization of signaling (Böhm et al., 1997). GPCRs are involved in a wide variety of physiological processes and are implicated in numerous human diseases (Gether, 2000; Gurrath, 2001;

Shichida and Imai, 1998). Understanding the structure of GPCRs and the mechanism by which they activate downstream signaling can help to determine the cause of many pathological diseases and possibly their treatments (Spiegel and Weinstein, 2004).

Although all GPCRs have seven transmembrane helices and are thought to have a general mechanism of activation, their sequences vary widely (Palczewski et al., 2000; Karnik et al., 2003; Rashid et al., 2004). Different GPCR families have different patterns of helix kinking (Yohannan et al., 2004) and they interact with diverse ligands from small molecules to proteins indicating significant structural diversity within the GPCR family (Bockaert and Pin, 1999).

The GPCR model of Baldwin (Baldwin, 1993), which is based on the constraints of the electron density of rhodopsin at low resolution and sequence alignment of the rhodopsin family of GPCRs, has been used as a template for generating models of peptide GPCRs (Mosberg, 1999; Pogozheva et al., 1997). Although there is only 3.3 Å of rms error between the Baldwin template and the crystal structure of rhodopsin (Palczewski et al., 2000), certain side chains are misaligned by the error in the helix orientation for some of the transmembrane segments. The overall arrangement of the transmembrane helices of GPCRs is in a counterclockwise manner as viewed from the extracellular side of the membrane (Baldwin, 1993; Liu et al., 1995; Wess, 1997) and is in accordance with the arrangement found in rhodopsin (Palczewski et al., 2000).

***Saccharomyces cerevisiae* as a model system for intercellular signal transduction**

To gain a better understanding of receptor structure and ligand-receptor interactions the yeast *Saccharomyces cerevisiae* has been used as a model system because of the abundant biochemical and genetic knowledge on this organism and efficient laboratory tools that are available (Blumer and Thorner, 1991). In the budding yeast, pheromone binding activates a signaling pathway that stimulates mating between opposite haploid mating-type cells, *MAT***a** cells and *MAT*α cells, to form diploid *MAT***a**/α cells (Kurjan, 1993). The *MAT*α cells, α-mating type, secrete a 13-residue peptide pheromone, α-factor, which is produced by proteolytic cleavage of large precursors and binds to the α-factor receptor (Ste2p) on the surface of the opposite mating type (*MAT***a** cells) only. The *MAT***a** cells secrete **a-**factor pheromone, a mixture of two 12-residue lipopeptides, which are processed from two similar precursors (Duntze et al., 1994) and bind to the **a**-factor receptor (Ste3p) on the surface of the *MAT*α cells.

Binding of the mating pheromones to their cognate receptor triggers cellular responses that allow cells to prepare for mating, including agglutination, cell cycle arrest, transcriptional activation of certain genes, morphological changes leading to the growth of a projection from the cell surface (Dohlman et al., 1991; Dohlman and Thorner, 2001; Blumer and Thorner, 1991; Bardwell et al., 1994), cell fusion, and nuclear migration to the center of the fused cell forming a zygote. Depending on the nutrients the zygote can undergo meiosis and sporulation to produce α and **a** cell types under limiting nutrient conditions or undergo asexual reproduction (mitotic

growth) to form daughter diploid buds under sufficient nutrient conditions (Sprague and Thorner, 1992) [Figure 1].

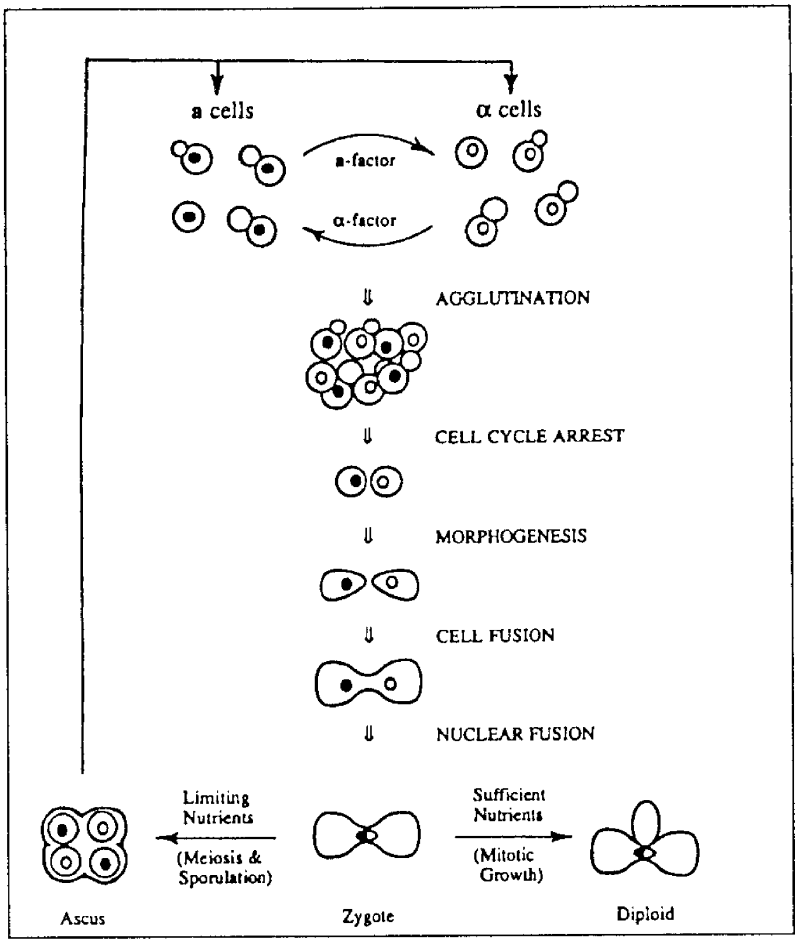

Figure 1. Representation of the mating process in *Saccharomyces cerevisiae**

* Taken from Sprague, G. F., and Thorner, J. W. The Molecular and Cellular Biology of the Yeast *Saccharomyces* Gene expression, Cold Spring Harbor Lab. Press, 1992, page 661.

Ste2p: α-factor pheromone receptor

The α-factor receptor (Ste2p), encoded by the *STE2* gene (Blumer and Thorner, 1991; Nakayama et al., 1985) interacts with a heterotrimeric guanine nucleotide-binding protein (G protein). This G protein is composed of three subunits

Gα (Gpa1p), Gβ (Ste4p), and Gγ (Ste18p), encoded by the *GPA1*, *STE4*, and *STE18* genes, which exhibit functional and sequence homology to the α, β, and γ subunits, respectively, of mammalian trimeric G proteins (Sommers and Dumont, 1997). Although the sequences of yeast and mammalian receptors are extremely different, it was demonstrated that coexpression of a mammalian beta 2-adrenergic receptor and a mammalian G protein alpha subunit in yeast can control yeast mating signal transduction by coupling to each other and to downstream effectors (King et al., 1990). It has also been shown that some heterologously expressed mammalian receptors are capable of directly activating the yeast G protein (Price et al., 1995, 1996), and the yeast α-factor receptor is capable of activating at least one mammalian G protein (Crowe et al., 2000). The Gα subunit has been suggested to anchor the Gβγ to the plasma membrane (Hirschman et al., 1997), and the C-terminus of Gα is believed to interact with the **a**- and α-factor receptors (Kallal and Kurjan, 1997). Activation of the receptor by pheromone binding leads to the dissociation of the trimer (Liri et al., 1998). The released βγ subunits initiate a downstream mitogen-activated protein kinase MAPK cascade, encoded by *STE20*, *STE11*, *STE7*, *and FUS3*, leading to the above-described mating responses (Leberer et al., 1997). Interaction of a serine-threonine protein kinase (Raf-1) with Gβγ after the release of the Gβγ complex from the membrane is believed to play a role in the mitogenic signal pathway (Pumiglia et al., 1995).

The pheromone receptors, Ste2p and Ste3p, involved in mating in the yeast *Saccharomyces cerevisiae* belongs to Class D (fungal pheromone receptors) of the large family of GPCRs (Dohlman et al., 1991; Fraser et al., 1994; Horn et al., 2001).

Although this class is unique to fungal species, mutational analysis indicates that the α-factor receptor and other GPCRs are structurally and functionally similar (Parrish et al., 2002; Dube et al., 2000; Eilers et al., 2005). However, this yeast GPCR has minimal homology with members of the rhodopsin family of GPCRs. The Ste2p and Ste3p receptors contain seven-transmembrane spanning domains and have been considered as a model system for studying the seven-transmembrane segments receptor family (Jeansonne, 1994; Dohlman, 2002). Although Ste2p and Ste3p share no significant sequence homology, their functional domains are organized similarly and they both activate the same G proteins and down stream signal transduction pathway (Blumer et al., 1988; Blumer and Thorner, 1991). However, there is a 50% similarity in the amino acid sequence of Ste2p to that of α-factor receptor of yeast *S. kluyveri*, which is closely related to *S. cerevisiae* (Marsh, 1992).

The predicted location of the domains of Ste2p based on the topology of accepted GPCRs models and the hydrophobicity index of the primary sequence of Ste2p is illustrated in Figure 2 (Reddy et al., 1994). A difference in the position of one or two residues at the membrane interface is observed in most of the transmembrane domains of different Ste2p topology models (Reddy et al., 1994; Sommers and Dumont, 1997; Martin et al., 2002). Recent studies of the substituted cysteine accessibility method (SCAM) analysis of Ste2p have provided the first direct evidence for Ste2p topology (Lin et al., 2003, 2004), where there is no difference in the position of residues at the extracellular membrane interface of the fourth, fifth, and seventh transmembrane domains as compared to Reddy's topology (Figure 2). However, there are differences in the position of one residue in the first and the

second transmembrane domains, two residues in the sixth transmembrane region, and three residues in the third transmembrane domain.

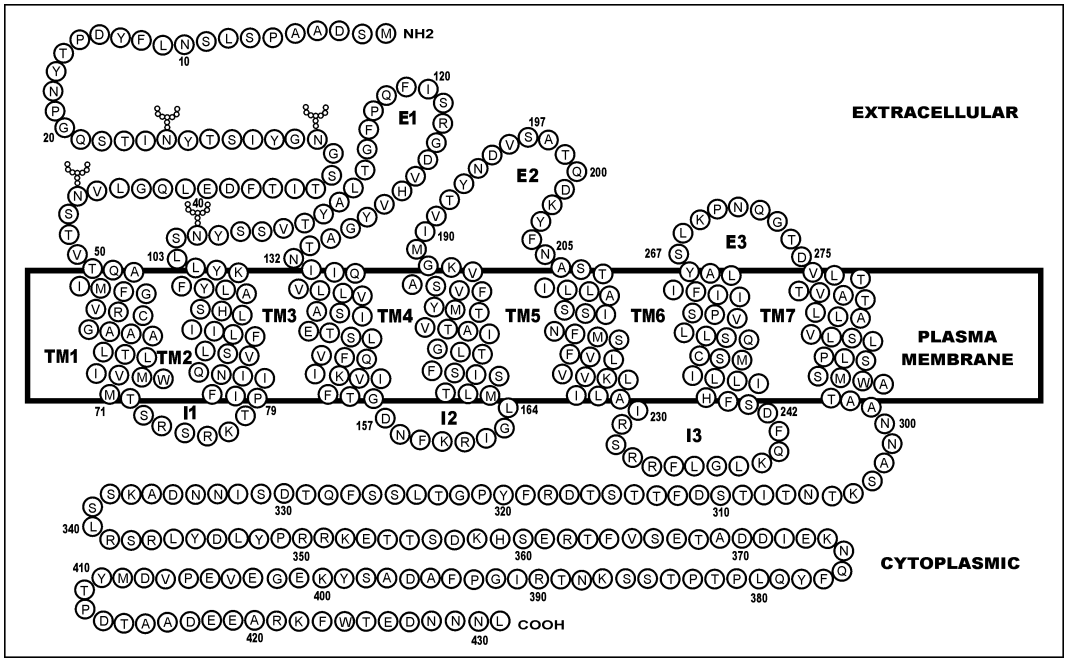

Figure 2. Cartoon of the α-factor receptor (Ste2p). Domains are indicated by the following: E, extracellular loops; I, intracellular loops; TM, transmembrane domains. Four of the Asn residues are represented as glycosylation positions.

The transmembrane segments and the extracellular loops of Ste2p are thought to form a pocket that acts as the ligand-binding domain because mutations within the transmembrane segments alter ligand-binding specificity (Konopka et al., 1996). The extracellular receptor surface (including the extracellular NH_2-terminal domain, different extracellular loops and the exofacial portions of various transmembrane domains) has been shown to be critical for ligand binding (Sen and Marsh, 1994; Sen et al., 1997). In Family A GPCRs, the third extracellular loop with an average length of 10-11 residues, which connects the sixth and seventh transmembrane domains, may constrain the transmembrane domains, thereby limiting their motion and preserving the ground state of the receptor (Lawson and Wheatley, 2004). Chimeric

angiotensin II (AngII) receptors constructed of portions of the AT_2 receptor substituted into the AT_1 receptor demonstrated that the AT_2 third extracellular loop is important for high-affinity binding of the AT_2 receptor-selective agonist CGP 42112A and activation in response to the AT_2-selective ligand (Hines et al., 2001).

Extracellular ends of the transmembrane domains of Ste2p have been also suggested to play a special role in ligand binding (Dosil et al., 2000; Lee et al., 2001; Akal-Strader et al., 2002). Aromatic residues at the extracellular ends of the fifth and sixth transmembrane domains of Ste2p have been suggested to play a role in promoting the interaction between α-factor and the ends of the transmembrane domains in the receptor, and similar aromatic residues may play an important role in other GPCRs (Lin et al., 2003). The identification of various residues in the extracellular regions of mammalian GPCRs that bind peptide ligand to be important for ligand binding suggests a possible common aspect to receptor activation by diverse ligands (Ji et al., 1998; Marshall, 2001).

The intracellular receptor surface is known to be involved in G protein recognition and activation (Staley and Kim, 1994; Wess, 1997). Mutational analysis suggested that the third intracellular loop is important for receptor and G protein activation in the yeast mating pathway (Clark et al., 1994; Stefan and Blumer, 1994; Weiner et al., 1993; Celic et al., 2003). The third intracellular loop has an amphipathic nature in Ste2p (Blumer and Thorner, 1991) and other GPCRs such as the m2 muscarinic receptor (Kostenis et al., 1997). Although this amphipathic cytoplasmic loop was predicted to favor a helical structure at the membrane interface (Blumer and Thorner, 1991), the only structural evidence available suggests a random

coil in water (Arshava et al., 1998). The possible helical structure as predicted by the helical wheel could be influenced by the phospholipid head groups, which could neutralize some of the positively charged side chains in the loop and influence its structure. Coupling of the G protein to the third intracellular loop of the receptor may induce a conformational change in the receptor, thereby increasing the affinity for the ligand (Shah and Marsh, 1996).

The transmembrane domains of GPCRs are essential for ligand binding and signal transduction (Konopka and Jenness, 1991). The seventh transmembrane domain of angiotensin II (AT_2) receptor has been demonstrated to be important for high-affinity binding and activation in response to the AT_2-selective ligand (Hines et al., 2001). The first and sixth transmembrane regions of Ste2p have been found to be important for ligand interaction, ligand specificity, and/or receptor activation to initiate the signal transduction pathway (Abel et al., 1998). Mutational studies have suggested a conserved role for the cytoplasmic end of the third transmembrane domain in regulating receptor activity of divergent GPCRs, including Ste2p receptor (Parrish et al., 2002).

As suggested by the two-stage model for the folding of integral membrane proteins independently stable transmembrane helices are inserted into the membrane and then the helices interact with one another to form the tertiary fold structure (Popot and Engelman, 1990). In GPCRs, interaction of transmembrane segments has been shown in several studies (Sommers and Dumont, 1997; Bourne, 1997; Farrens et al., 1996; Dube et al., 2000; Suryanarayana et al., 1992; Liu et al., 1995; Parrish et al., 2002; Dube and Konopka, 1998; Neumann et al., 2001; Lin et al., 1997; Biebermann

et al., 1998; Angelova et al., 2000). In a random mutational study of transmembrane regions of the Ste2p receptor, it was suggested that the fifth transmembrane domain (residue Met218) interacts with the third transmembrane domain (residues Ala140, Glu143, and Thr144), and that the third transmembrane domain (residue Glu143) interacts with the sixth transmembrane region (residue Tyr266) [Sommers and Dumont, 1997]. The Y266C and E143K mutations were each defective when tested individually but could be combined to create a partially functional receptor (Sommers and Dumont, 1997). Direct contact of the third and fifth transmembrane segments would be consistent with the proximity of these two transmembrane segments in the structural models of GPCRs (Baldwin, 1993). However, the first and third transmembrane segments are predicted to be on opposite sides of the second transmembrane segment, too far apart to interact directly (Baldwin, 1993). In addition, biochemical and biophysical studies have suggested that binding of ligand to GPCRs changes helix-helix interactions between the third and sixth transmembrane segments which, in turn, affects the conformation of intracellular regions of receptors, uncovering a site that interacts with the G protein (Bourne, 1997; Farrens et al., 1996). In Ste2p, interaction between the fifth (residue Val223) and the sixth (residue Leu247) transmembrane domains has been also reported (Dube et al., 2000). Mutational analyses on the adrenergic receptors (Suryanarayana et al., 1992) and the muscarinic acetylcholine receptor (Liu et al., 1995) reveal an important interaction between the first and the seventh transmembrane regions.

Previous studies have shown that transmembrane domains six and seven (TM6 and TM7) of GPCRs interact (Parrish et al., 2002; Dube and Konopka, 1998;

Neumann et al., 2001; Lin et al., 1997) and are likely involved in the stimulation of cellular responses (Konopka et al., 1996). The contact sites between polar residues Gln253 from TM6 and Ser288 and Ser292 from TM7 have been inferred from mutagenesis studies and are thought to play an important role in regulating Ste2p receptor function (Dube and Konopka, 1998). Molecular modeling of the inactivating TSH receptor mutation (D633R) in TM6 suggested the existence of a hydrogen bond between position 633 in TM6 and 674 in TM7, which was confirmed by the partial response to TSH stimulation in the TSH receptor double mutants (D633R/N674D) designed to break up possible H bonds (Neumann et al., 2001). Several pairs of strong interhelical side chain-side chain H bonds in TM6 and TM7, which have been proposed to stabilize the inactive receptor conformation, were revealed by a model of the LH receptor (Lin et al., 1997). During the activation of the TSH and LH receptors rearrangements of TM6 (Biebermann et al., 1998) and TM7 (Angelova et al., 2000), respectively, were mediated by hydrogen bond weakening or release.

The cytoplasmic C-terminus is a target for a receptor kinase that negatively regulates receptor signaling and functions in ligand-induced endocytosis (Dohlman and Thorner 2001; Feng and Davis, 2000). Receptor endocytosis and down-regulation leading to receptor desensitization is mediated by receptor phosphorylation by G protein receptor kinases (Böhm et al., 1997). After rapid endocytosis of the Ste2p receptor following exposure to the α-factor pheromone, new receptors are localized to the cell surface where they can bind pheromones (Hicke et al., 1998). In yeast, several Ser-Thr residues within the cytoplasmic C-terminus are phosphorylated upon exposure of the cells to the pheromone ligand (Feng and Davis, 2000).

Phosphorylation of serine residues in the SINNDAKSS sequence within the C-terminal tail of Ste2p has been involved in receptor endocytosis (Hicke et al., 1998; Rohrer et al., 1993). Ubiquitination is an endocytosis signal for Ste2p because mutation of lysine residues in the SINNDAKSS sequence within the C-terminal tail abolishes ubiquitination (Hicke et al., 1998). Ligand binding induces ubiquitination of the Ste2p C-terminal tail (Hicke and Riezman, 1996), which in turn induces endocytosis of Ste2p by a clathrin-dependent mechanism, and the internalized receptor is transported to and degraded in the vacuole (Tan et al., 1993; Schandel and Jenness, 1994).

GPCR activation involves rearrangement of the intracellular receptor surface, caused by ligand-induced changes in the relative orientation of individual transmembrane helices. Thus, rearrangement enables the G protein to interact with previously inaccessible residues on the receptor protein (Staley and Kim, 1994; Wess, 1997). Mutations that cause constitutive receptor activity have also been studied to learn about the process of GPCR activation. Constitutive mutations mimic the effects of ligand binding because they alter receptor structure such that G protein activation occurs in the absence of added ligand. Mutational analysis has shown that a limited number of mutations in Ste2p, some of them in the sixth transmembrane domain, have lead to a constitutively activated receptor (Sommers et al., 2000). Mutation of Pro258 to Leu and a double mutation of Pro258 to Leu and Ser259 to Leu in transmembrane domain 6 (TM6) of Ste2p resulted in constitutively active receptors with 45% and over 90% activation, respectively, compared to that of the maximal wild-type cells stimulated with α-factor (Konopka et al., 1996). Recent reports

indicate that proline residues in transmembrane domains can act as hinges that present significant conformational flexibility on these polypeptides (Williams and Deber, 1991; Li et al., 1996; Sansom and Weinstein, 2000). Since Pro258 occurs in the central portion of TM6, and since proline residues are expected to cause a kink in the α-helical domains, the Pro258Leu mutation is predicted to alter the structure of TM6 and thereby play an important role in the activation of Ste2p (Stefan et al., 1998) by changing the spatial relationship of sequences in the third intracellular loop (contiguous with TM6) that are thought to interact with the heterotrimeric G protein. The high-resolution structure of rhodopsin showed that there was a significant bend at Pro267 of the sixth transmembrane helix and that there were intramolecular interactions between various transmembrane regions, which were stabilized by a number of interhelical hydrogen bonds and hydrophobic interactions (Palczewski et al., 2000).

A conserved proline residue in transmembrane domain six is believed to play an important role in the activity and trafficking of GPCRs (Konopka et al. 1996). More than 90% of all GPCRs contain a proline residue at a similar position in TM6 with important functional consequences (Baldwin, 1993; Stefan et al., 1998), suggesting that this aspect of the receptor activation may be conserved in other receptors (Konopka et al., 1996).

Previous studies have proposed that the binding of peptides to membranes (Booth, 2003; Oren et al., 2002) and the insertion of transmembrane domains, and the initiation of the helix formation and folding (Lazarova et al., 2004) are driven by electrostatic interactions between peptides and negatively charged lipids. Sequence

variation in TM domains has been shown to result in different electrostatic effects among peptides (Lazarova et al., 2004). However, hydrophobic interactions have also been shown to be a major determinant for the folding of transmembrane domains (Lazarova et al., 2004).

Although previous studies have shown a correlation between the hydrophobicity and helical propensity of peptide segments in lipid systems (Li and Deber, 1993), peptides from the Ste2p and A_2a receptors show no correlation (Xie et al., 2000; Lazarova et al., 2004). For example, TM6 peptides were among the most hydrophobic, yet they were the least helical. Although TM7 peptides were the least hydrophobic, they were among the most helical in both micelles and vesicles (Xie et al., 2000; Lazarova et al., 2004). Neither the sequences of these peptides nor the presence of proline residues correlated with lower helicity. Proline, known as a helix breaker in aqueous solutions, showed greatly enhanced helical propensity in a membrane-mimetic environment and organic solvents (Li et al., 1996).

The involvement of TM6 in ligand binding and signal transduction in many GPCRs (Gether, 2000; Huang et al., 2001) suggests that the low helicity of TM6 peptides in the Ste2p receptor, the human adenosine A_2a receptor, and bacteriorhodopsin relates to their presumed role in signaling (Popot and Engelman, 2000; Lee et al., 2002), where flexibility for conformational changes and interaction with the other helices to facilitate stabilization of their helical structures are needed (Lazarova et al., 2004). Although TM3 from Ste2p receptor is also a poor helix and is believed to undergo significant conformational changes upon GPCR activation (Xie et al., 2000), TM3 in A_2a receptor is among the most helical, while TM2 is a poor

helix. The difference in these properties among Ste2p, A_2a, and the GPCR family, either indicates that helicity, conformational stability, and signaling mechanism are not conserved or that different TM helices can play different roles in different GPCRs (Yohannan et al., 2004). Understanding the structure and function of the different loop regions and TM segments of Ste2p may give insights on the signal transduction mechanism of Ste2p and other GPCRs.

Biophysical studies of membrane proteins

Structural features of membrane proteins have not been widely obtained by X-ray crystallography (Palczewski et al., 2000) because of the lack of success in obtaining good crystals of these proteins. Therefore, efforts have turned towards the utilization of NMR spectroscopy. However, the size of integral membrane proteins and the requirement of membrane-like vesicles for solubilization and stabilization of these proteins prohibit the direct determination of their structure by solution NMR at this time. To overcome difficulties in conducting X-ray and NMR studies on intact receptors many laboratories have begun studying fragments of these molecules to generate atomic level models of domains (Mierke and Giragossian, 2001; Naider et al., 2001). Since the stabilization of secondary structures is highly dependent on local interactions, it is expected that the conformational features of the receptor domains and of the complementary amino acid sequences in the intact receptor will be similar. Experimental evidence that supports this approach is provided by the fact that reconstituted or coexpressed fragments of Ste2p, bacteriorhodopsin, and rhodopsin assemble into non-covalently linked domains maintaining functions of the full protein

(Martin et al., 1999; Marti, 1998; Katragadda et al., 2001a; Ridge et al., 1995; Ridge et al., 1996; Ridge et al., 1999; Yeagle et al., 2001). Although there is some evidence that long-range effects may have an influence on local secondary structures (Hunt et al., 1997a), most studies support the idea that studies on GPCR fragments can provide insights into the structure of the intact receptor (Albert and Yeagle, 2000). The 3D structure of rhodopsin, calculated from the NMR structures of individual transmembrane regions (Katragadda et al., 2001b; Yeagle et al., 2000a; Chopra et al., 2000), was in reasonable agreement with the crystal structure of the full GPCR (Palczewski et al., 2000). However, comparison of the secondary structure of the N-terminus and the three interhelical loops of rhodopsin determined by NMR (Yeagle et al., 2000b) with the crystal structure was not possible because of the lack of crystallographic structural details on these regions. Biophysical studies on single domains of Ste2p (Reddy et al., 1994; Arshava et al., 1998; Xie et al., 2000; Ding et al., 2001) and NMR analysis of the seven transmembrane domains of this receptor indicated significant conformational diversity for these regions of the protein (Naider et al., 2001; Arshava et al., 2002).

Virtually all biophysical investigations on GPCR fragments have been limited to peptides containing 30-40 residues, and even for these relatively short peptides, no high-resolution structures of these fragments in detergent have been published. The reason for the dearth of information on longer regions of GPCRs and for the lack of high resolution studies in micelles are a result of the difficulties encountered in obtaining multi-milligram quantities of isotopically labeled peptides containing one or more transmembrane domains. Solution and solid-state NMR analyses of small

membrane proteins and peptide fragments in lipid bilayers and micelles are becoming more possible by the use of isotopic labeling and highly efficient protein expression systems. The uniformly [^{15}N]-labeled major coat protein of fd bacteriophage was examined in phospholipid bilayers using solid-state NMR (Marassi et al., 1997). Opella and coworkers have used solid-state NMR for several structural studies (Opella, 1997; Marassi and Opella, 1998). Multidimentional solution NMR experiments in sodium dodecylsulfate (SDS) micelles were used to study the Pf1 (Schiksnis et al., 1987) and M13 bacteriophage coat proteins (van de Ven et al., 1993).

Organic-aqueous solvents have been widely used as membrane mimetics. The relevance of structures determined in aqueous trifluoroethanol (TFE) to the biological state of a protein has been discussed almost since the introduction of this solvent in conformational analysis of peptides (Tamburro et al., 1968; Goodman et al., 1971). It is generally considered to induce helical secondary structures in peptides. Nevertheless, many studies on peptides and protein fragments indicate a reasonably good correlation between the TFE-induced structure of the peptides and the structure of these peptides in the intact protein (Sonnichsen et al., 1992). Unlike TFE/water mixtures, chloroform/methanol/water ($CDCl_3/CD_3OH/H_2O$) system has not been considered to be a helix inducing medium. Indeed a high-resolution structure of the *E. coli* EmrE transport protein in chloroform/methanol/water (6:6:1) has been taken to represent the native structure of this molecule (Schwaiger et al., 1998). The structure of subunit c of the F_1F_0 ATP synthase has also been determined in chloroform/methanol/water (4:4:1) (Rastogi and Girvin, 1999).

Given the absence of biophysical data on conformational changes in GPCRs that occur upon activation by ligands, studies on structural aspects of Ste2p may provide an understanding of the mechanism of signal transduction by Ste2p and other GPCR family members. Since TM6 and TM7 of GPCRs interact (Parrish et al., 2002; Dube and Konopka, 1998) and are likely involved in the stimulation of cellular responses (Konopka et al., 1996), studies on their structure and in particular on TM6 interacting with TM7 would have significant ramifications on elucidating Ste2p function. An important role has been suggested for the angiotensin II AT_2 third intracellular loop in AT_2-selective ligand binding (Hines et al., 2001), the third intracellular loop in receptor and G protein activation (Clark et al., 1994; Stefan and Blumer, 1994; Weiner et al., 1993; Celic et al., 2003), and the cytoplasmic C-terminus of Ste2p in regulating receptor signaling and desensitization (Dohlman and Thorner 2001; Feng and Davis, 2000; Hicke et al., 1998; Rohrer et al., 1993; Tan et al., 1993; Schandel and Jenness, 1994). Based on the above mentioned findings structural studies on a double domain peptide containing the third intracellular loop, the sixth transmembrane domain, the third extracellular loop, the seventh transmembrane domain, and part of the cytoplasmic C-terminus tail of Ste2p will be very useful in contributing to the tertiary structure of the intact Ste2p receptor and other GPCRs.

Escherichia coli as a host for peptide expression

Escherichia coli has been widely used as the host of choice for producing proteins for structural studies because of the availability of a variety of specialized

expression vectors, rapid cell growth, low cost per gene expressed, and availability of well-developed methods for isotopically labeling target proteins for NMR and X-ray diffraction studies (Edwards et al., 2000; Goulding and Perry, 2003). Even though the natural bilayer environment of GPCRs is somewhat different from *E. coli* lipids, GPCRs can be expressed and show native pharmacology in *E. coli* (Strosberg et al., 1992; Marullo et al., 1988; Chapot et al., 1990), suggesting that GPCRs can tolerate a variety of lipid interactions (Kiefer, 2003).

Because overexpression of functional eukaryotic membrane proteins in bacteria resulted in toxicity, partial proteolysis of many heterologous membrane proteins and consequently low yields (Grisshammer and Tate, 1995), an alternate method expressing membrane proteins in cytoplasmic aggregates (inclusion bodies) followed by refolding in vitro has been applied. Inclusion bodies can be promoted by slow folding of the fusion protein, which prevent targeting of the protein to membranes and thus yield a dramatic increase in expression levels, by a factor of 1000 to 10,000 as compared to functional expression in the membrane (Kiefer, 2003). Some proteins, such as mitochondrial transporters (Fiermonte et al., 1993), form aggregates spontaneously in the cytoplasm, while others, such as GPCRs, can form aggregates by introducing an N-terminal fusion, such as amino-terminal polyhistidine-tag fusion, or by introducing positively charged amino acids into the loop regions to prevent translocation (Kiefer et al., 1996; Kiefer et al., 2000).

Based on the above mentioned advantages of using *E. coli* as a host for peptide expression, we decided to carry our expression of isotopically labeled

peptides in *E. coli* cells using methods to promote protein aggregation into inclusion bodies to prevent toxicity.

Objective

Structural information on single and multiple domains of Ste2p in membrane mimetic solvents and micelles will be obtained by solution-state NMR, to gain high resolution information relevant to the structure of the intact receptor and to other membrane proteins. Domains of Ste2p will be biosynthesized and analyzed via CD and NMR. Although high-resolution information was obtained in NMR studies on single transmembrane domain peptides in TFE/H_2O (Arshava et al., 2002), a membrane mimetic solvent, difficulty in studying these peptides in detergent was encountered. This problem may be overcome by studying biosynthetic uniformly [^{15}N]-labeled and [^{13}C]-labeled peptides by heteronuclear NMR. Biosynthetic approaches will be used to prepare [^{15}N]- and [^{13}C]-labeled single and multiple domain peptides from Ste2p in 10 mg quantities.

First Aim

Since the Pro258Leu mutation in TM6 is predicted to alter the structure of TM6 and thereby play an important role in the activation of Ste2p (Stefan et al., 1998), wild type and mutant (P258L) TM domain 6 (TM6) peptides, residues 238-270, was chosen for structural studies. Success toward this objective will enable direct comparison of the structure of peptides corresponding to the wild type and constitutively active receptor. This part of the project was carried out as follows:

- Biosynthesize and [^{15}N]-label wild type and mutant (P258L) transmembrane domain 6 (TM6), residues 238-270.
- Express peptides as fusion proteins using histidine-tagged TrpΔLE as the N-terminus.
- Purify the wild type and mutant (P258L) unlabeled- and [^{15}N]-labeled M6 fusion proteins {[^{15}N]-M6FP and [^{15}N]-M6FP(P258L)} via reverse phase-high performance liquid chromatography (RP-HPLC).
- Cleave with cyanogen bromide (CNBr) at a methionine residue to remove the TrpΔLE.
- Purify the cleaved product further with HPLC.
- Confirm peptides molecular weights by mass spectrometry.
- Analyze the secondary structure of the wild type and the mutant unlabeled-M6 peptides via circular dichroism (CD) in organic aqueous media, micelles or detergents, and vesicles.
- Determine the three dimensional structure of the [^{15}N]-labeled-M6 and [^{15}N]-labeled-M6(P258L) peptides using heteronuclear magnetic resonance spectroscopy (NMR).

Second Aim

Although it is much more difficult than a single domain, our ultimate goal is to biosynthesize a double transmembrane domain (containing peptide I3-M6-E3-M7- and 40 residues of the tail) of Ste2p and study its structure. In order to accomplish this goal we started by studying [^{15}N]-labeled E3-M7-24-T40, a 73 amino acid

residue multidomain peptide. The peptide consists of the third extracellular loop (267-275), transmembrane domain 7 (276-299), and 40 amino acid residues of the carboxyl terminus (300-339) of Ste2p attached. Since the carboxyl terminus of Ste2p is highly hydrophilic, the hydrophilic 40 residues in the carboxyl terminus tail were expected to increase the solubility of the TM7 domain. For expression purposes comparison was made of the TrpΔLE fusion peptide method (Miozzari and Yanofsky, 1978; Kleid et al., 1981) and a thioredoxin fusion protein expression system that had been proposed as an efficient method to biosynthesize transmembrane domains (Therien et al., 2002). The second part of the project was carried out as follows:

- Express the unlabeled- and [^{15}N]-labeled-E3-M7-24-T40 peptides as a TrpΔLE fusion protein (M7FP).
- Purify the M7FP via RP-HPLC.
- Release the E3-M7-24-T40 peptide from the M7FP via CNBr cleavage.
- Purify the E3-M7-24-T40 peptides via RP-HPLC.
- Carry out CD studies on unlabeled-E3-M7-24-T40 peptide in organic-aqueous media, micelles or detergents, and vesicles.
- Determine the three dimensional structure of the E3-M7-24-T40 peptide via NMR in membrane mimetic solvents and in detergents.

Chapter II: Materials and Methods

Synthetic peptides

T40 [Ste2p(300-339)] and M7-24-T40[Ste2p(276-339)] were synthesized as previously described (Naider et al., 2003).

Vectors

The parent plasmid pMD194 containing the wild-type *STE2* gene used to generate Ste2p mutant plasmid pMD602 (Met250Ala) was obtained from Mark Dumont [University of Rochester] (Martin et al., 1999). pMD602 plasmid has a DNA sequence of *STE2* with C252S, and M250A mutations. Generation of Met^{250} mutant was carried out by Enrique Arevalo at Dr. Mark Dumont's laboratory, Rochester (personal communication) using the single stranded DNA template method (Kunkel et al., 1987). The plasmid pMMHa, which was used to construct plasmids pSW02 and pREJ02, was obtained as a gift from Dr. Peter Kim [Massachusetts Institute of Technology] (Staley and Kim, 1994). pMMHa plasmid expresses the 6.6 kDa bovine pancreatic trypsin inhibitor (BPTI), and is a pET-derived plasmid that contains the gene for β-lactamase, the T7 RNA polymerase binding site, the N-terminus TrpΔLE leader peptide derived from the tryptophan operon, and restriction enzyme cloning sites (*Bam*H I and *Hin*d III) (Staley and Kim, 1994).

The pMD602 plasmid (Figure 3A; Martin et al., 1999) and the pMMHa vector (Figure 3B; Staley and Kim, 1994) were used to create the pSW02 and pREJ02 plasmids (Figure 3C,D) expressing the TrpΔLE-M6 and TrpΔLE-E3-M7-24-T40

peptides, respectively. These vectors utilize a T7 expression system and express a polypeptide fused to a portion of the histidine-tagged TrpΔLE1413 polypeptide (Miozzari and Yanofsky, 1978; Kleid et al., 1981; Staley and Kim, 1994) at the N-terminus. The vector pSW02 (Figure 3C) contained the DNA sequence of M6 (residues Lys239 through Lys269 of Ste2p, Figure 2) that included the mutation Met250Ala and the addition of one non-native flanking lysine (KKQFDSFHILLIASSQSLLVPSIIFILAYSLKK) on each side of the natural sequence. The vector pREJ02 (Figure 3D) contained the DNA sequence of E3-M7-24-T40 [residues Ser267 through Ser339 of Ste2p comprising the third extracellular loop (E3), the 24-residue seventh transmembrane domain (M7) and 40 residues of the cytoplasmic tail (T40), Figure 2].

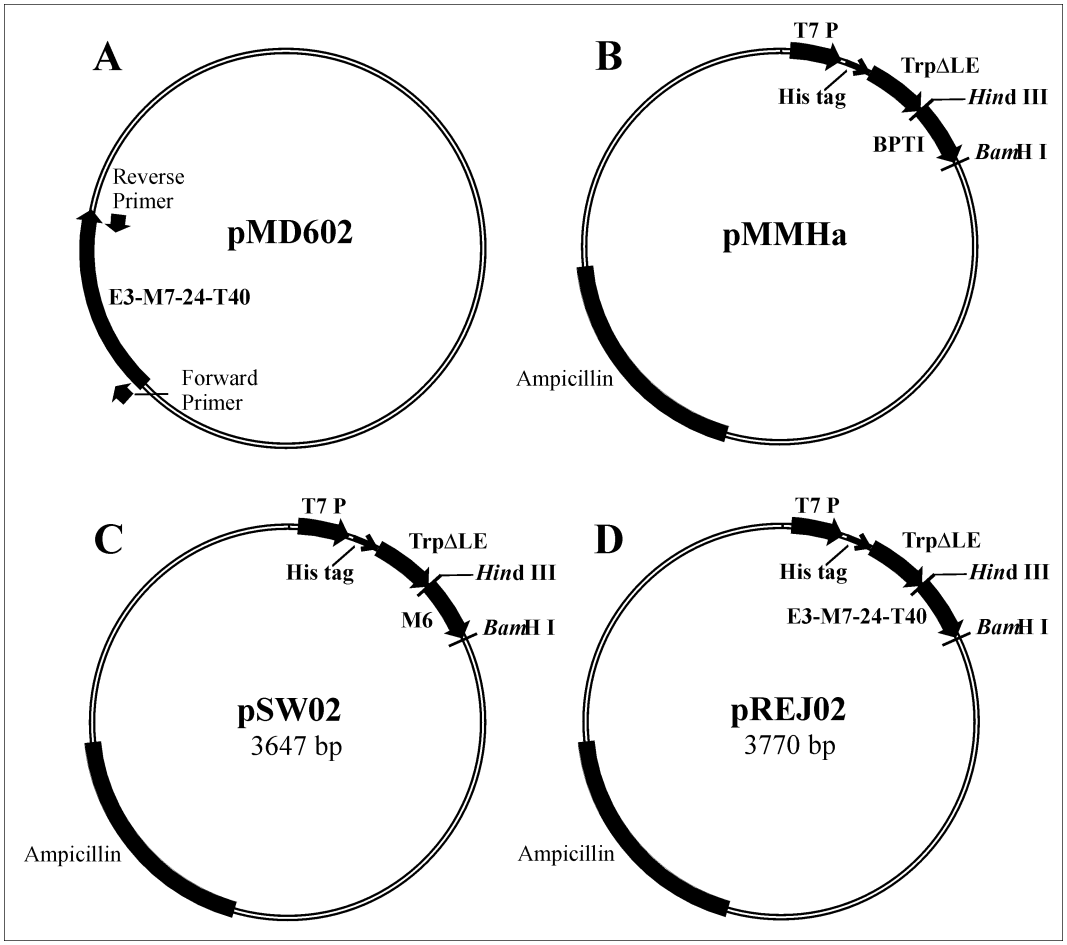

Figure 3. Plasmids used for the expression of M6 fusion protein (M6FP) and M7 fusion protein (M7FP). (A) The pMD602 plasmid contains the DNA sequence of the entire Ste2p and was used to PCR amplify the E3-M7-2-T40 region of Ste2p. (B) The pMMHa plasmid serves as the vector for the E3-M7-24 T40 insert. It contains the ampicillin gene, T7 RNA polymerase binding site, His-tag-TrpΔLE leader sequence and the bovine pancreatic trypsin inhibitor (BPTI) sequence. (C) The pSW02 plasmid was constructed from the pMMHa and pMD602 plasmids. It depicts the ampicillin gene, T7 RNA polymerase binding site, His-tag-TrpΔLE leader sequence and the M6 sequence between *Hin*d III and *Bam*H I restriction enzyme sites. (D) The pREJ02 plasmid was constructed from the pMMHa and pMD602 plasmids. It depicts the ampicillin gene, T7 RNA polymerase binding site, His-tag-TrpΔLE leader sequence and the E3-M7-24-T40 sequence between *Hin*d III and *Bam*H I restriction enzyme sites.

The vectors pSW02 and pREJ02 were constructed as previously described (Arevalo et al., 2003; Naider et al., 2004). A list of the primers that were used to amplify the coding sequences for single and multi transmembrane domains of Ste2p

are listed in Table 1. Polymerase chain reaction (PCR) was used to amplify the DNA sequence of M6 peptide from pMD602 (Figure 3A) using a forward 5' primer [Fwd(M6)] and a reverse 3' primer [Rev(M6)] (Table 1). The DNA sequence of E3-M7-24-T40 was PCR amplified from pMD602 using a forward 5' primer [Fwd(M7)] and a reverse 3' primer [Rev(M7)] (Table 1). Both of the forward primers, Fwd(M6) and Fwd(M7), added a cyanogen bromide (methionine) cleavage site as well as a unique *Hin*d III restriction site at the N-terminal of the target peptide sequence. Both of the reverse primers, Rev(M6) and Rev(M7), added a unique *Bam*H I restriction site followed by a termination codon at the C-terminal of the of the target peptide sequence. The amplified region of DNA was digested with *Hin*d III and *Bam*H I and ligated into the similarly digested pMMHa plasmid (Figure 3B) containing the His tag-TrpΔLE sequence. The resulting pSW02 and pREJ02 plasmids (Figure 3C,D) contained a His tag-TrpΔLE sequence followed by the DNA sequence of M6 and E3-M7-24-T40 of Ste2p, respectively (Table 2 & 3). In order to release E3-M7-24-T40 protein from the TrpΔLE with CNBr Met^{294} was mutated to Leu at the University of Tennessee. The Met250Ala mutation in M6 region of Ste2p was already achieved in pMD602 plasmid by Enrique Arevalo at Dr. Mark Dumont's laboratory, Rochester (personal communication). The correct insertion was confirmed by restriction digestion and DNA sequence analysis at the University of Tennessee.

Table 1

Primers used to amplify coding sequences for single and double transmembrane domains of *STE2*

Primer name	Primer sequence	Amplified region of Ste2p
Fwd(M6)*	5' TGG-(A/AGCTT)-TGG-ATG-AAA-AAGCAGTT CGATAGTTTCCATATTTTACTC 3' 5'TGG-(*Hin*d III)-TGG-Met-Lys-beginning of M6 3'	M6 (residues K-239-269-K)
Rev(M6)*	5' ATATT-(GGATC/C)-TCA-CTT-TTTCAAACTGT ATGCGAGGATGAATATTAT 3' 5' ATATT-(*Bam*H I)-stop codon-Lys-end of M6 3'	M6 (residues K-239-269-K)
Fwd(M7)	5' GGAT-(A/AGCTT)-TGG-ATG-AGTTTGAAACC AAACCAG 3' 5' GGAT-(*Hin*d III)-TGG-Met-beginning of E3 3'	E3-M7-24-T40 (residues 267-339)
Rev(M7)	5' CGAG-(G/GATCC)-CTA-ACTGCTTTTAGCATC GTTG 3' 5' CGAG-(Bam*H I*)-stop codon-end of T40 tail 3'	E3-M7-24-T40 (residues 267-339)

* Fwd(M6) and Rev(M6) primers were designed and used by Enrique Arevalo in constructing pSW02 plasmid, which expresses M6FP.

Table 2

cDNA sequence of the TrpΔLE-M6(K-239-269-K) fusion protein derived from pSW02 plasmid*

5’ ATG CAT CAC CAT CAC CAT CAC CAT CAC CAT AAA GCA ATT TTC GTA CTG AAA GGT TCA CTG GAC AGA GAT CTC GAC AGC CGT ATT GAA CTG GAA CTG CGT ACC GAT CAT AAA GAG CTG TCT GAA CAT CTG CTG CTG GTT GAT CTC GCC CGT AAT GAT CTG GCA CGC ATT GCT ACC CCC GGC AGC CGC TAC GTC GCC GAT CTC ACC AAA GTT GAC CGT TAT TCC TAT GTG CTG CAC CTC GTC TCT CGC GTA GTC GGC GAA CTG CGT CAC GAT CTT GAC GCC CTG CAC GCT TAT CGC GCC GCT CTG AAT CTG GGG ACG TTA AGC GGT GCG CCG AAA GTA CGC GCT (**A/AG CTT**) TGG ATG AAA AAG CAG TTC GAT AGT TTC CAT ATT TTA CTC ATA GCG TCA TCT CAA TCT TTG TTG GTT CCA TCG ATA ATA TTC ATC CTC GCA TAC AGT TTG AAA AAG TGA **(G/GA TCC)** 3'

* The cDNA sequence corresponds to [His-Tag - TrpΔLE - **(*Hin*d III)** - TGG - ATG - Lys-M6-Lys - stop codon-**(*Bam*H I)**].

Table 3

cDNA sequence of the TrpΔLE-E3-M7-24-T40 (267-339) fusion protein derived from pREJ02 plasmid*

5’ ATG CAT CAC CAT CAC CAT CAC CAT CAC CAT AAA GCA ATT TTC GTA CTG AAA GGT TCA CTG GAC AGA GAT CTC GAC AGC CGT ATT GAA CTG GAA CTG CGT ACC GAT CAT AAA GAG CTG TCT GAA CAT CTG CTG CTG GTT GAT CTC GCC CGT AAT GAT CTG GCA CGC ATT GCT ACC CCC GGC AGC CGC TAC GTC GCC GAT CTC ACC AAA GTT GAC CGT TAT TCC TAT GTG CTG CAC CTC GTC TCT CGC GTA GTC GGC GAA CTG CGT CAC GAT CTT GAC GCC CTG CAC GCT TAT CGC GCC GCT CTG AAT CTG GGG ACG TTA AGC GGT GCG CCG AAA GTA CGC GCT (**A/AG CTT) TGG ATG** AGT TTG AAA CCA AAC CAG GGT ACC GAT GTC TTG ACT ACT GTT GCA ACA TTA CTT GCT GTA TTG TCT TTA CCA TTA TCA TCA ATG TGG GCC ACG GCT GCT AAT AAT GCA TCC AAA ACA AAC ACA ATT ACT TCA GAC TTT ACA ACA TCC ACA GAT AGG TTT TAT CCA GGC ACG CTG TCT AGC TTT CAA ACT GAT AGT ATC AAC AAC GAT GCT AAA AGC AGT **TAG (GGA TC/C)** 3’

***** The cDNA sequence corresponds to [His-Tag - TrpΔLE - (***Hin***d III) -TGG-ATG - 3rd Extracellular Loop (E3) - 7th Transmembrane Domain (M7) - 40 Residues of the Tail (T40)-stop codon-(***Bam***H I)**].

The pKLS01 plasmid was constructed by PCR amplifying the E3-M7-24-T40 region of the *STE2* gene from pGA314.Cys-less Ste2p.FTHT plasmid (Akal-Strader et al., 2002) and inserting it into the vector pET-32b(+) (Novagen, Madison, WI) between the *Nco* I and *Xho* I restriction sites. Cloning into pET-32b(+) vector was performed following the manufacturer’s procedures (Novagen). The resulting clone contained a Trx•Tag, His•Tag and S•Tag followed by the E3-M7-24-T40 DNA sequence of Ste2p. The correct insertion was confirmed by DNA sequence analysis at

the University of Tennessee. The pET-32b(+) vector was designed for cloning and high-level expression of peptide sequences fused with the 109aa Trx•Tag™ thioredoxin protein (LaVallie et al., 1993).

Mutations in the plasmids pREJ02M, pKLS01, pKLS01-1, and pKLS01-2, containing small sequence variations in the DNA sequence of E3-M7-24-T40 of Ste2p with mutations Met294Leu, Met294Ala, Met294Leu, Met294Leu, respectively, were achieved through site-directed mutagenesis by Dr. Becker's research group at the University of Tennessee (Kunkel, 1985). A list of the plasmids that were used is summarized in Table 4. The amino acid sequence of the expressed proteins and their molecular weights are listed in Table 5.

Table 4

Summary of the names of the various constructs and the expressed proteins

Plasmid	Expressed protein/peptide
pMD194	Ste2p (wild type)
pMD602	Ste2p (C252S, M250A)
pMMHa	BPTI
pSW02	M6FP (M250A, C252S)
pSW02M	M6FP (M250A, C252S, P258L)
pREJ02	M7FP
pREJ02M	M7FP (M294L)
pKLS01	Trx•Tag™ thioredoxin-E3-M7-24-T40 (M294A, 6xHis)
pKLS01-1	Trx•Tag™ thioredoxin- E3-M7-24-T40 (M294L, 6xHis)
pKLS01-2	Trx•Tag™ thioredoxin-E3-M7-24-T40 (M294L)

Table 5

Summary of the names and the amino acid sequence of the expressed proteins of the Ste2p target regions and their molecular weights

Name of fusion protein	Amino acid sequence[a]	Molecular weight (Da)[b]	Molecular weight (Da)[c]
M6FP (M250A, C252S)	TrpΔLE-WMKKQFDSFHILLI**AS**S QSLLVPSIIFILAYSL KK	17,313	3,750
M6FP(P258L) (M250A, C252S, P258L)	TrpΔLE-WMKKQFDSFHILLI**AS**S QSLLV**L**SIIFILAYSLKK	17,329	3,766
M7FP (M294L)	TrpΔLE-WMSLKPNQGTDVLTT VATLLAVLSLPLSS**L**WATAAN NASKTNTITSDFTTSTDRFYPGT LSSFQTDSINNDAKSS	21,234	7,671
Trx•Tag™ thioredoxin-E3-M7-24-T40 (M294A, 6xHis)	Trx•Tag™ thioredoxin-SLKPNQG TDVLTTVATLLAVLSLPLSS **A**WATAANNASKTNTITSDFTTS TDRFYPGTLSSFQTDSINNDAKS SLEHHHHHH	26,022	_____
Trx•Tag™ thioredoxin- E3-M7-24-T40 (M294L, 6xHis)	Trx•Tag™ thioredoxin-SLKPNQG TDVLTTVATLLAVLSLPLSS**L**W ATAANNASKTNTITSDFTTSTD RFYPGTLSSFQTDSINNDAKSSL EHHHHHH	26,064	_____
Trx•Tag™ thioredoxin-E3-M7-24-T40 (M294L)	Trx•Tag™ thioredoxin-SLKPNQG TDVLTTVATLLAVLSLPLSS**L**W ATAANNASKTNTITSDFTTSTD RFYPGTLSSFQTDSINNDAKSS	24,999	_____

[a] The K residues and the LEHHHHHH sequence indicated by the dashed underlines are not in the native Ste2p sequence. Bolded residues were M, C or P in the native sequence.

[b] Molecular weights corresponding to the fusion proteins containing regions of Ste2p.

[c] Molecular weights corresponding to the peptide regions of Ste2p (underlined in the amino acid sequence) after their release from the TrpΔLE via CNBr cleavage.

* See Arevalo et al., 2003.

Strains

The A232 yeast strain (MAT**a** *ste2-Δcryl*^{R} *ade2-1 his4-580 lys2*$_{oc}$ *tyrl*$_{oc}$ *SUP4-3*^{ts} *leu2 ura3 bar1-1 FUS1::p[FUS1-lacZ TRP1]*) transformed with the plasmid DNA harboring the wild-type and mutant receptors was obtained as a courtesy from Mark Dumont (Martin et al., 1999). *E. coli* expression strains BL21(DE3)pLysS and BL21*trxB*(DE3) were purchased as competent cells from Promega (Madison, WI) and Novagen (Madison, WI), respectively. *E. coli* DH5α cells used to amplify the engineered plasmids were purchased from Gibco BRL Life Technologies (Grand Island, NY).

Chemicals

Polymerase chain reaction reagents, dNTPs, Taq polymerase, restriction enzymes, and ligase were obtained from Promega. Agarose and low melting point agarose were purchased from GIBCO BRL. Isopropyl-β-D- thiogalactopyranoside (IPTG) was purchased from Gold BioTechnology, Inc. (St. Louis, MO). Antibiotics such as ampicillin and chloramphenicol were purchased from Sigma (St. Louis, MO). Guanidinium hydrochloride (GuHCl), cyanogen bromide (CNBr), sodium deoxycholic acid, igepal CA60, and sodium dodecyl sulfate (SDS), and most of the buffer chemicals (Trizma-HCl, sodium phosphate, sodium chloride, potassium chloride, and magnesium chloride) were purchased from Sigma. Reagents used in lysis solution such as Tris-HCl, ethylenediaminetetraacetic (EDTA), phenylmethylsulfonyl fluoride (PMSF), and lysozyme were purchased from Sigma. Nitrocellulose membrane (0.45 μm pore size), and the SDS-PAGE reagents

(Coomassie brilliant blue G-250, acrylamide, N,N'-methylene-bis-acrylamide, glycerol, bromophenol blue, ammonium persulfate, N,N,N',N'-tetramethylethylenediamine (TEMED), SDS, and β-mercaptoethanol were purchased from Sigma. Antibody probe Luminol reagents for the development of His tags, bovine serum albumin (BSA), and HisProbe^{TM}-HRP were obtained from Pierce Biotechnology (Rockford, IL).

Media, buffers, and solvents

Luria Broth media (LB) was prepared by resuspending 25 g of Miller's LB broth (Sigma) in 1 L of water. Plates of LB medium were prepared as for the liquid medium with the addition of 15 g of agar. SOC medium was prepared from 20 g of tryptone, 5 g of yeast extract, 0.5 g of sodium chloride, 186 mg of potassium chloride, and 20 mM glucose in 1 L of H_2O. M9 minimal medium was made by mixing 200 mL of 5X M9 salts (15 g of potassium phosphate (monobasic), 34 g of sodium phosphate (dibasic), 2.5 g of sodium chloride, 5 g of ammonium chloride in 1 L of water, pH 7.2), 8 mL of 50% glucose (filtered-sterilized), 1 mL of 2 M magnesium sulfate, 0.2 mL of 0.5 M calcium chloride, and 788.3 mL of deionized water (Marley et al., 2001). The final glucose concentration was 0.4%. The [^{15}N]-labeled minimal medium was made by substituting ammonium chloride with [^{15}N]-labeled ammonium chloride (Spectra Stable Isotopes, Columbia, MD) in M9 medium. The [^{15}N]- and [^{13}C]-labeled M9 minimal medium was made by substituting ammonium chloride with [^{15}N]-labeled ammonium chloride and glucose with 4 mL of 50% solution of

uniformly labeled [^{13}C] D-glucose (Cambridge Isotope Laboratories, Andover, MA). The final [^{13}C] D-glucose concentration was 0.2%.

All media and their components were sterilized by autoclaving except the glucose solution, which was filtered sterilized and added to the cooled media. After autoclaving of the media, filter-sterilized antibiotics were added to the cool media to a final concentration of 0.2 mg/mL ampicillin and 0.035 mg/mL chloramphenicol.

Lysis solution contained lysis buffer (50 mM Tris-HCl, 1 mM EDTA), 1 mM phenylmethylsulfonyl fluoride, and 300 μg/mL lysozyme, pH 8.7. Inclusion body washing buffer contained 1% igepal Ca-630 and 1% deoxycholic acid in lysis buffer. Tris-buffered saline (TBS buffer) contained 20 mM Tris-HCl and 500 mM NaCl, pH 7.5. Tris-buffered saline + Tween-20 (TBST) was made of TBS + 0.05% Tween-20 (w/v). Blocking buffer contained TBST + 3% bovine serum albumin (w/v). Dilution buffer was made of TBST + 1% bovine serum albumin (w/v). PBS contained 10 mM sodium phosphate, 138 mM NaCl, and 2.7 mM KCl, pH 7.4. TE buffer was made of 8.3 mM Tris-HCl and 40 μM EDTA, pH 7.5. Guanidinium buffer was made of 6 M guanidinium hydrochloride (GuHCl), 10 mM Na_2HPO_4, 138 mM NaCl, and 2.7 mM KCl, pH 7.4. 1X Phosphate buffered saline (PBS) contained 137 mM NaCl, 2.7 mM KCl, 10.1 mM of Na_2HPO_4, 1.8 mM of KH_2PO_4, pH 7.4. 4X SDS sample buffer was prepared from 4 mL of a 20% SDS solution, 2.4 mL of glycerol, 400 μL of 2-mercaptoethanol, 1 mL of Tris-HCl, pH 6.8, and a small amount of BrilliantBlue G in 20 mL of H_2O.

Solvents used for HPLC purification such as acetonitrile and H_2O were HPLC grade and were purchased from VWR International (Bridgeport, NJ). TFA and TFE

were purchased from Sigma. Phospholipids used in circular dichroism (CD) such as 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and 1,2-dimyristoyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)] (DMPG) and detergents such as 1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)] (PPG) and dodecylphosphocholine (DPC) were purchased from Avanti Polar Lipids (Alabaster, AL). Phosphate buffer contained 10 mM Na_2HPO_4 /NaH_2PO_4, 0.02% NaN_3, pH 6.4. Deuterated solvents used in NMR such as d_2-trifluoroethanol (d_2-TFE) and D_2O were purchased from Cambridge Isotope Laboratories and chloroform ($CDCl_3$) and methanol (CD_3OH) were purchased from Sigma. Deuterated detergents used in NMR such as d_{38}-DPC or d_{22}- 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (d_{22}-DHPC) and d_{25}-SDS were purchased from Avanti Polar Lipids and Cambridge Isotope Laboratories, respectively. Phospholipids that were used to prepare lipid vesicles (DMPC and DMPG) were purchased from Avanti Polar Lipids.

Protein expression

Protein expression using pSW02, pSW02M, and pREJ02M plasmids was based on previously published expression procedures (Arevalo et al., 2003; Naider et al., 2004). *E. coli* BL21(DE3)pLysS cells were transformed with the pREJ02M, pSW02, or pSW02M plasmid. Two colonies of the transformed cells were inoculated into 5mL of LB medium containing ampicillin (200 µg/mL) and chloramphenicol (35 µg/mL) and allowed to grow at 37^{o}C and 225 rpm to late log phase (OD_{600} ~0.7). Cells transformed with pREJ02M took about 5 h to reach late log phase (OD_{600} ~0.7) in LB medium. Cells transformed with pSW02 or pSW02M took about 10-11 h to

reach late log phase (OD_{600} ~0.7). Two milliliters of the grown culture were then harvested at 6000 RPM for 3 min and inoculated into 1 L of the same medium, minimal medium, [^{15}N]-labeled minimal medium, or [$^{13}C/^{15}N$]-labeled minimal medium. The cells were grown to late log phase (OD_{600} ~0.6) [~9-10 h when transformed with pREJ02M and 12-13 h when transformed with pSW02 or pSW02M], induced with 1 mM IPTG, incubated for 6 h, and harvested by centrifugation. The time course of expression was performed by collecting 1 mL samples every hour of induction in order to optimize the time of expression. One-mL aliquots of the BL21(DE3)pLysS containing pREJ02M, pSW02, or pSW02M were prepared using the inclusion body isolation protocol below. Western blotting was performed following published procedures (Walker and Gaastra, 1988). His-tag detection was carried out following the manufacturer's procedures (Pierce Biotechnology).

Comparison of the expression of pKLS01, pKLS01-1, pKLS01-2, and pREJ02M

Protein expression using pKLS01, pKLS01-1, and pKLS01-2 plasmids was based on the manufacturer's procedures (Novagen). Two colonies of *E. coli* BL21(DE3)pLysS transformed with pREJ02M and *E. coli* BL21*trxB*(DE3) transformed with pKLS01, pKLS01-1, or pKLS01-2 were inoculated into 5 mL LB medium with the appropriate antibiotics (ampicillin for both cell strains, kanamycin for the BL21*trxB*(DE3) strain, and chlorampenicol for the BL21(DE3)pLysS strain). The cells were then grown with shaking at 225 rpm and 37 °C to late log phase, OD_{600} of 0.6-1, and stored at 4 °C overnight. Two mL of each culture were harvested

by centrifugation at 5,000 rpm for 2 min and 14,000 rpm for 30 sec, respectively. Cell pellets were resuspended into 50 mL cultures and incubated to late log phase, OD_{600} of 0.6-1, for 3-6 h. Cells were induced with IPTG (1 mM) for 3-6 h. One-mL aliquots were collected every hour for SDS-PAGE analysis of induction of protein expression. A 1 mL aliquot of the BL21(DE3)pLysS containing pREJ02M and the BL21*trxB*(DE3) containing pKLS01, pKLS01-1, or pKLS01-2 was prepared using the inclusion body isolation protocol and the total cell protein fraction protocol, respectively.

Inclusion body isolation and protein purification

Large culture

The isolation of inclusion bodies was performed using a modification of a previously published procedure (Staley and Kim, 1994). All of the following steps were carried out at 4 °C. After growing and inducing the expression cells with IPTG in either rich or M9 mimimal medium, for expressing unlabeled or isotopically-labeled polypeptides, respectively, we harvested the cells by centrifugation at 8,000 rpm for 20 min. The wet cell pellet was weighed and then resuspended in lysis solution at a ratio of 5 mL of solution per gram of wet cells. The cell suspension was sonicated until a uniform suspension was achieved and then the lysate was centrifuged at 18,000 rpm for 20 min. The resulting pellet was sequentially resuspended by sonication in 3 mL of lysis buffer, 3 mL of inclusion body washing buffer, and then 3 mL of distilled water per gram of wet cells. After each resuspension and sonication, the resuspended pellet was centrifuged as specified

above. Finally, the resulting inclusion body pellet was weighed and resuspended by sonication in 3-4 mL of guanidinium buffer. The solution was sonicated and then purified via HPLC, which yielded ~10-100 mg of fusion protein per liter of cell culture.

Small culture

The 1-mL aliquots of the TrpΔLE fusion proteins, which were collected for SDS-PAGE analysis, were centrifuged at 8,000 rpm for 5 min and the pellets were resuspended completely in 100 μL of lysis solution. Cell suspensions were sonicated for 20-30 s and then the lysate solution was centrifuged at 14,000 rpm for 5 min. The resulting pellet was resuspended by sonication in 200 μL of inclusion body washing buffer and then centrifuged as specified above. All of the above steps were perforemed at 4 °C. Finally, the inclusion body pellet was resuspended by sonication in 100 μL of 4X SDS sample buffer. The samples were stored at –20 °C and then boiled for 1-2 min just before running in SDS-PAGE.

Total cell protein fraction

The 1-mL aliquots which were collected for SDS-PAGE analysis of the induction of protein expression as 109aa Trx•Tag™ thioredoxin fusion proteins were centrifuged at 8,000 rpm for 5 min. The pellets were resuspended completely in 100 μL of 1 X PBS to yield a concentration factor of 10X. Then 100 μL of 4X SDS sample buffer was added. The samples were sonicated and then boiled for 1-2 min and stored at –20 °C.

Protein cleavage

Cyanogen bromide (CNBr) was used to release the transmembrane peptide from the N-terminal TrpΔLE leader protein (Walker and Gaastra, 1988; Arevalo et al., 2003). The lyophilized pure fusion protein was cleaved with 1 M CNBr. Approximately a 500- or 5000-fold molar excess of CNBr was added to the M7FP and M6FP or M6FP(P258L) solutions, respectively. The lyophilized pure M6FP and M6FP(P258L) fusion proteins and the CNBr were dissolved in 70% TFA, while the M7FP was dissolved in 50% TFA. The CNBr was dissolved separately in TFA and was then added to the fusion protein. The cleavage reaction of M7FP and wild-type or mutant M6FP was carried out in the dark, at room temperature for 4.5 h and 24 h, respectively. The reaction was stopped by lyophilization. The lyophilized M6 peptides were dissolved in 10% TFA, 60% acetonitrile, and 30% H_2O, wheras the E3-M7-24-T40 peptide was dissolved in 10% TFA, 40% acetonitrile, and 50% H_2O, and then 3-4 mg of peptide was injected on a preparative C18 reverse-phase column and purified.

Peptide purification

HPLC purification was carried out using a 1050 Hewlett Packard-Agilent instrument equipped with a gradient-based solvent system. Chromatograms were monitored at 220 nm. Fusion protein purification was carried out using a Vydac 259VHP82215 preparative reversed-phase polymer column (22 mm x 150 mm; 8 μm; 300 Å) with a water jacket at 50 °C and a water-acetonitrile (0.1% TFA) gradient

from 30% to 60% acetonitrile in 80 min at a flow rate of 4 mL/min. The guanidinium chloride solution of the inclusion bodies containing fusion protein was injected directly on the column. Usually 2-3 mL were injected per run. Cleaved peptides were purified using a Waters μBondpak^{TM} preparative C18 reversed-phase column (19 mm x 300 mm; 10 μm; 125 Å) with a water jacket at 50 °C and a water-acetonitrile (0.1% TFA) gradient from 40% to 80% acetonitrile in 80 min at a flow rate of 4 mL/min. Fusion proteins and cleaved peptides were analyzed using a 1050, 1100, or 1090 Hewlett Packard-Agilent instrument equipped with a gradient-based solvent system and a Vydac 259VHP54 reversed-phase polymer column (4.6 mm x 150 mm; 5 μm; 300 Å) for the fusion proteins and a Waters Delta Pak C18 reversed-phase column (3.9 mm x 150 mm; 5 μm; 100 Å) for the cleaved peptides, at 50 °C and water-acetonitrile (0.1% TFA) gradient from 30% to 60% acetonitrile in 20 min at a flow rate of 1 mL/min. All peptides were purified to over 98% homogeneity as judged by RP-HPLC. The molecular weights of the final products were assessed by ESI-MS at the mass spectrophotometer facility at Hunter College, CUNY, New York, NY.

CD sample preparation

CD samples in TFE/H_2O-

E3-M7-24-T40 (1.92 mg), M6 (0.81 mg), and M6(P258L) (0.79 mg) were dissolved in 200 μL of 95% TFE to obtain stock solutions. Five 20 μL or 25 μL aliquots of the stock solutions of E3-M7-24-T40 or M6 or M6(P258L), respectively, were diluted to 450 μL with the appropriate amount of TFE and H_2O to obtain final peptide concentration of about 50 μM in 95%, 75%, 50%, 25%, and 5% of TFE/H_2O

(v/v) solutions. The final concentration of the peptides in 5-95% of TFE/H_2O was determined by UV spectroscopy at 280 nm using an extinction coefficient of 6890 $M^{-1}cm^{-1}$ for E3-M7-24-T40 (one tyrosine and one tryptophan residue) and 1340 $M^{-1}cm^{-1}$ for M6 and M6(P258L) [one tyrosine residue] (Naider et al., 2003).

CD samples in DMPC/DMPG (4:1) vesicles-

E3-M7-24-T40 (0.97 mg), M6 (0.38 mg) and M6(P258L) (0.44 mg) were dissolved in 500 μL of TFE/H_2O (1:1) to obtain stock solutions. The concentrations of the stock solutions of E3-M7-24-T40, M6, and M6(P258L) were determined by UV spectroscopy to be 165 μM, 146 μM and, 111 μM, respectively. To obtain a 50 μM final concentration of E3-M7-24-T40, M6, and M6(P258L) in phosphate buffer containing vesicles, first, 151 μL, 171 μL, and 225 μL of the stock solutions, respectively, were added to 2 mg of DMPC/DMPG (4:1) in 1 mL of $CHCl_3$. The resulting mixture was dried under N_2 flow. Residual traces of organic solvent were removed by placing the dried film under vacuum overnight. Then the peptide/detergent mixture was resuspended in 500 μL of 10 mM phosphate buffer, pH 6.4, to give a final peptide concentration of 50 μM. The suspension was sonicated at 50 °C for 60 min using a W-385 unit (Misonix, Farmingdale, NY) equipped with a 2.5 inch cup horn sonicator operated at 40% output power (~ 200 watts).

CD samples in DPC, PPG and SDS-

E3-M7-24-T40 (2.82 mg), M6 (2.03 mg), and M6(P258L) (1.96 mg) were dissolved in 2.5 mL of TFE/H_2O (1:1) to obtain stock solutions. The concentrations

of the stock solutions of E3-M7-24-T40, M6, and M6(P258L) were determined by UV spectroscopy to be 74, 118, and 112 μM, respectively. The E3-M7-24-T40 stock solution was aliquoted into 2 vials and lyophilized overnight. Then 200 μL of 100 mM DPC or PPG in 10 mM phosphate buffer were added to E3-M7-24-T40 to obtain a peptide concentration of 390 μM in DPC and 500 μM in PPG. The E3-M7-24-T40 sample having a lower detergent to peptide ratio was prepared in a similar manner but with a detergent concentration of 10 mM. Because of the poor solubility of the M6 peptides in aqueous detergent solutions, 200 μL of 10-15 mM DPC or PPG dissolved in TFE/H_2O (1:1) were added to aliquots of the M6 (85 μL) and M6(P258L) [100 μL] peptides. The peptide-detergent mixture was co-lyophilized and then 200 μL of 10 mM phosphate buffer was added to obtain a peptide concentration of about 50 μM. The concentrations of M6 and M6(P258L) peptides in PPG, which were determined by UV to be 124 μM and 112 μM, respectively, were much larger than that determined from the calculated value (50 μM). Because of the inaccuracy in the UV concentrations of these two peptides in PPG, calculated concentrations were used in CD measurements. To prepare a sample of M6 peptide in SDS micelles, M6 (0.35 mg) was dissolved in 600 μL of 5% TFA, 35% TFE, and 60% H_2O. SDS (34.6 mg) was dissolved in 1750 μL of 0.02% NaN_3 in water, the pH was adjusted to 4.0 using TFA, to give a final detergent concentration of 69 mM. Then 450 μL of the above SDS preparation was added to the dissolved M6 peptide and the peptide-detergent mixture was colyophilized. To the colyophilized peptide-detergent mixture, 450 μL of H_2O, pH 4.0, was added to obtain a target peptide concentration of about 200 μM. The concentration of M6 peptide in SDS was determined by UV spectroscopy to be

94 μM. All peptide-detergent mixtures were sonicated at 50 °C for 15 min prior to the CD measurements.

CD measurements

The CD spectra of the peptides were recorded on an AVIV model 62-DS CD instrument (AVIV Associates, Lakewood, NJ). Quartz cuvettes with pathlengths of 1 mm and 0.2 mm were used for peptides in aqueous TFE solutions and DMPC/DMPG (4:1) vesicles, respectively. The concentration of the peptides in aqueous-TFE solutions and DMPC/DMPG (4:1) vesicles was about 50 μM. Cuvettes with pathlengths of 0.2 and 0.1 mm were used for peptides with concentration of 50 μM and 400 μM, respectively, in detergents (15 mM or 100 mM). All spectra were the average of 3-5 scans between 260-280 nm and 185 nm at an interval of 1 nm with a 3-5 sec integration time at each wavelength. The bandwidth for each measurement was set to 2 nm. CD spectra on blanks corresponding to the different media without dissolved protein were collected and subtracted from the spectra containing the protein. Protein concentrations were obtained by UV spectroscopy. CD intensities are expressed as mean residue ellipticities (deg cm^2 $dmol^{-1}$). Mean residue ellipticity values were calculated according to the formula:

$$[\theta] = \Delta\varepsilon/(10 \times C \times l \times AA) \qquad \text{(Eq. 1)}$$

where $\Delta\varepsilon$ is the measured ellipticity, C is the concentration of the peptide in moles/L, l is the path length in centimeters and AA is the number of amino acids of the peptide. Assuming that the transmembrane and tail portions of E3-M7-24-T40 behaved

independently, we may estimate the mean residue ellipticity of E3-M7-24 using the following equation,

$$[\theta]_{E3\text{-}M7\text{-}24} = \{([\theta]_{E3\text{-}M7\text{-}24\text{-}T40} \times 73) - ([\theta]_{T40} \times 40)\}/33 \quad \text{(Eq. 2)}$$

Mathematical subtraction of the CD spectra of T40 from the E3-M7-24-T40 peptide spectra was carried out using the calculated molar ellipticity values of each spectrum. Finally, the mean residue ellipticity ([θ]) of E3-M7-24 was obtained by dividing the molar ellipticity by the appropriate number of residues. The percentage of α-helicity was calculated using the method of Wu et al. (1981) and Chen et al. (1974).

NMR sample preparation

NMR samples in TFE/H_2O or $CDCl_3$/CD_3OH/H_2O-

[^{15}N]-E3-M7-24-T40 (1.9 mg) was dissolved in 500 μL TFE/H_2O (1:1) to yield a peptide concentration of 488 μM. The peptide was first dissolved in d_2-TFE and then H_2O (0.1% TFA) was added to give the final ratio of TFE and H_2O (v/v). Similarly, [^{15}N]-E3-M7-24-T40 (1.9 mg) was dissolved in 500 μL $CDCl_3$/CD_3OH/H_2O (4:4:1) to yield a peptide concentration of 488 μM. The peptide was first suspended in $CDCl_3$, and then CD_3OH and finally H_2O (0.1% TFA) were added to give the final ratio of the solvent components. The solution clarified upon addition of the aqueous component. Acidic water was used to aid in water suppression. The peptide concentrations were based on the measured weights of the samples.

NMR samples in detergents: At low peptide concentrations (0.17-0.2 mM) in SDS, DPC, DHPC and PPG at high detergent/peptide ratios (~450/1)-

[^{15}N]-E3-M7-24-T40 (1.6 mg) was dissolved in TFE/H_2O (1:1) then divided into 3 vials and lyophilized. To each vial (containing 0.53 mg of [^{15}N]-E3-M7-24-T40) was added 350 μL of d_{25}-SDS, d_{38}-DPC, or d_{22}-DHPC (86 mM) detergents in H_2O/D_2O (9:1), containing 0.02% NaN_3, the pH was adjusted to 4.0 using TFA, to yield a final peptide concentration of 194 μM and a peptide/detergent ratio of 1/440. In a separate experiment 350 μL of PPG (86 mM) in H_2O/D_2O (9:1), containing 0.02% NaN_3, pH 4 was added to [^{15}N]-E3-M7-24-T40 (0.46 mg) to yield a peptide concentration of 170 μM and a peptide/detergent ratio of 1/500. The peptide/detergent mixtures were sonicated at 50 °C for 10 min and then transferred to an NMR Shigemi tube.

NMR sample in detergents: At high peptide concentration (0.5 mM) in PPG at high peptide/detergent ratio (1/17)-

300 μL of 8.5 mM PPG (1.3 mg) in H_2O/D_2O (9:1), containing 0.02% NaN_3 was added to [^{15}N]-E3-M7-24-T40 (1.14 mg) to yield a peptide concentration of 488 μM and a peptide/detergent ratio of 1/17. The peptide/detergent mixture was sonicated at 50 °C for 10 min and then transferred to an NMR Shigemi tube.

NMR sample in detergents: At high peptide concentration (0.5 mM) in DPC at high detergent/peptide ratios (~450/1)-

To 1.14 mg of [^{15}N]-E3-M7-24-T40 was added 350 μL of 200 mM d_{38}-DPC (27.28 mg) in H_2O/D_2O (9:1), containing 0.02% NaN_3, the pH was adjusted to 4.0 using TFA, to yield a peptide concentration of 475 μM and a peptide/detergent ratio of 1/421. The peptide/detergent mixture was sonicated at 50 °C for 10 min and then transferred to an NMR Shigemi tube.

NMR spectroscopy

^{1}H-NMR spectra were recorded on a three-channel Varian UNITY INOVA 600 MHz NMR spectrometer (Varian NMR Instrument, Palo Alto, CA) equipped with a z-axis pulsed-field-gradient and a Varian 5-mm ^{1}H/^{13}C/^{15}N triple resonance probe. For all 1D proton spectra and 2D homonuclear proton experiments of [^{15}N]-labeled E3-M7-24-T40 peptide ^{15}N-decoupling was applied. In the case of the [^{15}N]-E3-M7-24-T40 peptide in TFE-d_2/H_2O (1:1 v/v), $CDCl_3/CD_3OH/H_2O$ (4:4:1 v/v), d_{25}-SDS, d_{38}-DPC, d_{22}-DHPC and PPG, heteronuclear single quantum correlation (HSQC) was used (Bodenhausen et al., 1980). HSQC spectra were obtained at 25 °C in organic-aqueous media and at 50 °C in detergent media. HSQC-nuclear Overhauser effect spectroscopy (NOESY) and HSQC-total correlation spectroscopy (TOCSY) experiments (Kay et al., 1992; Zhang et al., 1994) were used for the [^{15}N]-E3-M7-24-T40 peptide in TFE-d_2/H_2O (1:1 v/v) and $CDCl_3/CD_3OH/H_2O$ (4:4:1 v/v) at 25 °C. Identification of specific amino acid signals was made using 3 dimensional TOCSY-HSQC spectra with contact times of 70 ms to identify spin systems. Sequential assignments of amino acids were determined using 3 dimensional NOESY-HSQC spectra. Two dimensional TOCSY (Braunschweiler et al., 1983) and

NOESY (Jeener et al., 1979) spectra were also recorded and were helpful in identifying crosspeaks or connections belonging to proline residues in 3D NOESY. All NOESY experiments were performed with a 300 ms mixing time and the results were used in modeling calculations.

NMR structure refinement

Molecular modeling was performed using the DYANA program (Güntert et al., 1997) and all calculations were performed on a Linux workstation. All necessary file conversions from the Varian format and further data processing to the NMRView format were performed using the NMRPipe software (Delaglio et al., 1995). NOESY cross-peak assignments, integration of crosspeak volumes, and creation of distance restriction files in the DYANA input format were aided by the NMRView program (Johnson et al., 1994). The upper NOE distance constraints were calculated by the NMRView program using the median peak intensity calibration method, all lower distance constraints were set to 1.8 Å. Dihedral angle constraints were generated for Φ torsion angles using NH-CHα coupling constants ($^3J_{HN\alpha}$) and equation 3, which was first described by Karplus (1959) and then modified by Bystrov (1976) and (Pardi et al., 1984).

$$^3J_{HN\alpha} = 6.4 \cos^2 \theta - 1.4 \cos \theta + 1.9 \qquad \text{(Eq. 3)}$$

where $\theta = |\Phi - 60°|$, and $^3J_{HN\alpha}$ is given in hertz. The DYANA modeling starts by generating 50 to 100 molecules with random conformations. We used the standard DYANA protocol that includes simulated annealing for every initial random conformation molecule with 4000 torsion angle dynamics steps. The first 800 of these

steps were performed at an initial high temperature that is software determined, followed by slow cooling of the molecule. One half of the refined structures representing those molecules in the ensemble with the lowest target function (lowest sum of NMR distance constraint violations) was selected for generating the output protein data base files and for further analysis, including RSMD calculation and visualization by the MOLMOL program (Koradi et al., 1996).

Chapter III: Results*

III. 1. Cloning, biosynthesis, isolation, and purification of fusion proteins corresponding to domains of Ste2p

In order to obtain high-resolution structural information on single and multiple domains of Ste2p in membrane mimetic media via NMR, milligram quantities of isotopically labeled peptides are needed. Because of the important role that TM6 plays in the activation of Ste2p (Stefan et al., 1998), and for structural comparison of peptides corresponding to the wild type and constitutively active receptor, wild type and mutant (P258L) TM6 peptides, residues 238-270, have been chosen for structural studies. Since TM6 and TM7 of GPCRs interact (Parrish et al., 2002; Dube and Konopka, 1998) and are likely involved in the stimulation of cellular responses (Konopka et al., 1996), studies on the structure of TM7 in addition to TM6 would have significant implications on revealing Ste2p function. Because of the suggested role of the cytoplasmic C-terminus of Ste2p in regulating receptor signaling and desensitization (Dohlman and Thorner 2001; Feng and Davis, 2000; Hicke et al., 1998; Rohrer et al., 1993; Tan et al., 1993; Schandel and Jenness, 1994), and since the carboxyl terminus of Ste2p is highly hydrophilic, 40 residues of the cytoplasmic tail (T40; residues 300-339) that were expected to increase the solubility of the hydrophobic TM7 domain in aqueous solution (Naider et al., 2003) were included for structural studies.

* Significant parts of the Results and Discussion appear in publications referenced on page 130.

In order to biosynthesize and [^{15}N]-label wild type TM6 peptide (residues 238-270) and E3-M7-24-T40 multidomain peptide [consisting of the third extracellular loop (267-275), TM7 (276-299), and 40 amino acid residues of the carboxyl terminus (300-339) of Ste2p attached], pSW02 and pREJ02 plasmids, respectively, were constructed. The pSW02 and pREJ02 plasmids code for fusion proteins consisting of a N-terminus histidine-tagged TrpΔLE leader sequence joined via a methionine residue to TM6, a 33 amino acid residue single domain peptide (238-270) of Ste2p, or E3-M7-24-T40, a 73 amino acid residue multidomain peptide (267-339) of Ste2p, respectively (Figure 3C,D). The TrpΔLE fusion protein directs the production of the fusion protein into inclusion bodies, which protects the cell from potential protein toxicity and facilitates isolation of a fusion protein (Altman et al., 1991; Uhlen et al., 1990). The methionine residue used to fuse the C-terminus of TrpΔLE with the M6 or E3-M7-24-T40 peptides serves as a cleavage site so that the target peptide could be released with cyanogen bromide. This necessitated the mutation of Met250Ala in the M6 region of Ste2p and Met294Leu in the M7 domain. The Met250Ala mutation in the M6 region of Ste2p was achieved in the parent plasmid pMD194 containing the wild-type *STE2* gene by Enrique Arevalo at Dr. Mark Dumont's laboratory, University of Rochester (personal communication) to generate Ste2p mutant plasmid pMD602 (Met250Ala). The Met294Leu mutation in the M7 region of Ste2p was achieved through site-directed mutagenesis by Dr. Becker's research group at the University of Tennessee (Kunkel, 1985). Previous studies have shown that these mutations do not impair the function of Ste2p (Arevalo, 2002; Martin et al., 2002).

Restriction digestion analysis of the constructed pSW02 and pREJ02 plasmids digested with *Bam*H I and *Hin*d III restriction enzymes produced a 114 bp band and a 238 bp band, respectively (Figure 4). The 114 bp band that came from two of the isolated and digested pSW02 plasmids corresponds to the M6 DNA sequence. The 238 bp band that came from two of the isolated and digested pREJ02 plasmids corresponds to the E3-M7-24-T40 DNA sequence. The appearance of the 114 bp band and the 238 bp band indicates that the ligations of the M6 DNA sequence and the E3-M7-24-T40 DNA sequence to the TrpΔLE in the pMMHa vector were successful (Figure 4). The DNA sequence of the plasmids (see Materials and Methods section) was confirmed by sequencing analysis.

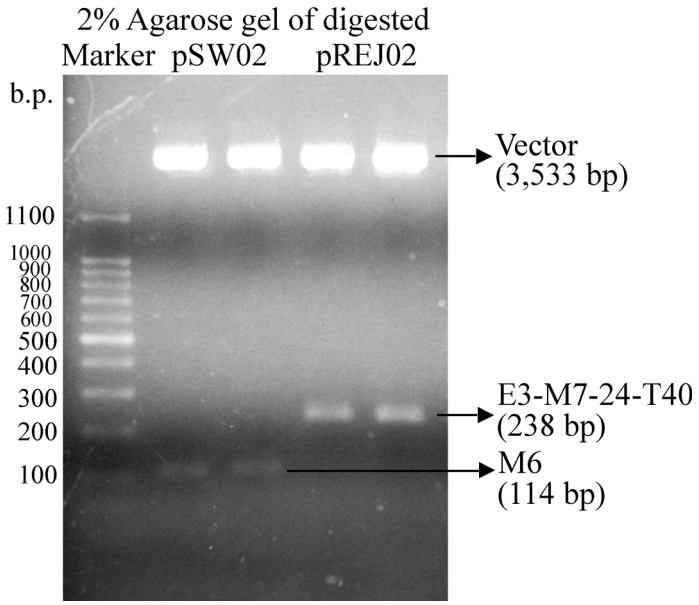

Figure 4. 2% agarose gel of digested pSW02 and pREJ02 plasmids. pSW02 (lanes 2 & 3) and pREJ02 (lanes 4 & 5) plasmids were digested with *Hin*d III and *Bam*H I restriction enzymes and analyzed on a 2% agarose gel. The expected length of the M6 insert DNA sequence is 114 bp (lanes 2 & 3) and the E3-M7-24-T40 insert DNA sequence is 238 bp (lanes 4 & 5).

After the confirmation of the correct constructs the pSW02 and pREJ02 plasmids were sent to the University of Tennessee for site directed mutagenesis

introducing the mutation Pro258Leu to the DNA sequence of M6 creating the pSW02M plasmid and introducing mutation Met294Leu to the DNA sequence of E3-M7-24-T40 creating the pREJ02M plasmid. The DNA sequence of the plasmids containing the above mentioned mutations was reconfirmed by sequencing analysis.

Expression of unlabeled and isotopically labeled fusion proteins was induced with 1 mM of IPTG for up to 6 h in LB rich medium and M9 minimal medium, which contained [^{13}C]-labeled glucose and/or [^{15}N]-labeled ammonium chloride. Inclusion bodies, expressed from the induced cells containing pSW02, pSW02M, or pREJ02M, were isolated and the expression of [^{15}N]-M6FP, [^{15}N]-M6FP(P258L), [^{15}N]-M7FP or [^{13}C/^{15}N]-M7FP in M9 minimal medium and unlabeled-peptides in rich medium were analyzed via SDS-PAGE (Figure 5).

The M7FP fusion protein containing the E3-M7-24-T40 region of Ste2p exhibited excellent expression properties (Figure 5A-C) whereas the M6FP and M6FP(P258L) fusion proteins were expressed at lower levels in both rich and minimal media (Figure 5D,E). The absence of the band corresponding to the molecular weight of the peptide of interest at zero hour induction (0 h) indicates that the target protein is not being expressed or expression is undetectably low before induction. In all cases maximal expression was observed four to six hours after induction. Based on these results it was decided to use 6 h as the standard expression time for the fusion proteins. Western blotting using a hexa-histidine probe confirmed the presence of the M6 fusion protein (Figure 6) and the M7 fusion protein (data not shown).

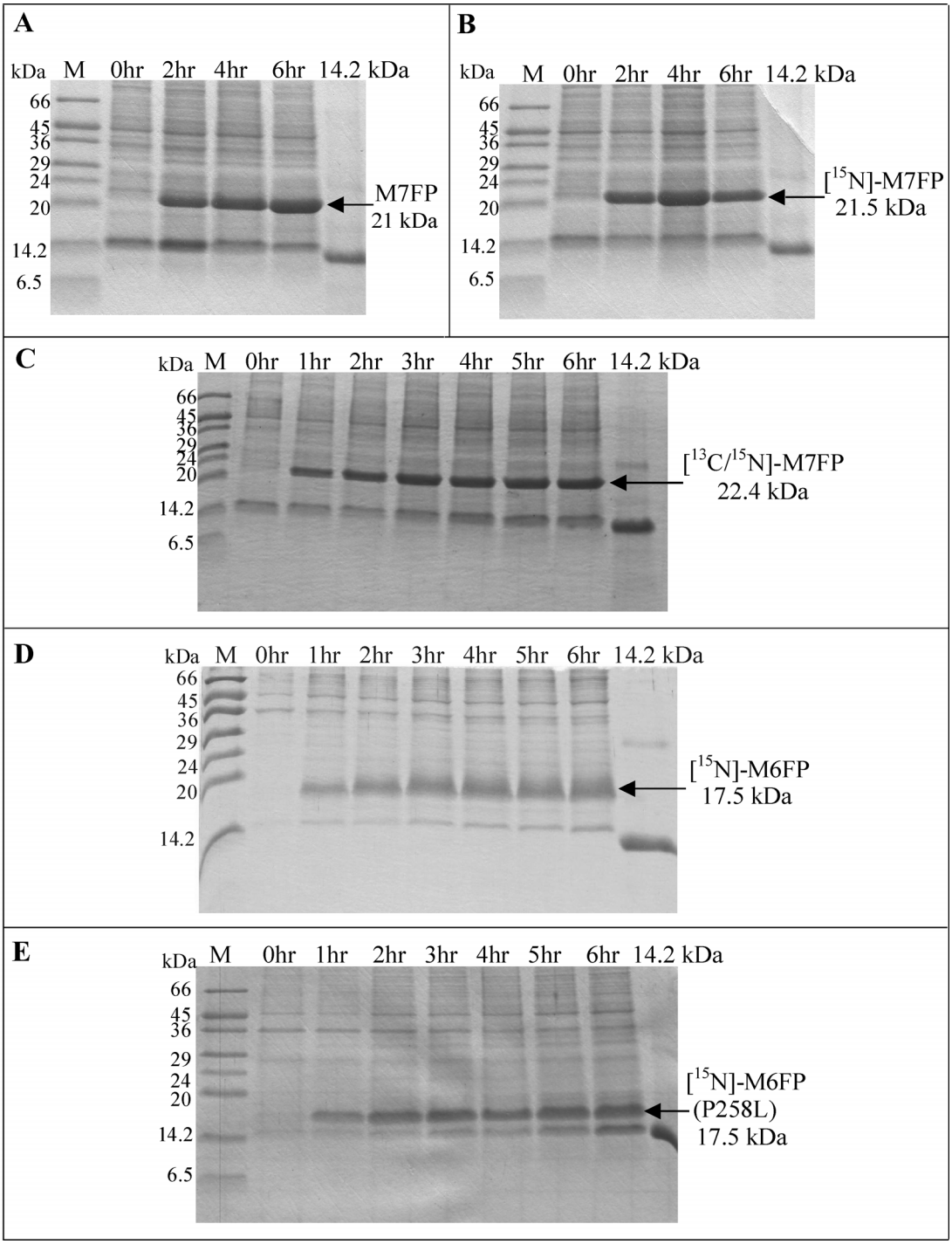

Figure 5. Expression of unlabeled and isotopically labeled TrpΔLE-M6, TrpΔLE-M6(P258L), and TrpΔLE-E3-M7-24-T40 fusion proteins from the pSW02, pSW02M, and pREJ02M plasmids, respectively. The expression of the M6FP and the M7FP was monitored every hour for up to six hours via SDS-PAGE analysis. (A) Unlabeled-M7FP has a molecular weight of 21kDa. (B) [^{15}N]-labeled M7FP has a molecular weight of 21.5 kDa. (C) [^{13}C/^{15}N]-labeled M7FP has a molecular weight of 22.4 kDa. (D) [^{15}N]-labeled M6FP has a molecular weight of 17.5 kDa. (E) [^{15}N]-labeled M6FP(P258L) has a molecular weight of 17.5 kDa.

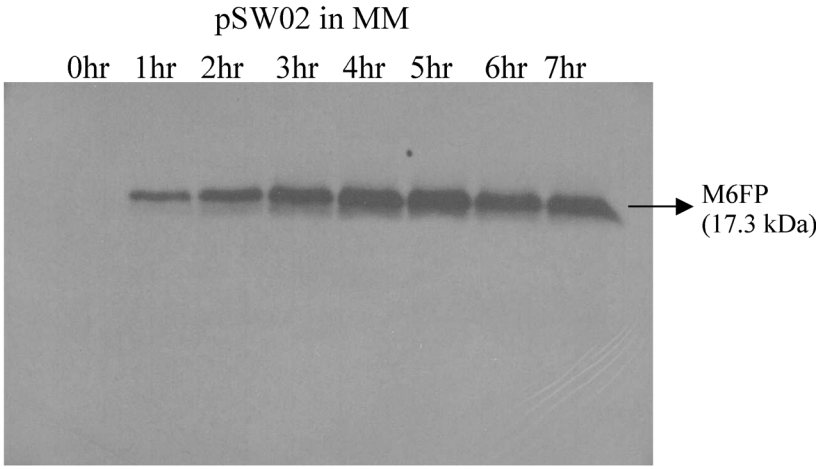

Figure 6. Western blot of inclusion bodies, isolated after induction with 1mM of IPTG in M9 minimal medium (M.M.) for up to 7 h. M6FP (17.313 kDa) was detected using a histidine probe (INDIA HisProbe-HRP).

A thioredoxin fusion system has been used in the biosynthesis of transmembrane domains of the CFTR protein (Therien et al., 2002). To evaluate the efficiency of the thioredoxin (Therien et al., 2002) and TrpΔLE expression systems (Miozzari and Yanofsky, 1978; Kleid et al., 1981) we compared the biosynthesis of M7FP by the pREJ02M and pKLS01 plasmids. pREJ02M encodes the E3-M7-24-T40 region of Ste2p with the Met294Leu mutation, fused with the TrpΔLE peptide (TrpΔLE-E3-M7-24-T40 or M7FP), a 21 kDa peptide, whereas the pKLS01 plasmid encodes the E3-M7-24-T40 region of Ste2p, with a Met294Ala mutation and an additional 8 amino acids at the C-terminus of the peptide fused with the Trx•Tag™ thioredoxin (Trx•Tag™ thioredoxin-E3-M7-24-T40), a 26 kDa peptide (Table 5). Both expression systems were tested in rich and minimal media. Expression of the TrpΔLE-E3-M7-24-T40 fusion protein driven by pREJ02M was significantly higher than the expression of the Trx•Tag™ thioredoxin-E3-M7-24-T40 fusion protein driven by pKLS01 in rich medium (Figure 7A,B). Moreover, the TrpΔLE fusion protein

represented a much higher fraction of the proteins as judged by SDS-PAGE. Strikingly, the expression of M7FP by pREJ02M remained high in minimal medium (Figure 7C) whereas no expression of fusion protein was driven by pKLS01 in minimal medium (Figure 7D). Two other plasmids, pKLS01-1 and pKLS01-2 coding for variations of the E3-M7-24-T40 domain, one virtually identical to that coded for by pREJ02M (Table 5), were tested and gave similar results to that of pKLS01 (data not shown). Based on this observation, in large scale expressions the

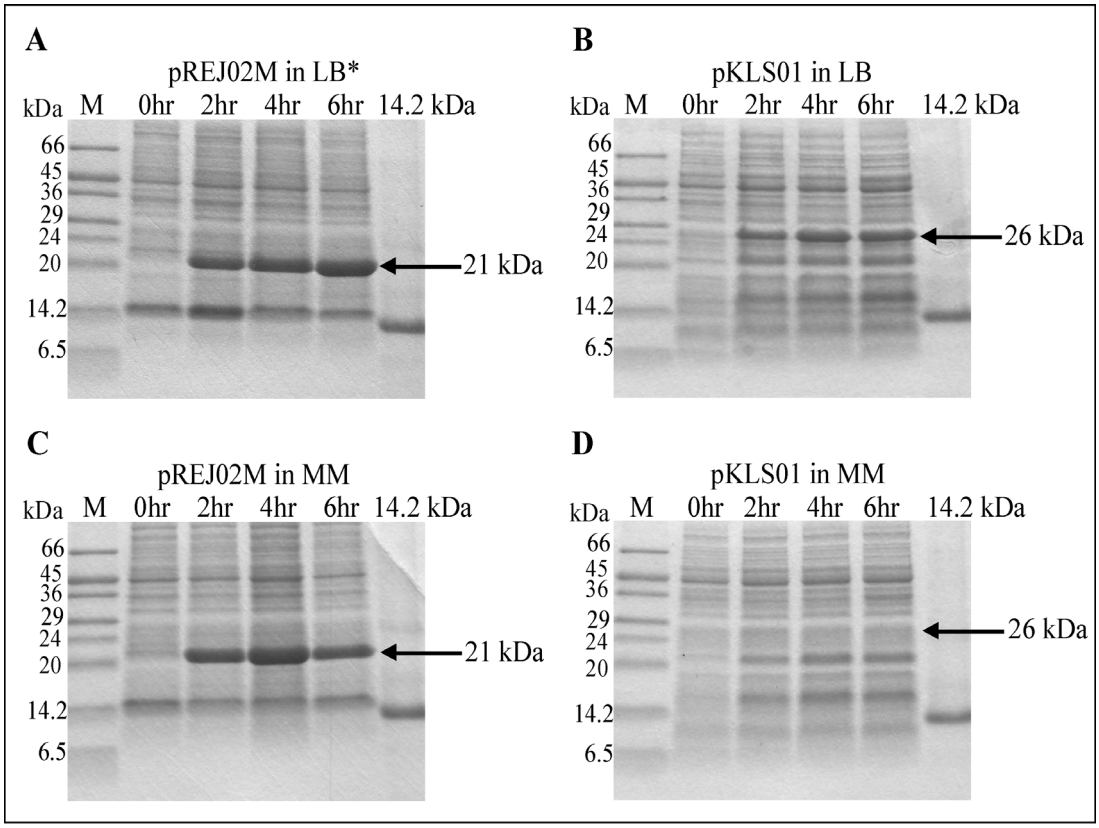

Figure 7. SDS-PAGE of E3-M7-24-T40 fusion proteins biosynthesized in rich (LB) or minimal medium (MM). Cells were induced with 1mM of IPTG in LB or M9 MM for up to 6 h and the fraction containing the target fusion protein was isolated as per the Materials and Methods. TrpΔLE-E3-M7-24-T40(Met294Leu) (A & C) and Trx•Tag™ thioredoxin-E3-M7-24-T40(Met294Ala) (B & D) have molecular weights of 21 and 26 kDa, respectively. SDS-PAGE gels were stained with Coomassie dye.

* Published in Naider et al., 2004.

system that expressed proteins fused to a portion of the TrpΔLE1413 polypeptide (Miozzari et al., 1978; Kleid et al., 1981) was used to prepare the M6 and M7 regions of Ste2p.

After the expression and isolation of the fusion protein inclusion bodies, the final pellet was resuspended by sonication in 6M GuHCl and then purified via HPLC (see Materials and Methods). The crude inclusion bodies were highly enriched in the target fusion protein (Figure 8), which was readily isolated to yield a pure fusion protein of greater than 95% homogeneity as judged by reversed phase HPLC (Figure 9), SDS-PAGE (Figure 10) and mass spectrometry (Figure 11). Biosynthesis, isolation and purification of unlabeled- and [^{15}N]-labeled M6FP and M6FP(P258L) yielded about 8-12 mg (> 95% homogeneous, Figure 9A,B) of fusion protein per liter of bacterial culture in both rich and minimal media. Growth in rich media yielded 100 mg of pure unlabeled-M7FP (> 95% homogeneous) per liter of bacterial culture while growth in isotopically [^{15}N]-labeled minimal media produced about half that amount (50 mg) of [^{15}N]-labeled M7FP (> 95% homogeneous, Figure 9C). However, when the fusion protein was expressed in isotopically [^{13}C/^{15}N]-labeled minimal media the yield of pure M7FP was only 30 mg per liter of growth (> 95% homogeneous, Figure 9D).

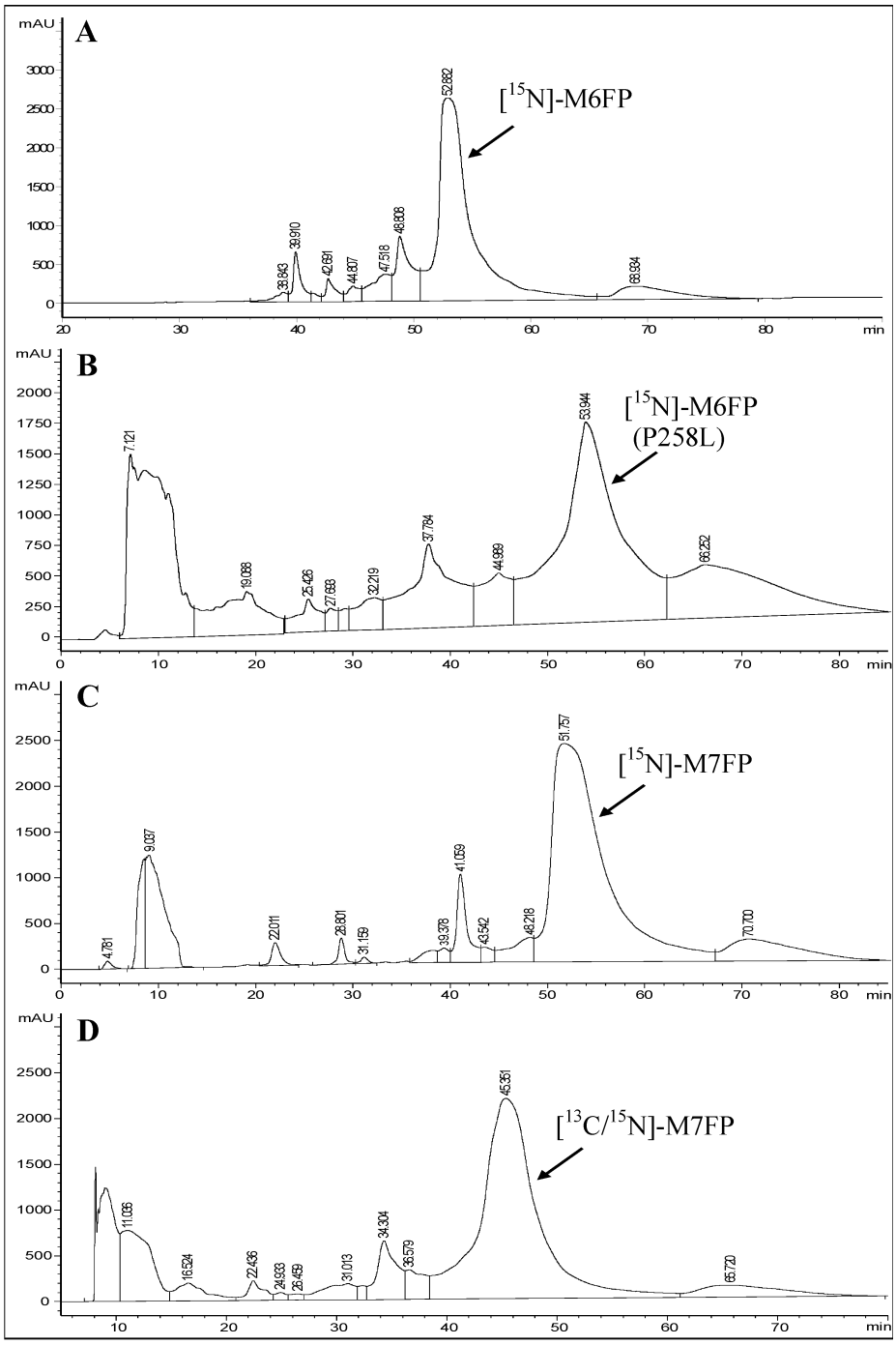

Figure 8. HPLC chromatograms of crude fusion proteins. After induction with 1 mM of IPTG in M9 minimal medium for 6 h, inclusion bodies were isolated. The pellets were resuspended by sonication in 6 M GuHCl, then purified using a preparative reversed-phase polymer column in an acetonitrile-water gradient. The [^{15}N]-M6FP (A), [^{15}N]-M6FP(P258L) (B), [^{15}N]-M7FP (C), and [^{13}C/^{15}N]-M7FP (D) peptides are observed at a retention time of around 53, 54, 52, and 45 min, respectively. The [^{15}N]-M7FP (C) and [^{13}C/^{15}N]-M7FP (D) were observed at different retention times because they were purified on different HPLC columns.

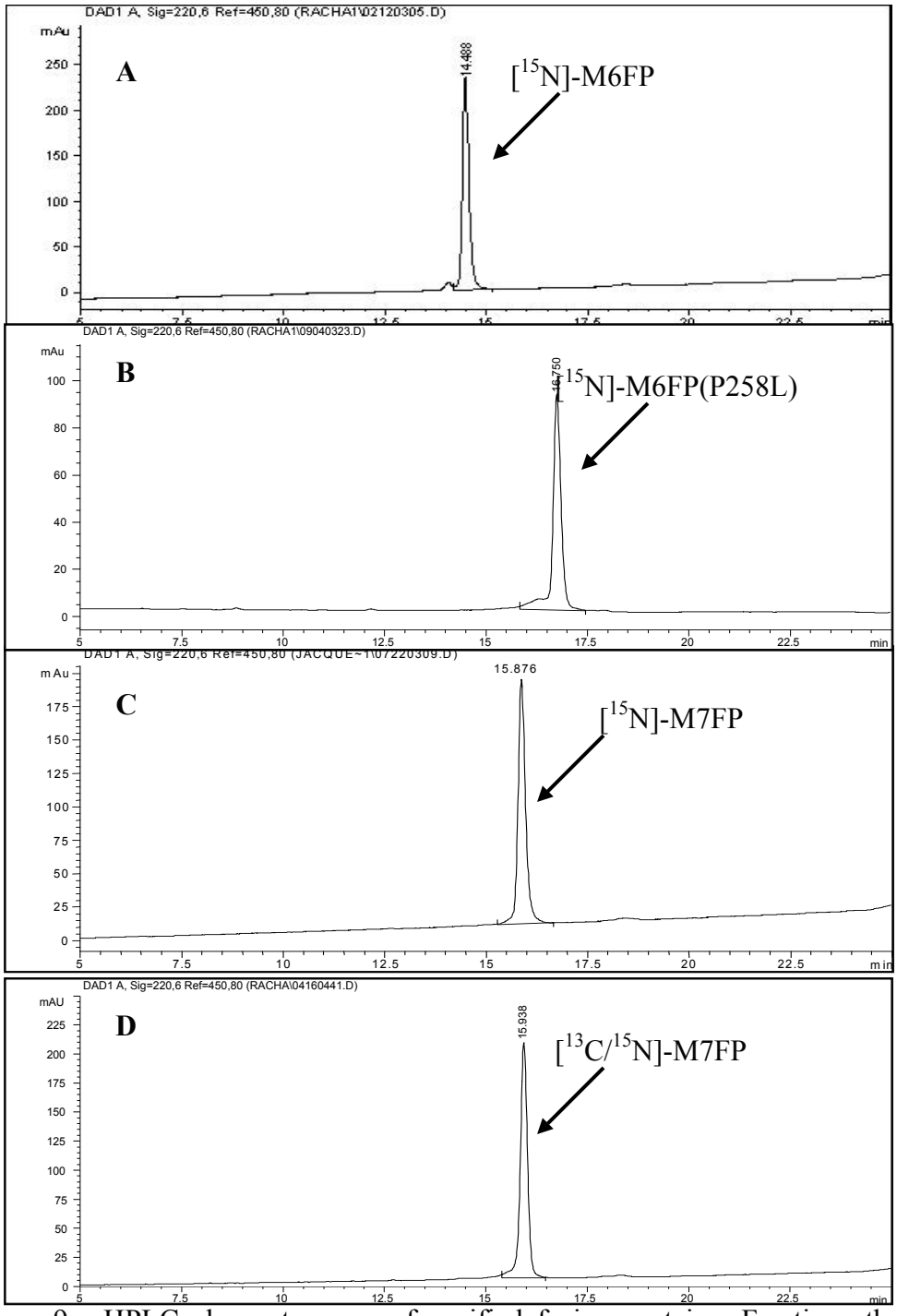

Figure 9. HPLC chromatograms of purified fusion proteins. Fractions that were collected during purification were analyzed on an analytical polymer reversed-phase column in an acetonitrile-water gradient + 0.1% TFA (30-60% in 20 min). [^{15}N]-M6FP (A), [^{15}N]-M6FP(P258L) (B), [^{15}N]-M7FP (C), and [^{13}C/^{15}N]-M7FP (D) peaks are observed at a retention time of around 14.5, 16.7, 15.8, and 15.9 min, respectively.

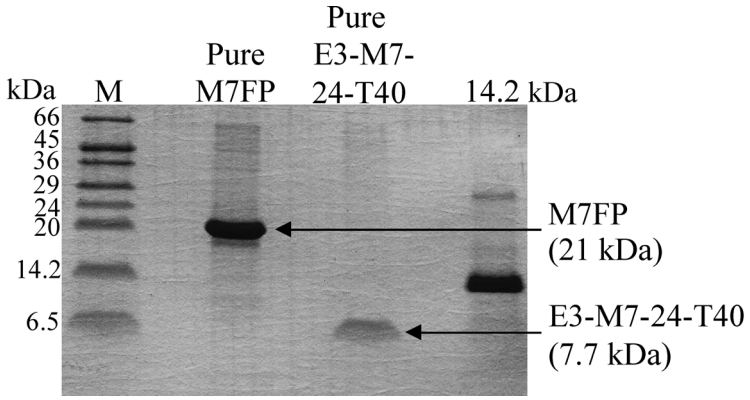

Figure 10. SDS-PAGE of pure M7FP and pure E3-M7-24-T40 after purification with HPLC. The lyophilized proteins were resuspended in 100 μL sample buffer and analyzed on a 16% polyarcylamide gel. The M7FP has a molecular weight of 21 kDa whereas the E3-M7-24-T40 peptide has a molecular weight of 7.7 kDa.

The molecular weights of all fusion proteins were measured by ESI-MS. Figure 11 shows the non-deconvoluted mass spectrum of the M6FP, M6FP(P258L), and M7FP fusion proteins. Each peak is denoted with a charge (Z) [+13, +14, …, +30) that corresponds to the mass-to-charge (m/*z*) ratio values. Calculation of the mass of M6FP, M6FP(P258L), and M7FP using the formula [(m/*z*) x *z* – *z*] yielded average values of 17312.6 Da, 17331.2 Da, and 21234.9 Da, respectively. Therefore, the molecular weights of the fusion proteins determined by mass spectrometry are consistent with the calculated values (17.313 kDa for M6FP(WT), 17.329 kDa for M6FP(P258L), and 21.234 kDa for M7FP).

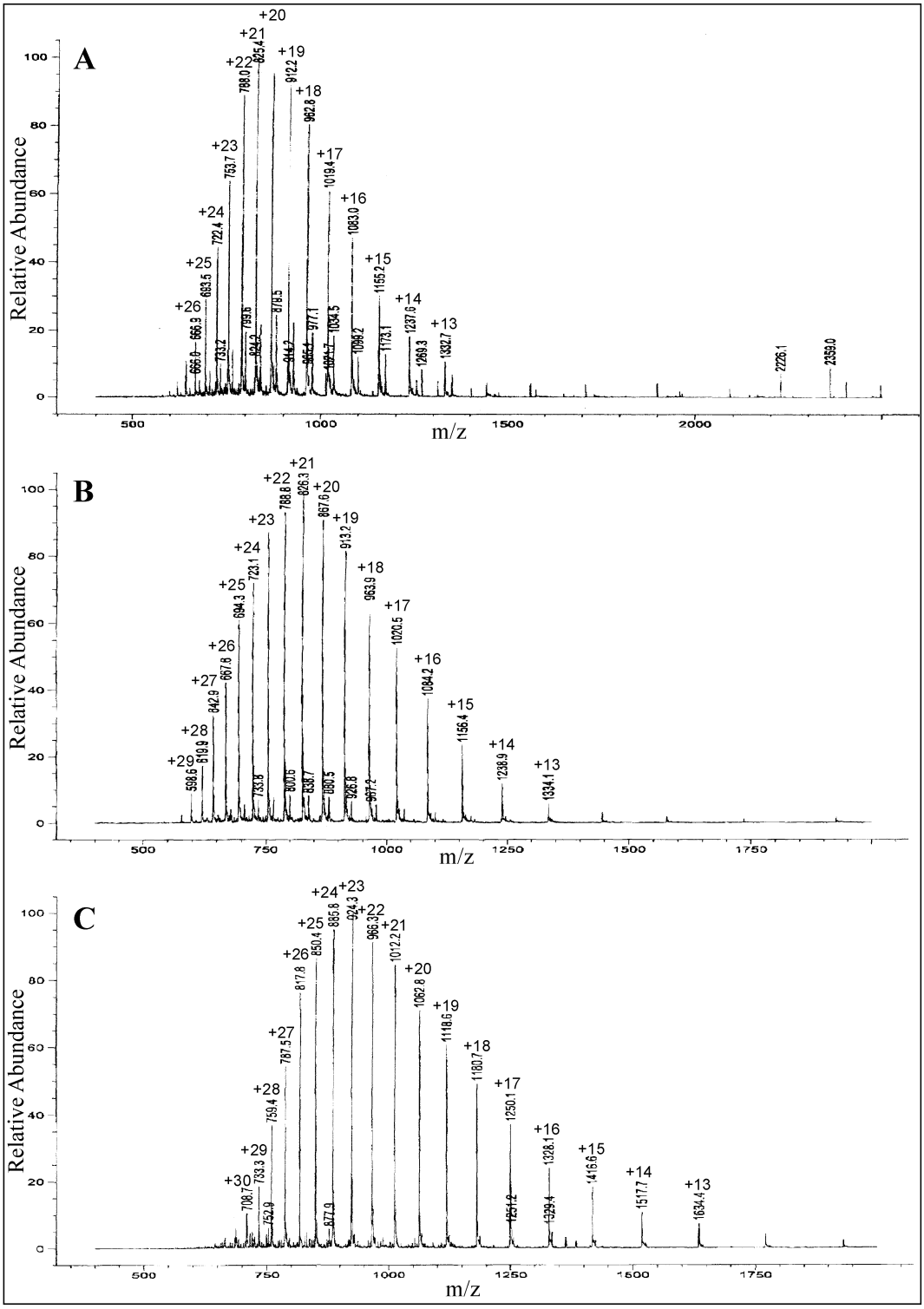

Figure 11. Non-deconvoluted mass spectra of the M6FP, M6FP(P258L), and M7FP fusion proteins. Each peak represents a specific mass-to-charge (m/*z*) ratio from the electrospray ionization. The calculation of the mass of (A) M6FP, (B) M6FP(P258L), and (C) M7FP yielded 17312.6 Da, 17331.2 Da, and 21234.9 Da, respectively. ESI-MS analysis of the fusion protein was performed at the mass spectrometry facilities at Hunter College, CUNY.

III.1.A. *Release of membrane peptides using CNBr cleavage*

The CNBr cleavage reaction was performed in TFA using a 5000 molar excess when cleaving M6FP and M6FP(P258L) and 5000 and 500 molar excesses of CNBr when cleaving M7PF. In order to optimize release of peptides from fusion proteins the time course of their release by CNBr was followed using HPLC. Optimal cleavage was obtained with 1M CNBr (5000 molar excess) in 70% TFA for 24 h for M6FP (Figure 12, see Table 6) and M6FP(P258L) [Figure 13, see Table 6].

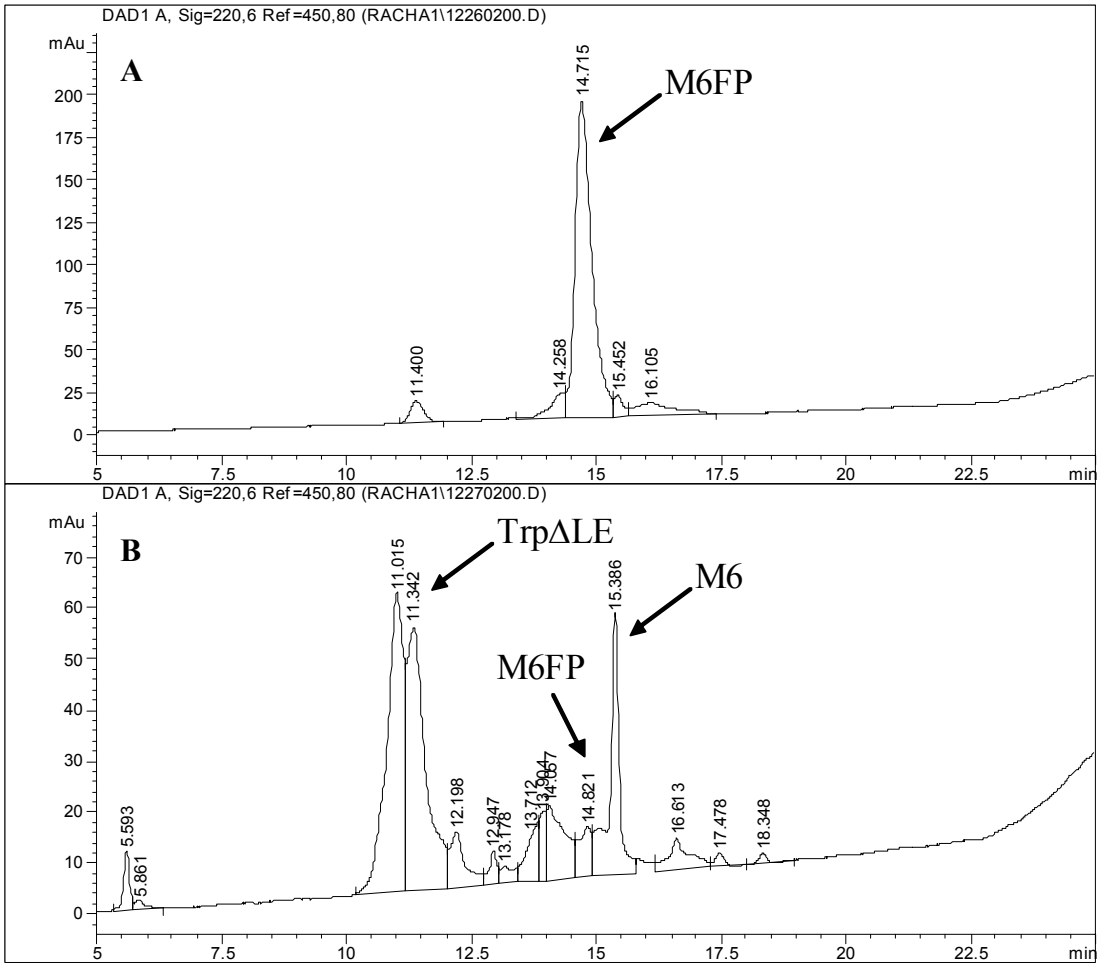

Figure 12. HPLC chromatograms of the [^{15}N]-M6FP fusion protein cleavage. The [^{15}N]-M6FP fusion protein was cleaved with 1M CNBr (5000 molar excess) in 70% TFA for 24 h in the dark at room temperature. The cleavage reaction was monitored at zero hour (A) and 24 h (B) using a WATERS C18 analytical column and an acetonitrile-water gradient + 0.1% TFA (30-60% in 20 min).

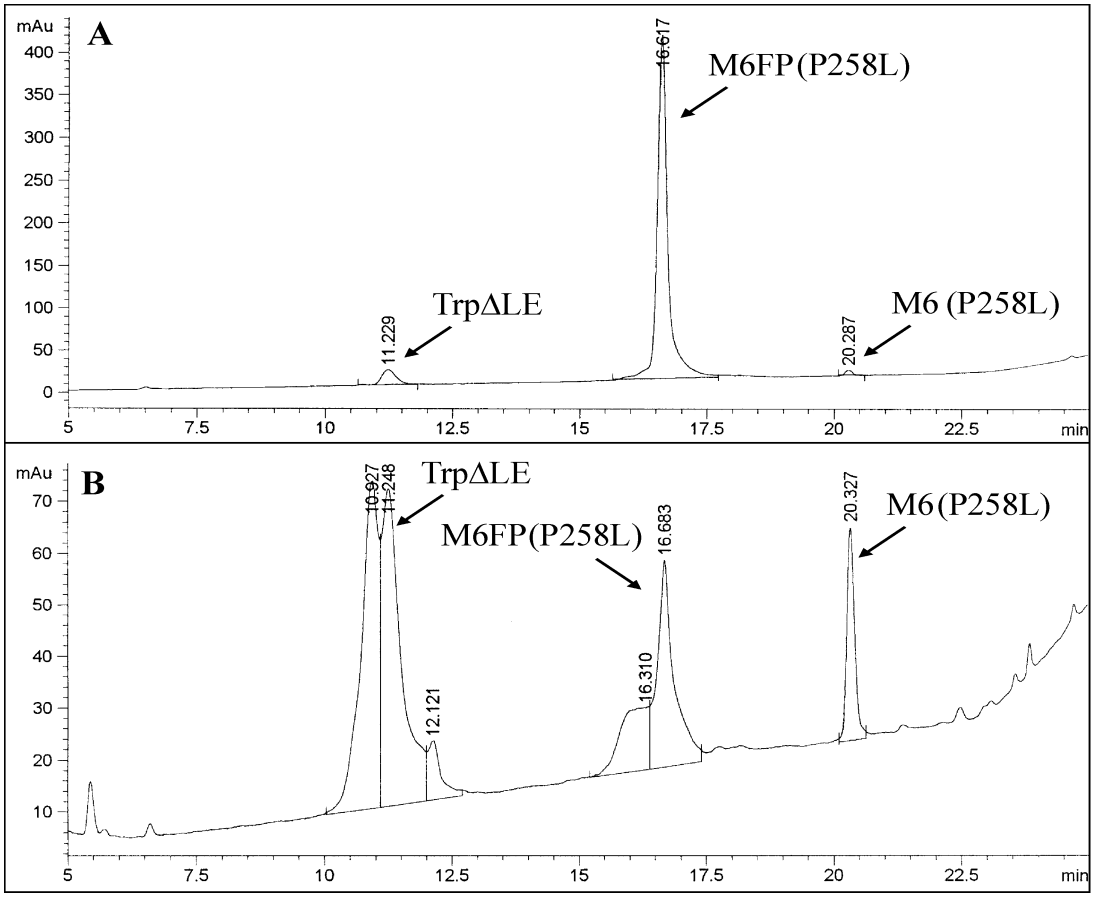

Figure 13. HPLC chromatograms of the unlabeled-M6FP(P258L) fusion protein cleavage. The unlabeled-M6FP(P258L) fusion protein was cleaved with 1M CNBr (5000 molar excess) in 70% TFA for 24 h in the dark at room temperature. The cleavage reaction was monitored at zero hour (A) and 24 h (B) using a WATERS C18 analytical column and an acetonitrile-water gradient + 0.1% TFA (30-60% in 20 min).

Table 6

Time course of M6FP and M6FP(P258L) fusion protein cleavage using 5000 M excess of CNBr in 70% TFA[a]

Hours	M6FP	M6	M6FP(P258L)	M6(P258L)
0 hr	60%	4.8%	93%	1%
6 hr	12.7%	9%	____	____
12 hr	11.7%	10.3%	____	____
24 hr	8.4%	17.6%	16.8%	8%
27 hr	____	____	16%	7.6%
30 hr	8%	15%	____	____

[a] Percentage of peptides was measured from the total area of the HPLC chromatogram peaks.

The HPLC monitoring indicated that maximum product was obtained for M6 and M6(P258L) peptides at about 24 h with most of the fusion proteins cleaved at this time (Table 6). Further incubation apparently resulted in no further cleavage or degradation of the product. Thus each reaction was stopped at its maximum cleavage time (24 h). Purification via HPLC of the cleaved product from the 5000-fold molar excess CNBr reaction resulted in a yield of ~1 mg of pure peptide released per 10 mg of pure fusion protein cleaved.

Chromatograms produced from the time series of the 5000 M excess of CNBr in 70% TFA cleavage reaction indicated that maximum cleavage of E3-M7-24-T40 from the TrpΔLE occurred by 2 h (Figure 14), while chromatograms from the 500 M excess reaction time series in 70% TFA showed maximum cleavage between 4 and 5 h (Figure 15, see Table 7). The chromatograms for CNBr cleavage of the unlabeled M7FP and [^{15}N]-M7FP fusion protein cleavage reactions (data not shown) were very similar to that of the [^{13}C/^{15}N]-M7FP fusion protein (Figure 15). The TrpΔLE in Figure 15 is observed at a different retention time (5.5 min) than that observed in Figure 14 because these cleavage reactions were analyzed on different HPLC instruments.

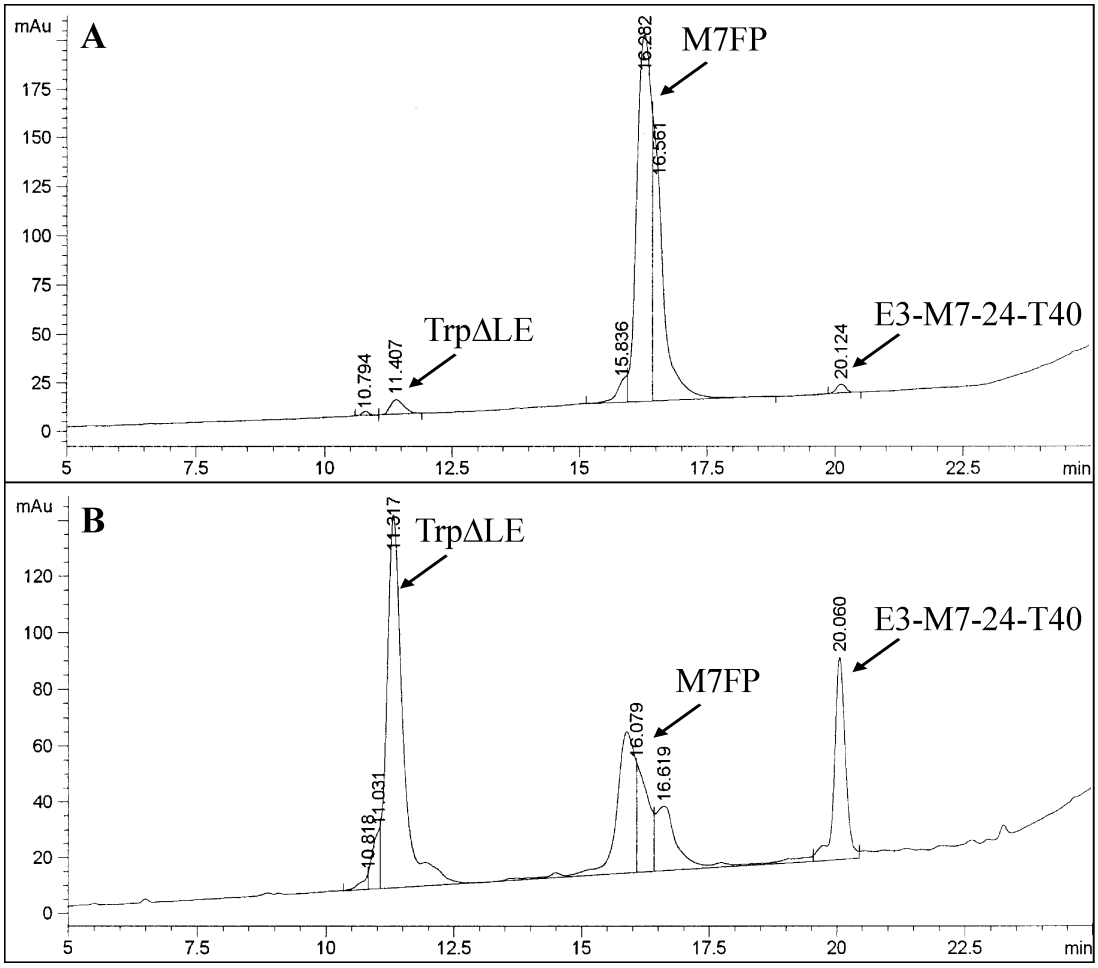

Figure 14. HPLC chromatograms of the unlabeled-M7FP fusion protein cleavage. The unlabeled-M7FP fusion protein was cleaved with 1M CNBr (5000 molar excess) in 70% TFA for 2 h in the dark at room temperature. The cleavage reaction was monitored at zero hour (A) and 2 h (B) using a WATERS C18 analytical column and an acetonitrile-water gradient + 0.1% TFA (30-60% in 20 min).

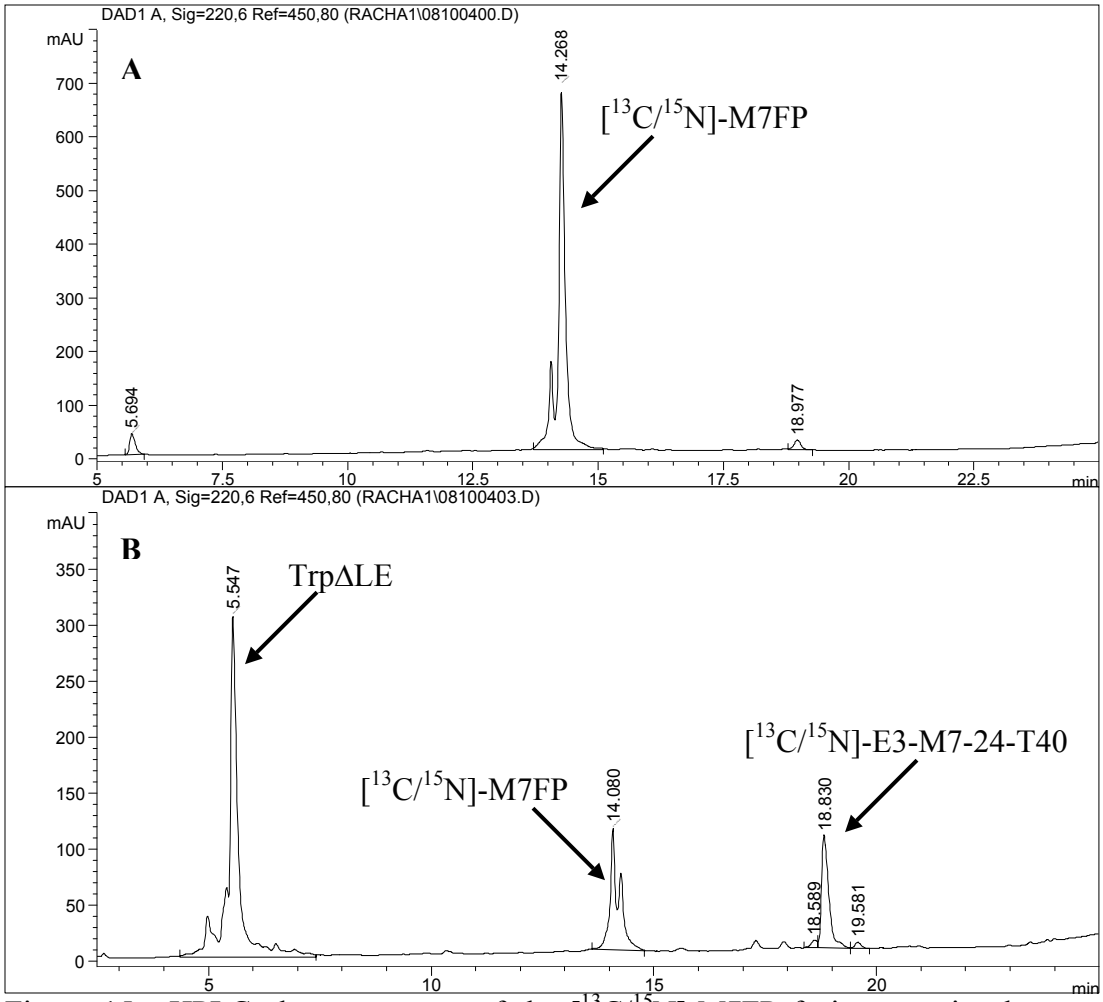

Figure 15. HPLC chromatograms of the [^{13}C/^{15}N]-M7FP fusion protein cleavage. The [^{13}C/^{15}N]-M7FP fusion protein was cleaved with 1M CNBr (500 molar excess) in 50% TFA for 4.5 h in the dark at room temperature. The cleavage reaction was monitored at zero hour (A) and 4.5 h (B) using a WATERS C18 analytical column and an acetonitrile-water gradient + 0.1% TFA (30-60% in 20 min).

Table 7

Time course of M7FP fusion protein cleavage using 500 M excess of CNBr in 50% and 70% TFA [a]

	M7FP		E3-M7-24-T40	
Hours	50% TFA	70% TFA	50% TFA	70% TFA
0 hr	89%	94.3%	3%	1.7%
2 hr	40%	40%	14.8%	14%
3.5 hr	29%	28%	16%	15.2%
4.5 hr	22.3%	20.5%	16.6%	14%
5 hr	20.4%	20.5%	16.3%	14.3%
6 hr	17.4%	19.2%	15.5%	13.5%
12 hr	13.7%	12.9%	12.9%	10.8%
24 hr	5%	5.6%	8.7%	6%

[a] Percentage of peptides was measured from the total area of the HPLC chromatogram peaks.

Similar results were obtained using sublimed CNBr or newly purchased reagent. Interestingly, the HPLC monitoring indicated that maximum product was obtained for E3-M7-24-T40 at about 4.5 h despite the fact that significant fusion protein remained at this time (Table 7). Further incubation apparently resulted in degradation of the product. Although each reaction was stopped at its maximum cleavage time, the reaction using a 5000 M excess of CNBr produced a much lower yield of pure E3-M7-24-T40 peptide than the reaction using only a 500 M excess. Purification via HPLC of the cleaved product from the 500-fold molar excess CNBr reaction resulted in a yield of 1.1 mg of pure E3-M7-24-T40 per 10 mg of pure M7FP cleaved compared to <0.5mg for the 5000 M excess reaction.

Once it was determined that cleavage with a 500 M excess of CNBr produced better yields, the effect of the TFA percentage on the reaction was determined. A cleavage time course study was performed in which the M7FP was cleaved in 50%

TFA and 70% TFA over 24 h. The reactions were monitored at different times and the chromatograms were analyzed for the percent cleaved E3-M7-24-T40 protein (Table 7). The results indicated that there was no significant difference between the solvents with regards to the percent of cleaved E3-M7-24-T40. In both cases, at maximum cleavage, the M7FP still maintains a significant presence in the reaction solution. In order to obtain almost full cleavage of the fusion protein the reaction has to be left for 24 h. However, when this is done the cleaved product begins to degrade. Based on these results it was decided to use the 50% TFA solvent system because it cleaves the peptide at a slower rate than the 70% TFA system, and to stop the cleavage reaction after 4 to 5 h.

The cleaved fusion proteins were separated and purified using a preparative C18 column (Figure 16). All final peptides were ≥ 95% homogeneous as judged by analytical HPLC (Figure 17) and had the correct calculated molecular weight as judged using electrospray ionization mass spectrometry (ESI-MS) [Figure 18]. In addition, the purity of the E3-M7-24-T40 peptide was verified by SDS-PAGE analysis (Figure 10). However, SDS-PAGE analysis was not a good tool to confirm the purity of the M6 and M6(P258L) peptides due to their low molecular weight (~ 3.7 kDa), poor solubility, and aggregation in SDS-PAGE. Mass spectrometry indicated a 98% isotope incorporation for the [^{15}N]-E3-M7-24-T40 and [^{15}N]-M6 peptides and a 97% isotope incorporation for the [^{13}C/^{15}N]-E3-M7-24-T40 peptide (Table 8, Figure 18). The percent isotope incorporation was calculated by comparing the experimental molecular weight difference between the labeled and unlabeled peptides to that of the calculated difference. (Table 8, Figure 18). The molecular weights of all cleaved

peptides were measured by ESI-MS. Figure 16 shows the non-deconvoluted mass spectrum of pure unlabled-M6, unlabeled-M6(P258L), [^{15}N]-E3-M7-24-T40, and [^{13}C/^{15}N]-E3-M7-24-T40 peptides. Each peak is denoted with a charge (Z) [+3, +4, …, +9) that corresponds to the mass-to-charge (m/*z*) ratio values. Calculation of the mass of M6, M6(P258L), [^{15}N]-E3-M7-24-T40, and [^{13}C/^{15}N]-E3-M7-24-T40 peptides using the formula [(m/*z*) x *z* – *z*] yielded average values of 3748.8 Da, 3766.6 Da, 7756.7 Da, and 8078.9 Da, respectively. Therefore, the molecular weights of the peptides are consistent with the calculated values (3.750 kDa for M6, 3.766 kDa for M6(P258L), 7.759 kDa for [^{15}N]-E3-M7-24-T40, and 8.093 kDa for [^{13}C/^{15}N]-E3-M7-24-T40 peptides).

The CNBr cleavage of the [^{15}N]-labeled M7FP using a 500-fold molar excess of CNBr and 50% TFA produced similar results to that of the unlabeled fusion protein cleavage reaction, with an average of 1 mg of pure E3-M7-24-T40 per 10 mg of pure M7FP cleaved. However, cleavage of the [^{13}C/^{15}N]-labeled fusion protein using the same conditions gave a lower yield (0.62 mg per 10 mg of pure M7FP) because the separation ability of the column was diminished. Despite the lower yield we could conclude that the cleavage reaction was successful based on the similarity of the HPLC chromatograms to those of the unlabeled and [^{15}N]-labeled M7FP cleavages (Figure 16C,D).

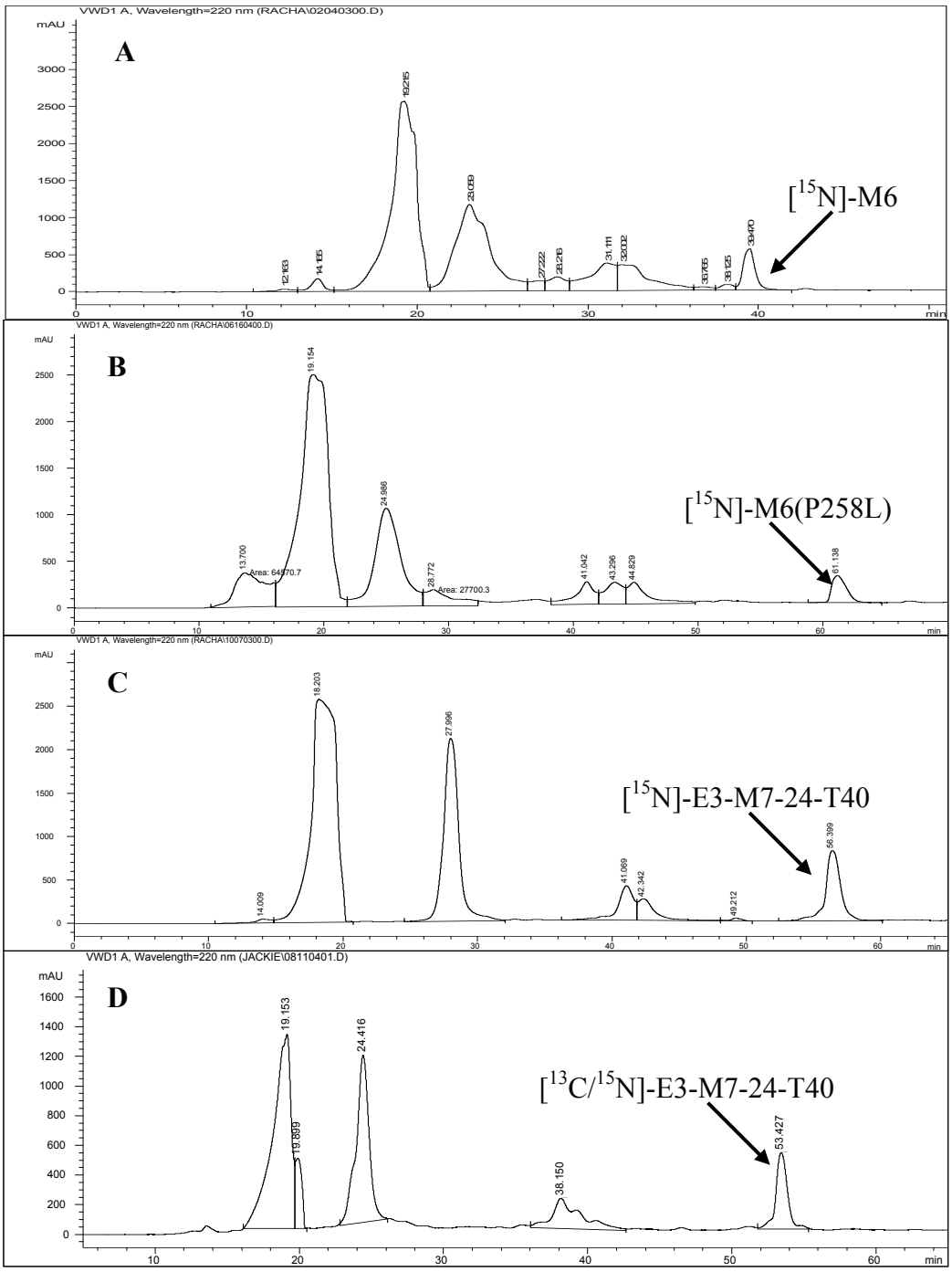

Figure 16. HPLC chromatogram of crude cleaved fusion proteins. The ^{15}N-M6FP or ^{15}N-M6FP(P258L) and the ^{15}N- or ^{13}C/^{15}N-M7FP were cleaved with 1M CNBr in 70% TFA for 24 h and 50% TFA for 4.5 h, respectively (see Materials and Methods section), and then purified using a preparative C18 reversed-phase column and an acetonitrile-water gradient + 0.1% TFA (40-80% in 80 min). The ^{15}N-M6 (A), ^{15}N-M6(P258L) (B), ^{15}N-E3-M7-24-T40 (C), and ^{13}C/^{15}N-E3-M7-24-T40 (D) peptides are observed at a retention time of around 39, 61, 56, and 53 min, respectively.

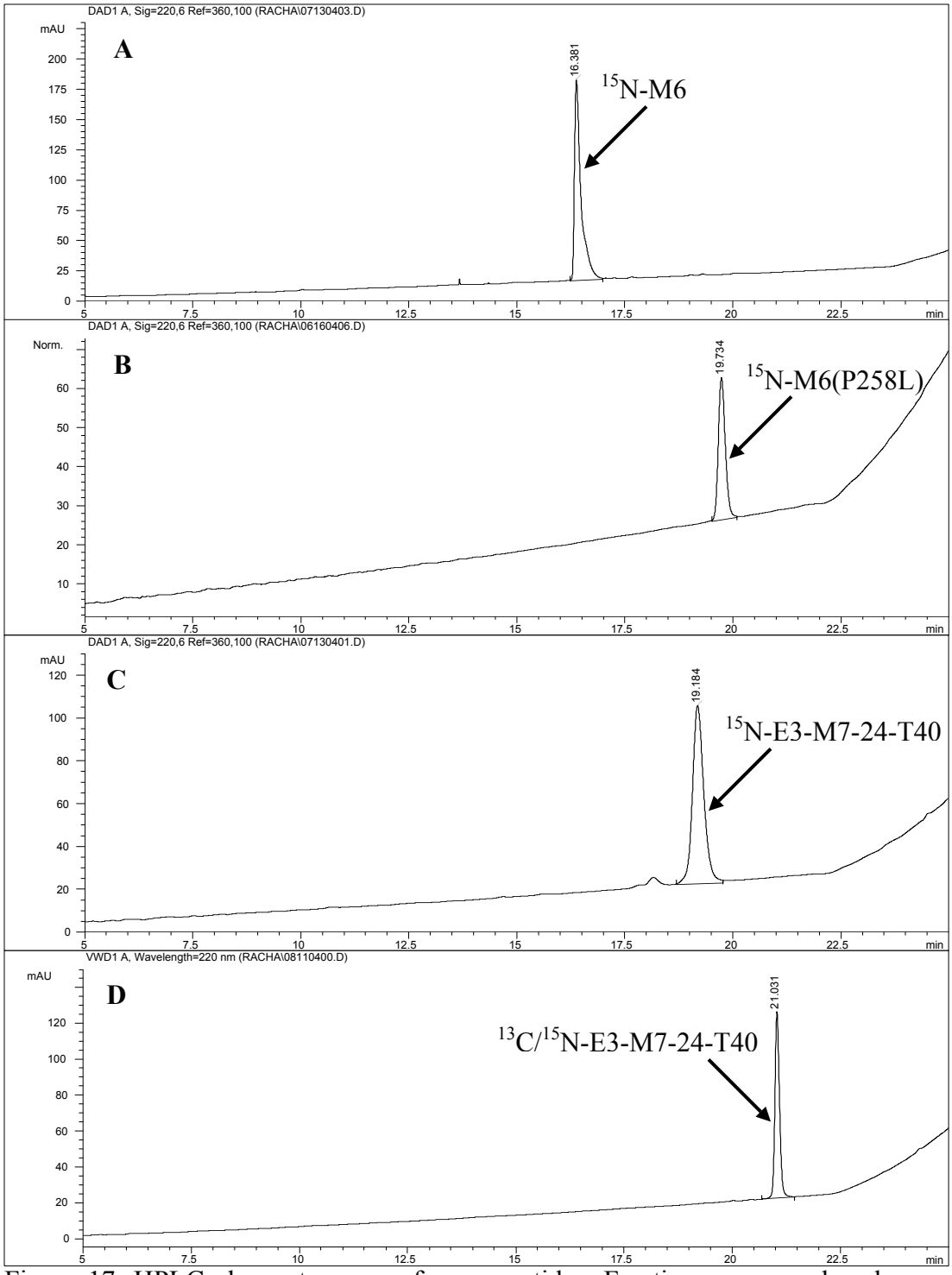

Figure 17. HPLC chromatograms of pure peptides. Fractions were analyzed on an analytical C18 reversed-phase column in an acetonitrile-water gradient + 0.1% TFA (30-60% in 20 min). The ^{15}N-M6 (A), ^{15}N-M6(P258L) (B), ^{15}N-E3-M7-24-T40 (C), and ^{13}C/^{15}N-E3-M7-24-T40 (D) peptides are observed at a retention time of around 16.4, 20.0, 19.8, and 21 min, respectively. The ^{15}N-E3-M7-24-T40 (C) and ^{13}C/^{15}N-E3-M7-24-T40 (D) peptides are observed at different retention times because they were analyzed using different HPLC instruments.

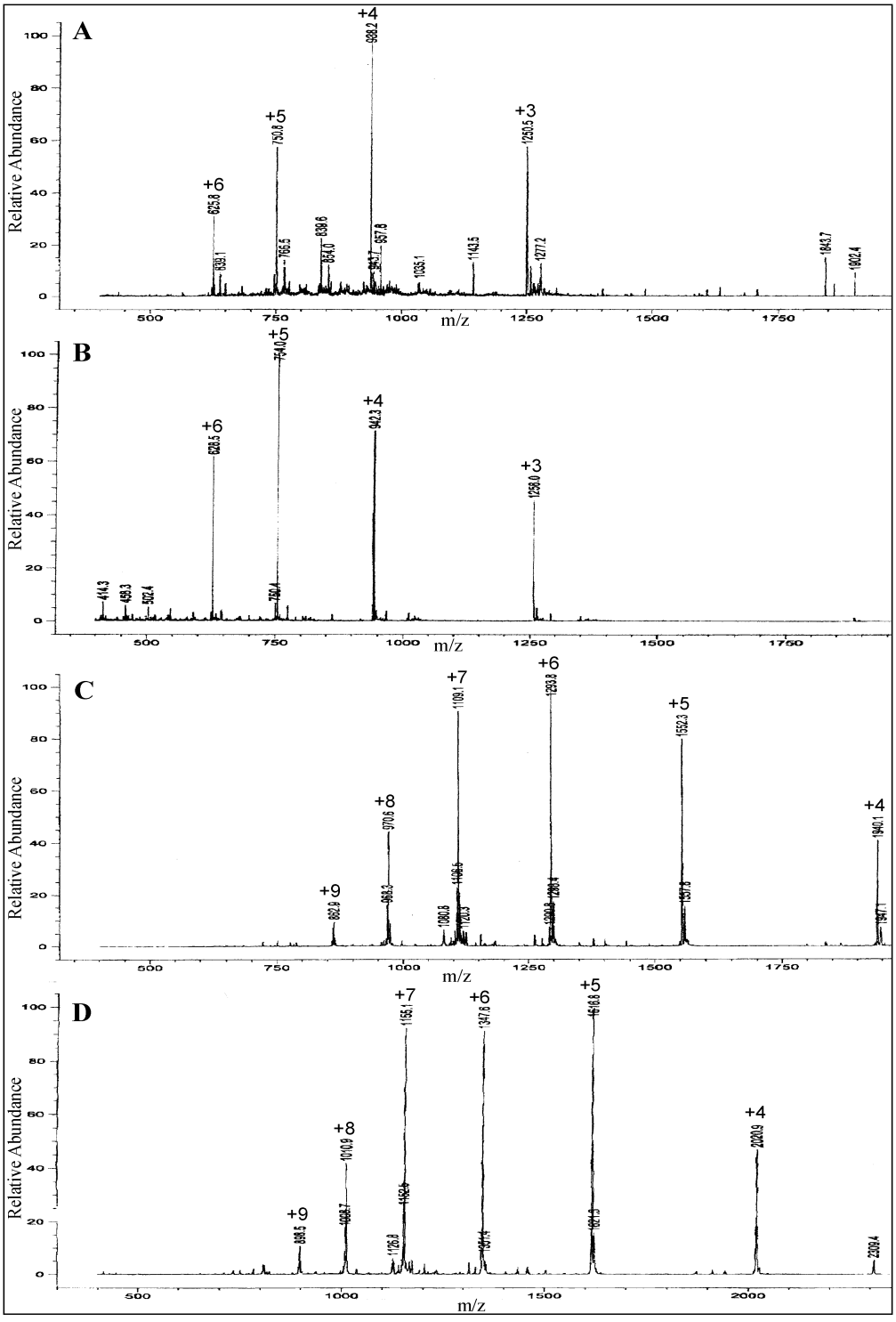

Figure 18. Non-deconvoluted mass spectra of pure unlabled-M6, unlabeled-M6(P258L), [^{15}N]-E3-M7-24-T40, and [^{13}C/^{15}N]-E3-M7-24-T40 peptides. Each peak represents a specific mass-to-charge (m/*z*) ratio from the electrospray ionization. The calculation of the mass of (A) unlabled-M6, (B) unlabeled-M6(P258L), (C) [^{15}N]-E3-M7-24-T40, and (D) [^{13}C/^{15}N]-E3-M7-24-T40, yielded 3748.8 Da, 3766.6 Da, 7756.7 Da, and 8078.9 Da, respectively.

Table 8

Summary of recoveries and molecular weights of peptides from purification after release from TrpΔLE fusion proteins via CNBr cleavage

Peptides	Recovery per liter of fermentation (mg)	Calculated molecular weight (Da)	Experimental molecular weight (Da)[a]	% Isotopic incorporation
Unlabeled M6	1.0 [b]	3,750	3,749	____
[^{15}N]-labeled M6	1.0 [b]	3,791	3,790 [d]	98%
Unlabeled M6(P258L)	1.2 [b]	3,766	3,766	____
Unlabeled M6(P258L)	2.1 [c]	3,766	3,765	____
[^{15}N]-labeled M6(P258L)	0.7 [b]	3,807	____	____
Unlabeled E3-M7-24-T40	15 [c]	7,671	7,671	____
[^{15}N]-labeled E3-M7-24-T40	3.9 [b]	7,759	7,757	98%
[^{13}C/^{15}N]-labeled E3-M7-24-T40	2.0 [b]	8,093	8,079	97%

[a] Molecular weights were determined by ESI-MS.
[b] Fermentation in minimal media (see Materials and Methods).
[c] Fermentation in rich media (see Materials and Methods).
[d] Molecular weight of [^{15}N]-labeled M6 measured by ESI-MS was determined by Enrique Arevalo (Arevalo, 2002).

We were able to isolate about 1 to 2 mg of the M6 peptides per liter of fermentation in rich or minimal medium. In contrast we isolated about 4 mg/per liter of [^{15}N]-labeled E3-M7-24-T40 and 2 mg/liter of [^{13}C/^{15}N]-labeled E3-M7-24-T40 from minimal medium and nearly 15 mg/liter of unlabeled E3-M7-24-T40 peptide from rich medium (Table 8).

III.2. Biophysical Studies

III.2.A. *Circular Dichroism*

III.2.A.1. *Studies on M6 Peptides.* Circular dichroism (CD) provides information on the overall secondary structure of a peptide or protein, but not on the individual residues in the chain. Previously it has been observed that peptides corresponding to the M6 domain of Ste2p had a significant tendency to aggregate in aqueous media (Xie et al., 2000). This made it difficult to carry out high resolution NMR analyses of the structure of these peptides in the presence of detergent. Our ability to produce [^{15}N]-labeled M6 and M6(P258L) led us to re-evaluate membrane mimetic conditions that would be amenable for the study of these peptides. CD curves for both M6 and M6(P258L) in various TFE/H_2O mixtures from 95% to 25% TFE exhibited minima near 208 nm and 222 nm and a maximum near 190 nm characteristic of highly helical peptides (Figure 19). In contrast in 5% TFE/95% H_2O the spectra of these peptides had one broad minimum near 215 nm, a broad maximum near 190 nm and were of very low intensity, indicating both aggregation and a loss of peptide from solution. UV measurements confirmed that the peptide concentration had decreased.

In the presence of PPG micelles CD spectra with shapes roughly characteristic of those expected for helices were observed for both M6 and M6(P258L) (Figure 20). In SDS micelles M6, also exhibited helical CD patterns as judged by the double minimum at 208 and 222 nm and the splitting of the $\pi \rightarrow \pi^*$ transition (Figure 20).

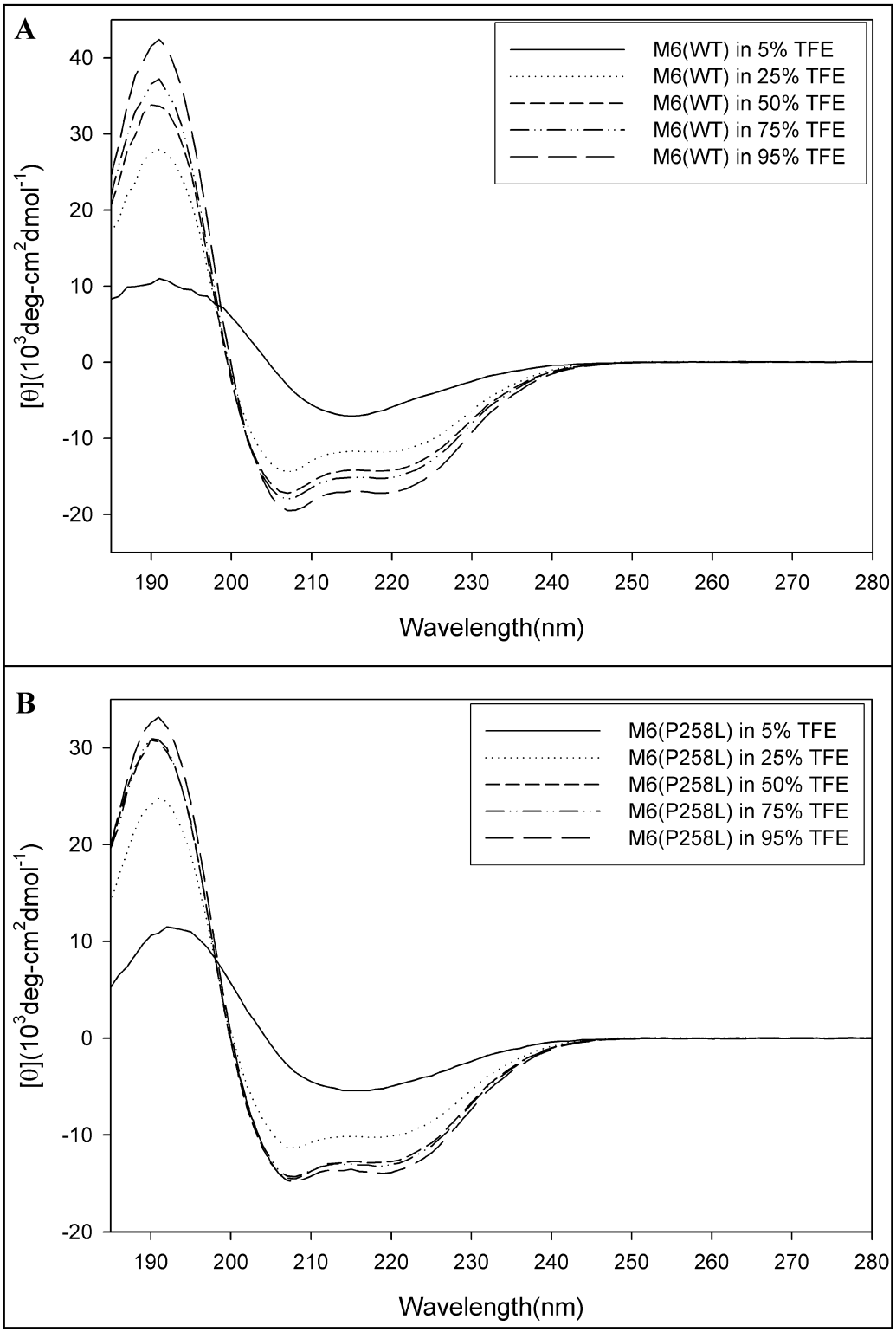

Figure 19. CD spectra of M6 peptides in TFE/H_2O. CD of M6(WT) (A) and M6(P258L) (B) were measured in TFE/H_2O at ratios of 95/5, 75/25, 50/50, 25/75, and 5/95 and at peptide concentration of 60 μM. The CD is calculated as mean residue ellipticity.

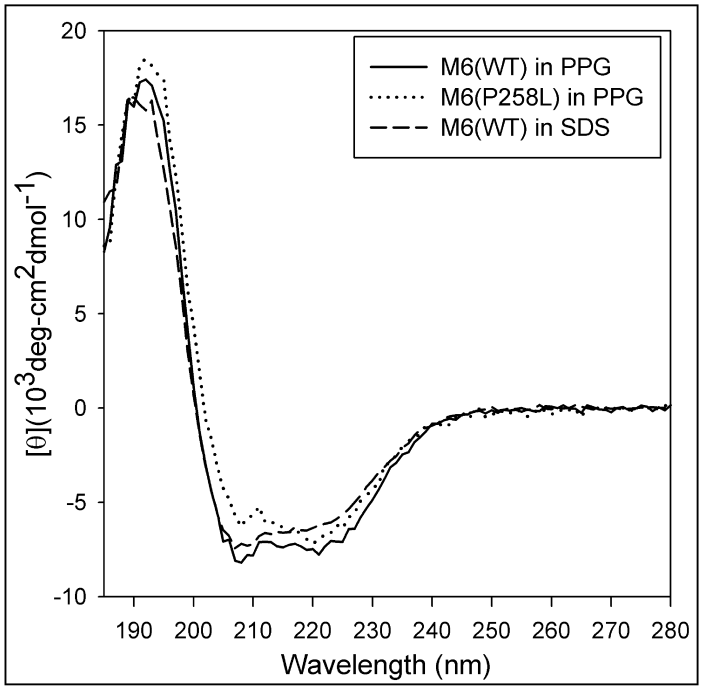

Figure 20. CD spectra of M6 peptides in detergent micelles. CD was measured in the presence of 15 mM PPG micelles for M6(WT) and M6(P258L) at a peptide concentration of 50 μM and for M6(WT) at a peptide concentration of 94 μM in the presence of 69 mM SDS micelles. The CD is calculated as mean residue ellipticity.

However, we were unable to get concentrations higher than ~90 μM in detergent media. The mean residue ellipticities [θ] measured on these peptides are very low for a helical peptide. The low [θ] values obtained indicate that we may be losing peptide during the course of the sample preparation or analysis. The conclusion on helicity is made on the basis of overall spectral appearance and the ratio of the 195 and 208 nm peaks. We found it very difficult to accurately determine the concentration of the M6 peptides in SDS detergent because the peptide concentrations measured by UV spectroscopy were surprisingly larger than the theoretical concentrations that were calculated using weight measurements. Our results suggested that it would be very difficult to study these M6 domain peptides in the presence of detergent using NMR

spectroscopy. In the presence of DMPC/DMPG vesicles the CD patterns for both M6 and M6(P258L) had one broad minimum from about 230 to 205 nm indicative of the presence of β-sheets or aggregation (Figure 21). Both M6(WT) and M6(P268L) peptides had very poor solubility in medium containing DMPC/DMPG vesicles.

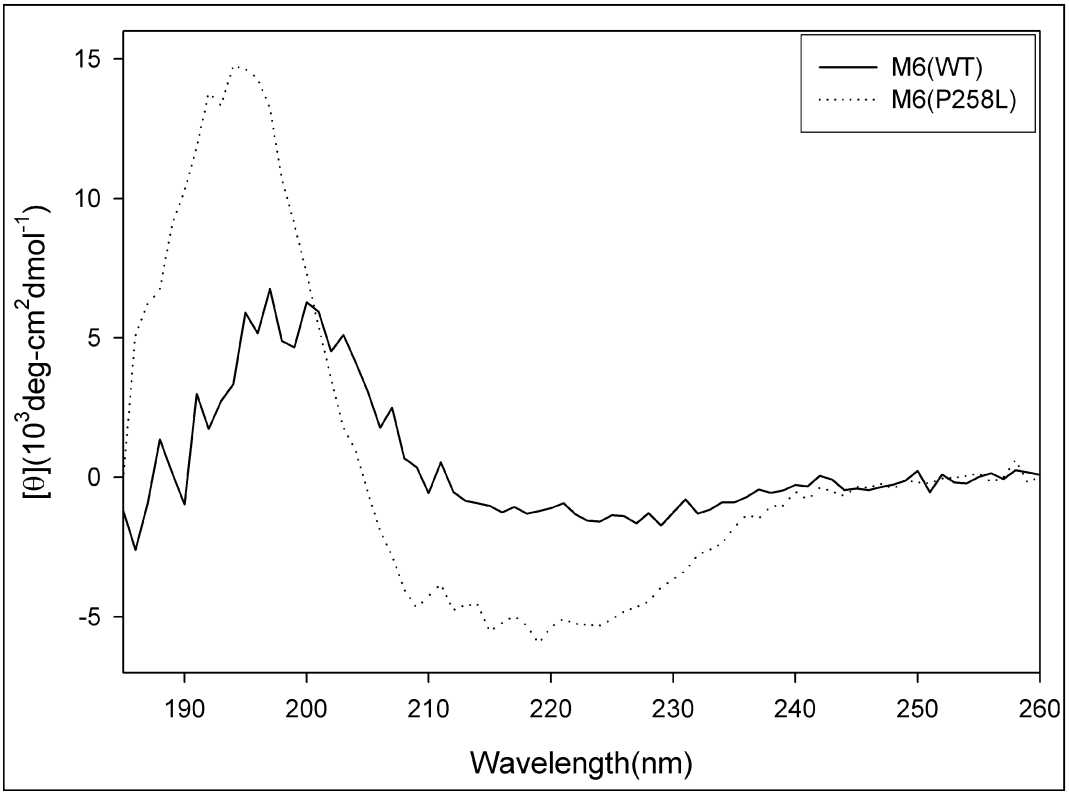

Figure 21. CD spectra of M6 peptides in DMPC/DMPG vesicles. CD was measured in the presence of DMPC/DMPG vesicles (4 mg/mL) at a peptide concentration of 50 μM. The CD is calculated as mean residue ellipticity.

III.2.A.2. *Studies on the E3-M7-24-T40 peptide.* CD studies on E3-M7-24-T40 were carried out in TFE/H_2O mixtures, DPC and PPG micelles and DMPC/DMPG (4:1) vesicles. The CD pattern of the 73-residue peptide (E3-M7-24-T40) was indicative of a partially structured peptide in 5% TFE/95% H_2O and in 25% TFE/75% H_2O two minima at about 208 nm and 220 nm, characteristic of a helical conformation, were evident (Figure 22A). As the water content decreased to 5% the

intensity of the 208 nm and 222 nm minima and of the 190 nm maxima increased. Quantitatively similar spectra were measured in aqueous media containing 50% or 75% TFE (Figure 22A).

The CD of E3-M7-24-T40 in TFE/H_2O (1:1) was measured at various temperatures (Figure 22B). Over the range 10ºC to 110ºC a significant decrease in the ellipticity at 208 nm and 222 nm was observed. However, even at 110ºC this peptide exhibited a discernable n→π* transition at 222 nm, and a π→π* minimum that had shifted from 208 nm at 10ºC to 203 nm at 110ºC was observed. The ellipticity at 222 nm at 110ºC (-10,067 deg-cm^2 $decimole^{-1}$) indicated 23% helix as compared to 43% measured at 25ºC (Figure 22B and A, respectively).

To determine the feasibility of studying the structure of isotopically-labeled peptides in a more membrane-like environment by NMR, we measured CD spectra in the presence of micelles. The CD spectra of E3-M7-24-T40 in DPC (data not shown) and PPG (Figure 22C) were very similar and showed that at 390-500 μM peptide concentration in the presence of these detergents the 73-residue peptide was partially helical with no significant change in the CD pattern when using low (10 mM) and high (100 mM) detergent concentrations. The CD pattern of T40 at 500 μM peptide concentration in the presence of 10 mM PPG micelles in phosphate buffer is indicative of a disordered peptide exhibiting one negative band centered at 198 nm (Figure 22C). This allowed estimation of the CD pattern for the E3-M7-24 domain of the peptide by difference spectroscopy (see Materials and Methods, equation 2). The CD pattern calculated for E3-M7-24 indicated a high degree of helicity (54%) [Figure

22C]. No difference in these spectra was observed at low (20:1) and high (200:1) detergent/peptide molar ratios.

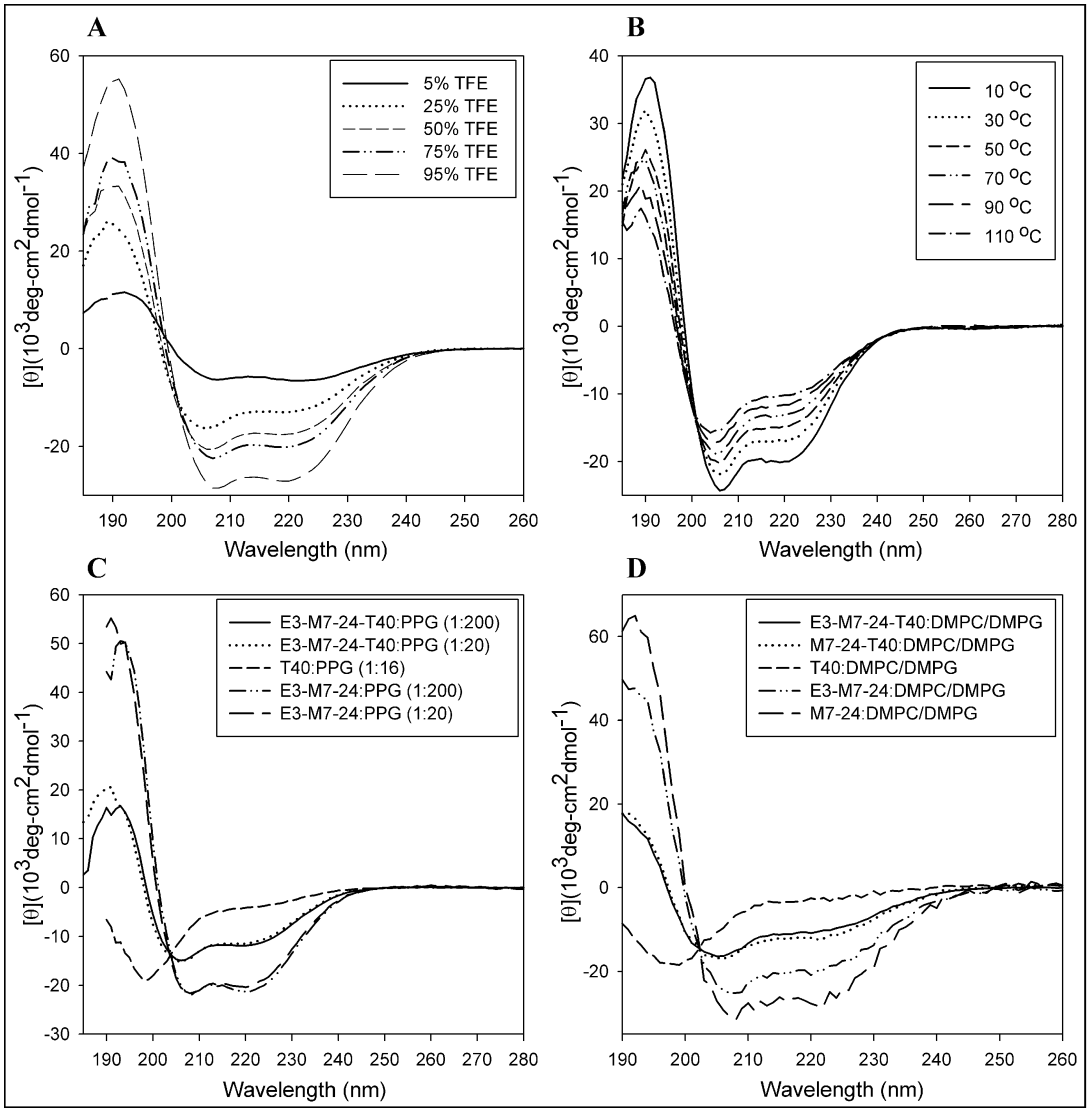

Figure 22. CD of E3-M7-24-T40 in different solvents. (A) In TFE/H_2O at different TFE percentages (5-95% by volume), (B) TFE/H_2O (1:1) at different temperatures (10-110 °C). (C) In the presence of PPG micelles at low and high peptide to lipid ratios. The peptide:lipid ratio is indicated in the parenthesis and (D) In the presence of DMPC/DMPG (4:1) vesicles. The peptide:lipid ratio is 1:117. The concentrations of peptides were 40, 20, 500, and 50 μM for E3-M7-24-T40 (A, B, C, and D, respectively), 50 μM for M7-24-T40 (D), and 500 and 50 μM for T40 (C and D, respectively). The concentrations of detergent or phospholipid were 10 mM and 100 mM in C and 4 mg/mL in D. Experimental spectra were measured for T40, E3-M7-24-T40, and M7-24-T40 and the difference spectrum was calculated for E3-M7-24 (E3-M7-24-T40 minus T40) in C & D and M7-24 (M7-24-T40 minus T40) in D [See Materials and Methods, equation 2].

The E3-M7-24-T40 peptide was also reconstituted into DMPC/DMPG (4:1) vesicles in phosphate buffer (see Materials and Methods). Under these conditions T40 is disordered exhibiting one negative band centered at 198 nm (Figure 22D). Both E3-M7-24-T40 and M7-24-T40, a synthetic 64-residue peptide corresponding to the seventh transmembrane domain and the tail of Ste2p showed a double minimum at 208 nm and 222 nm and positive absorbance at 190 nm indicative of significant α-helical secondary structure (Figure 22D). Difference CD spectra representing the transmembrane fractions of E3-M7-24 and M7-24 of Ste2p were obtained by subtracting the T40 tail from E3-M7-24-T40 and M7-24-T40, respectively. Both the E3-M7-24 and M7-24 fragments showed double minima at 208 nm and 222 nm and positive absorbance at 193 nm (Figure 22D). The calculated percent helix for M7-24 was 75% and the calculated percent helix for E3-M7-24 was 50%.

III.2.B. *NMR Spectroscopy*

NMR spectroscopy can provide highly detailed structural information concerning individual residues in a peptide chain. One and two-dimensional proton NMR spectroscopy has been quite useful for determining the 3D structure of small peptides. However, the determination of the 3D structure of large peptides and small proteins requires the use of multidimensional heteronuclear 2D, 3D, and 4D NMR spectroscopy and of isotopically labeled peptides.

To investigate the E3-M7-24-T-40 peptide we have used 1D proton and 2D homonuclear nuclear Overhauser effect spectroscopy (NOESY) and total correlation spectroscopy (TOCSY) NMR experiments and heteronuclear 2D ^{1}H-^{15}N Heteronuclear Single Quantum Coherence (HSQC) and 3D HSQC-TOCSY and HSQC-NOESY experiments. HSQC is a two-dimensional one-bond heteronuclear correlation experiment in which the two frequency coordinates of a resonance are the chemical shifts of a heteronucleus (^{15}N or ^{13}C) and a proton that are directly scalar coupled through one bond. In principle, the cross-peaks in ^{1}H-^{15}N HSQC spectrum belong to backbone NH groups, side-chain NH_2 groups (doublets) from Asn and Gln residues, and side-chain NH groups from Arg, Trp, and His residues.

Running a ^{1}H-^{15}N HSQC experiment is a simple and efficient way to check the condition of the sample. Changes in peak position or linewidth indicate changes in sample condition, such as sample degradation, changes in the pH or the composition of the solvent, peptide aggregation and/or denaturation.

III.2.B.1. *Evaluating the feasibility of NMR spectroscopy of E3-M7-24-T40 in organic-aqueous and detergent media.* A set of heteronuclear ^{1}H-^{15}N HSQC experiments have been performed to find out which solvent or medium is more appropriate for performing time consuming 3D heteronuclear NMR experiments which are necessary to make signal assignments and molecular modeling calculations for a 73-residue peptide.

Two organic-aqueous systems have been chosen: TFE/H_2O (1:1) and $CHCl_3$/CD_3OH/H_2O (4:4:1). These solvents have been widely used for examination of membrane peptides by NMR spectroscopy. In both of the organic-aqueous solvents well-resolved HSQC spectra of 0.5 mM [^{15}N]-E3-M7-24-T40 peptide with good signal to noise ratio were obtained in 40 min at 25 °C (Figure 23). In these media out of 69 expected crosspeaks corresponding to the backbone NHs (73 minus amine terminal serine and three prolines) between 65 and 70 crosspeaks were observed. In addition to the undetectable amine terminal and proline residues in an HSQC, the fewer number of crosspeaks observed than were expected could be due to an overlap of crosspeaks of two or more residues. Extra peaks might be due to conformational inhomogeneity and/or the presence of possible *cis*-isomers of proline residues. The expected number of doublets for the side chain NH_2 of six asparagine and two glutamine residues and one crosspeak corresponding to the side chain NH of an arginine residue were observed between 6.6 ppm and 7.7 ppm, and one crosspeak corresponding to the side chain NH of one tryptophan residue was observed at about 9.45 ppm. The high field crosspeaks corresponding to side chain NH_2 doublets of

asparagine and glutamine residues and the crosspeaks corresponding to the side chain NH of arginine and tryptophan residues were not plotted in Figure 23.

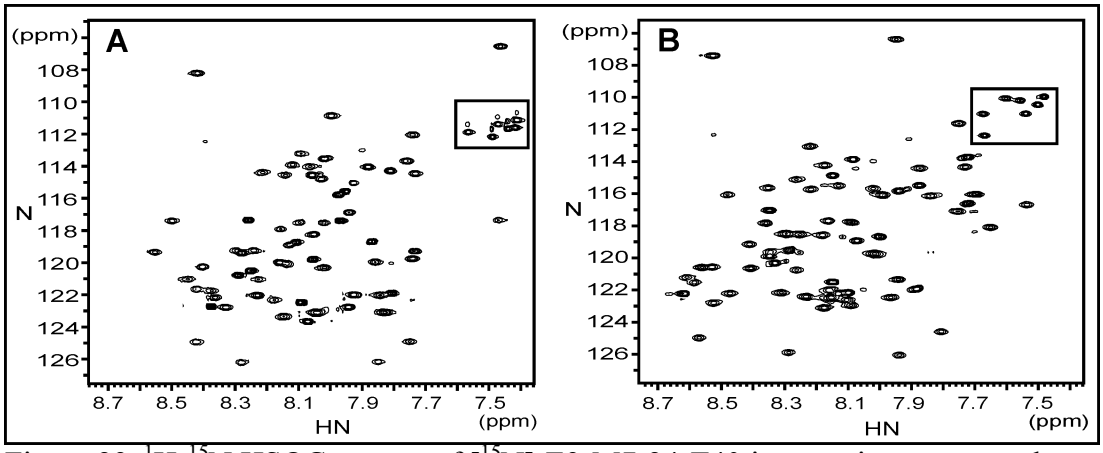

Figure 23. ^{1}H-^{15}N HSQC spectra of [^{15}N]-E3-M7-24-T40 in organic-aqueous solvents measured for 40 min. (A) TFE/H_2O (1:1), (B) $CHCl_3$/CH_3OH/H_2O (4:4:1). The peptide concentration was 0.5 mM. Backbone NH portion of the spectra are shown. Boxed crosspeaks belong to side chain amide groups of asparagine and glutamine residues. Not all of the crosspeaks belonging to the side chains of asparagine, glutamine, arginine, and tryptophan residues were plotted.

In contrast to the HSQC spectra measured in organic-aqueous media, at similar concentrations a poorly resolved, highly overlapped HSQC spectrum was observed for a ~0.5 mM preparation of E3-M7-24-T40 in 8.5 mM PPG detergent measured at 45 °C for 4 h with detergent/peptide molar ratio of 20:1 (Figure 24A). We chose PPG as a first choice of detergent media because it has been considered by several researchers as one of the best detergents for NMR spectroscopy of peptides because of its very small micelle size (Krueger-Koplin et al., 2004; Vinogradova et al., 1998).

In an attempt to improve spectral resolution of E3-M7-24-T40 in PPG detergent the TROSY-HSQC pulse sequence (Pervushin et al., 1997) was employed. In our case just marginal improvement of HSQC spectra quality was achieved even

with the use of much more experimental time due to the low sensitivity of the TROSY method (Figure 24B). Because of this observation we decided not to attempt additional TROSY-based experiments on samples in detergent.

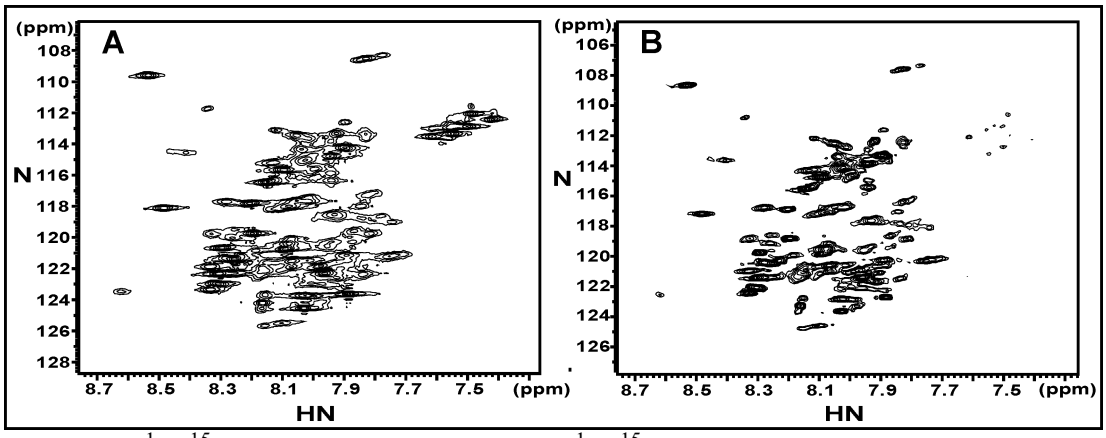

Figure 24. ^{1}H-^{15}N HSQC spectrum (A) and ^{1}H-^{15}N TROSY spectrum (B) of 490 μM [^{15}N]-E3-M7-24-T40 in 8.5 mM PPG micelles with 1:17 peptide to detergent ratio at 45 °C. The spectrum in (A) was measured for 4 h and the spectrum in (B) was measured for 24 h. Both spectra depict low resolution, and slightly improved linewidth of crosspeaks in (B).

Major improvement in resolution was achieved when the detergent excess was increased to 400-500 fold as has been suggested in Opella's work (McDonnell & Opella, 1993). To screen the detergents for NMR experiments a set of HSQC spectra was measured for 4 h using low peptide concentration (~0.2 mM) in ~80-90 mM d_{25}-SDS, d_{38}-DPC, d_{22}-DHPC, or PPG at 50 °C (Figure 25A-D). At this high detergent excess 67, 64, 65, and 62 crosspeaks were clearly resolved for these d_{25}-SDS, d_{22}-DHPC, d_{38}-DPC and PPG samples, respectively. The pattern of HSQC crosspeaks was slightly different in each of the latter detergent preparations. HSQC peaks were about the same width and peak intensity for all four detergents (Figure 25A-D), with DHPC and PPG being judged as the best based on line-width and peak dispersion. HSQC spectra measured in detergents at low temperature (25 °C) resulted in a

significant decline in intensity and broadening of many peaks in the spectrum. Taking into account that running a three-dimensional NMR experiment may take a week, we have tested the stability of our peptide preparations in detergents. As judged by HSQC NMR spectra, preparations of E3-M7-24-T40 in SDS and DPC were found to be more stable with time whereas those measured in PPG and DHPC were less stable (data not shown).

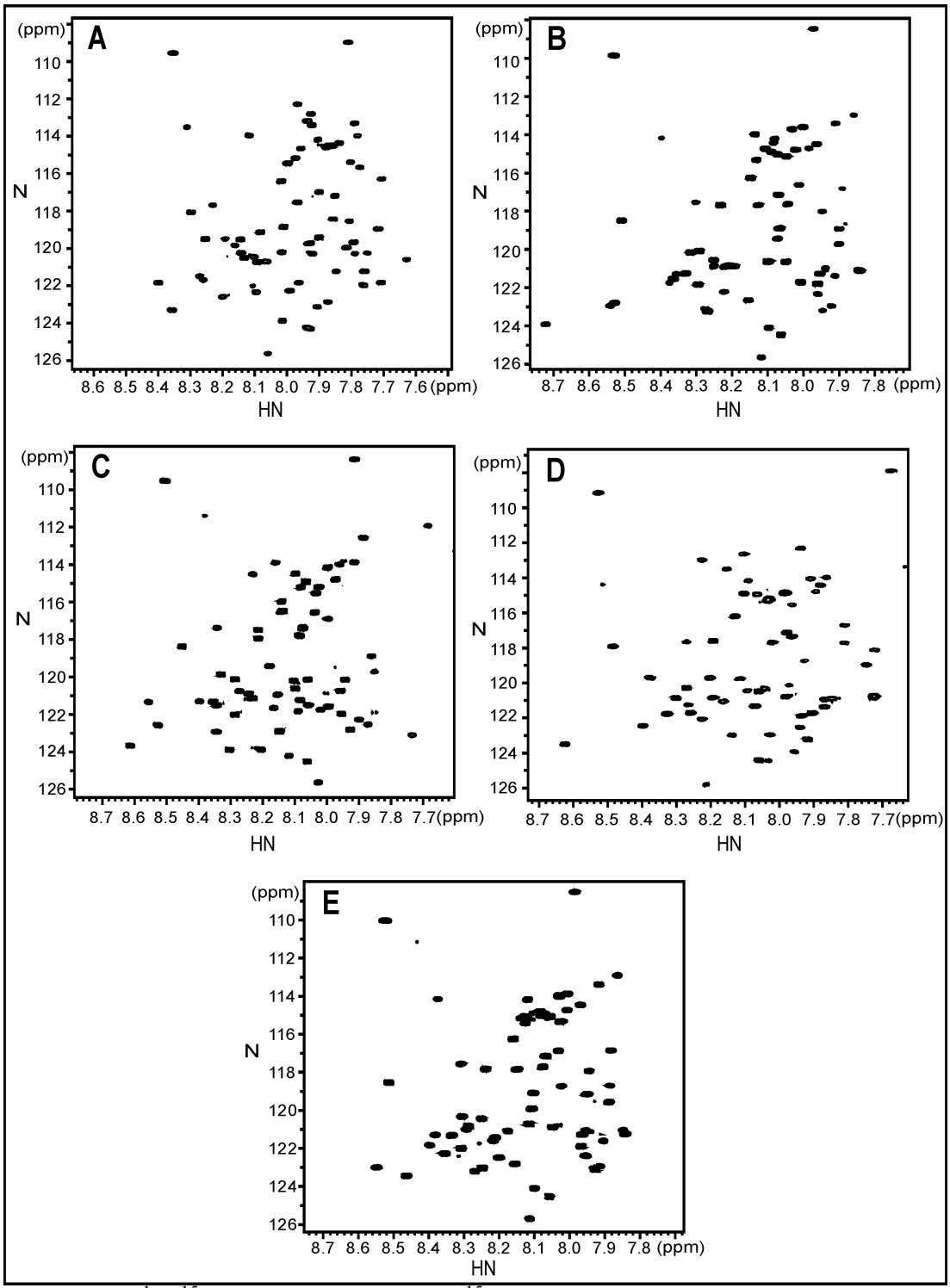

Figure 25. ^{1}H-^{15}N HSQC spectra of [^{15}N]-E3-M7-24-T40 measured for 4 h in detergents at high detergent/peptide ratios (A) d_{25}-SDS micelles, (B) d_{38}-DPC micelles, (C) d_{22}-DHPC micelles and (D) PPG micelles at 50 °C. The peptide to detergent ratio was 1:440 in A, B, & C and 1:500 in D. The peptide concentration was approximately 0.2 mM. (E) 0.5 mM [^{15}N]-E3-M7-24-T40 in 200 mM d_{38}-DPC micelles (1:400 peptide to detergent ratio).

An attempt was made to increase peptide concentration in the detergent preparations to 0.5 mM to make 3D NMR measurements possible using our 600 MHz instrument. DPC was chosen based on its better sample stability in time observed at 0.2 mM peptide concentration and to avoid using SDS, which often denatures proteins and peptides (Vinogradova et al., 1998). The HSQC spectrum obtained for a fresh solution of 0.5 mM 73-residue peptide in 200 mM DPC (Figure 25E) was similar in resolution to that of the 0.2 mM peptide in the same detergent (Figure 25B). However, the 0.5 mM sample was not stable and became turbid and gel-like after one week of running 3D NMR experiments at 50 °C. The HSQC spectrum of this sample lost about 40% of peak intensity, some peaks broadened and extra peaks appeared in the spectrum. Similar instability of other peptides in lipid-like environments also has been reported (Krueger-Koplin et al., 2004). In contrast, the lower concentration sample in DPC was stable for several weeks at the same temperature. Unfortunately, at least 0.5 mM peptide concentration is needed to perform a 3D NMR measurement using our 600 MHz NMR instrument in a reasonable amount of time. Therefore, we decided to commence our studies of the 3D structure of E3-M7-24-T40 in an organic-aqueous membrane mimetic media.

III.2.B.2. *Analysis of E3-M7-24-T40 in organic-aqueous media.* In addition to HSQC spectra, 1D proton NMR spectra of the [^{15}N]-labeled E3-M7-24-T40 peptide in TFE/H_2O (1:1) and $CDCl_3$/CD_3OH/H_2O (4:4:1) at 25 °C (Figure 26) were also recorded as a fast evaluation of instrument resolution and overall quality of sample preparation. Comparing peak width and integral intensity of 1D proton NMR spectra

before and after running a long 3D NMR experiment allows us to monitor sample stability. The E3-M7-24-T40 peptide solution in TFE/H_2O (1:1) was found to be stable for several months and for one month in $CDCl_3/CD_3OH/H_2O$.

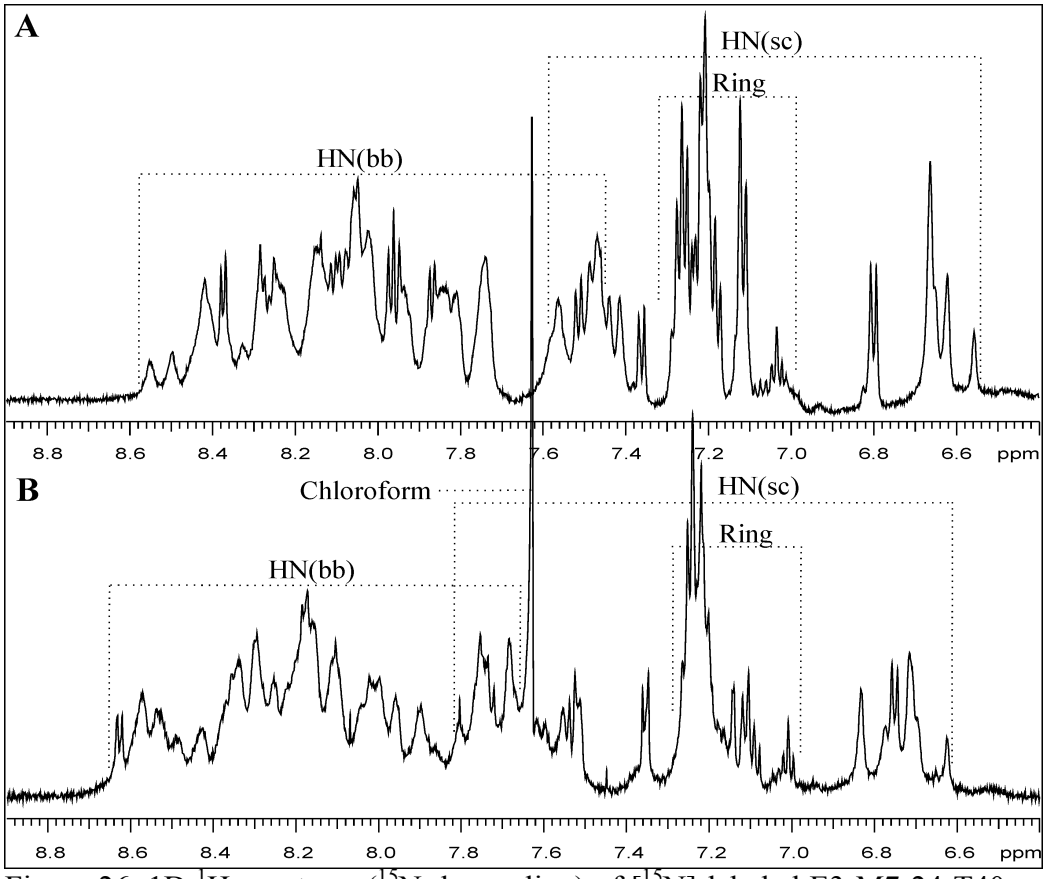

Figure 26. 1D ^{1}H spectrum (^{15}N-decoupling) of [^{15}N]-labeled E3-M7-24-T40 peptide in (A) TFE/H_2O (1:1) and (B) $CDCl_3/CD_3OH/H_2O$ (4:4:1) at 25 °C. The chemical shift ranges of backbone amide protons, side-chain amide protons, and aromatic protons are indicated by HN (bb), HN (sc), and ring, respectively. In the case of $CDCl_3/CD_3OH/H_2O$ (B), the strong peak around 7.63 ppm is from the chloroform solvent.

A set of homonuclear 2D NMR experiments was performed on E3-M7-24-T40 peptide solution in TFE/H_2O (1:1) and $CDCl_3/CD_3OH/H_2O$ (4:4:1) at 25 °C as a preliminary step to evaluate the samples suitability for 3D heteronuclear NMR. Better sensitivity and better digital resolution in F1 (indirect) dimension can be achieved in a

2D NMR experiment and thus may helpful in the assignment of the 3D spectra. Figure 27 depicts a 2D NOESY spectra showing a number of well-resolved and narrow peaks but at the same time several clusters with extreme overlap of peaks. Such overlap prevented unambiguous assignment of chemical shifts in the 73-residue peptide using only 2D experiments and required the use of 3D NOESY experiments.

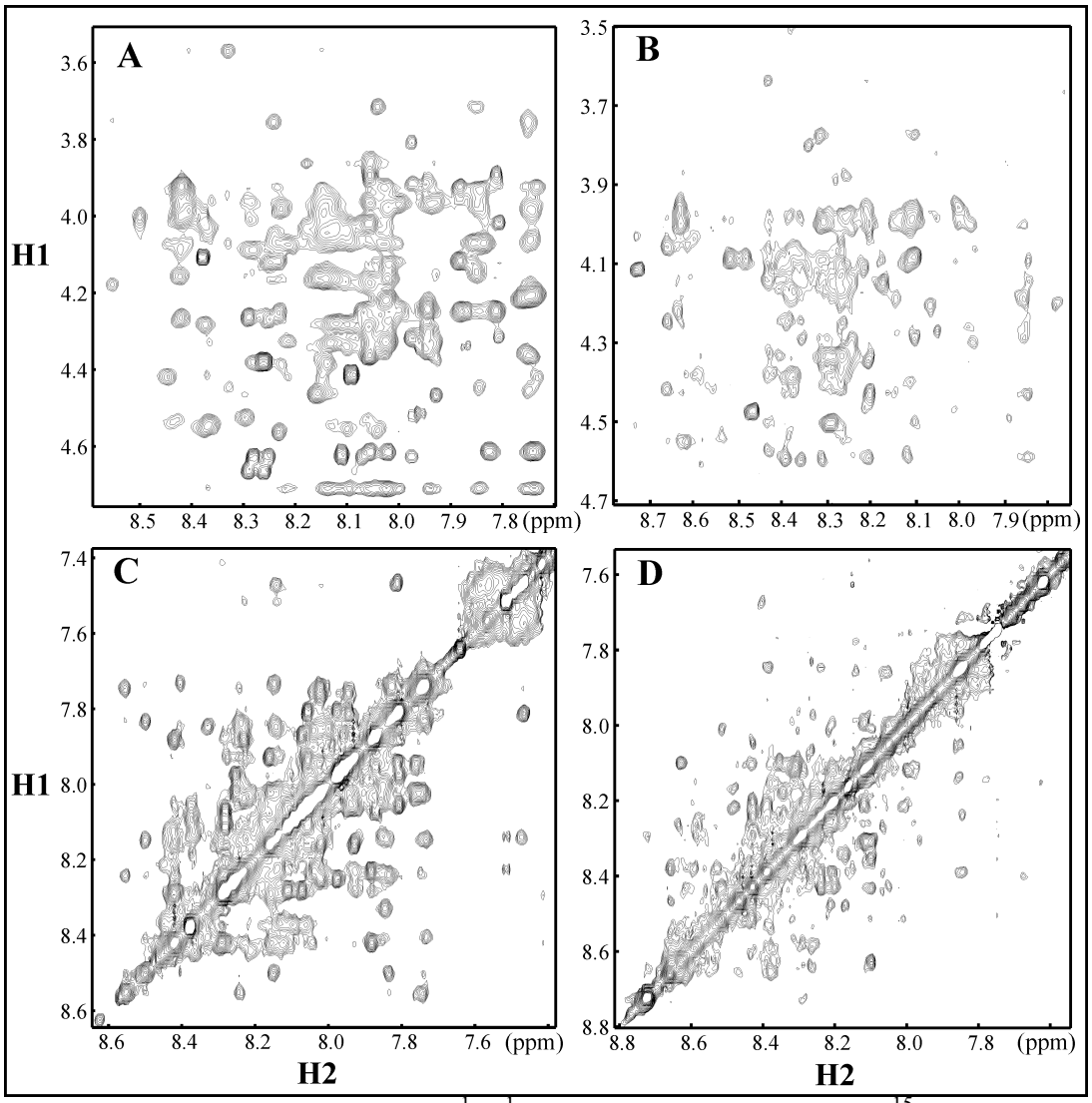

Figure 27. Selected regions of the ^{1}H-^{1}H 2D NOESY spectrum of [^{15}N]-labeled E3-M7-24-T40 peptide in TFE/H_2O (1:1) [A & C] and $CDCl_3$/CD_3OH/H_2O (4:4:1) [B & D] at 25 °C. (A & B) regions show NH to CHα and some CHβ connectivities. (C & D) regions show backbone NH-NH connectivities.

3D ^{15}N-edited experiments, such as TOCSY-HSQC and NOESY-HSQC (Kay et al., 1992; Zhang et al., 1994), which resolve crosspeaks between ^{1}H spins according to the chemical shift of ^{15}N bonded directly to the ^{1}H spins, were used to make complete assignments of backbone and side chain protons of the [^{15}N]-labeled E3-M7-24-T40 peptide in TFE/H_2O (1:1) and $CDCl_3$/CD_3OH/H_2O (4:4:1) at 25 °C. 3D TOCSY-HSQC and 3D NOESY-HSQC spectra give the same kind of information obtained from the connectivity of NH to NH and NH to CHα-δ regions of 2D ^{1}H-^{1}H TOCSY and NOESY spectra, but they are resolved into a third dimension according to the ^{15}N chemical shift associated with the amide ^{1}HN. Identification of specific amino acids in the HSQC spectrum was made using the 3D TOCSY-HSQC spectrum. Sequential assignments of amino acids were determined using 3D NOESY-HSQC and 3D TOCSY-HSQC spectra.

Heteronuclear-edited 3D NMR experiments, data processing, and sequential assignments of signals in 3D spectra of ^{15}N-ubiquitin are presented in Cavanagh et al., 1996. First several unique and rare amino acid residues in the peptide sequence, such as one arginine, one tyrosine, two glutamine, and 3 valine residues, were used as starting points to assemble and assign peptide fragments. Each HSQC peak was assigned to a possible amino acid type by comparing the chemical shifts of crosspeaks in the 3D TOCSY spectrum to the ^{1}H chemical shifts for the 20 common amino acid residues (Wüthrich K, 1986; Schwarzinger et al., 2000).

An example of how assignments of residues (in this case Thr^{41}) and their short- and medium-range connectivities were made in TFE/H_2O (1:1) and $CDCl_3$/CD_3OH/H_2O (4:4:1) is presented in Figure 28. A strip plot of the 3D NOESY

spectrum and a strip plot of the 3D TOCSY spectrum corresponding to one residue (e.g. Thr^{41}) are compared to find crosspeaks with matching chemical shifts and thereby assigning these NOESY crosspeaks to the residue itself [intraresidue connectivity (*i,i*)]. Short- and medium-range NOEs result from the correlation through space between protons of amino acid residues that are close in space due to the presence of a specific conformation. The short- and medium-range connectivities in the already assigned residue (e.g. Thr^{41}) were also assigned by first, searching through all 3D TOCSY strips of all amino acids for the best matches of crosspeaks with the ones in the 3D NOESY of already assigned residue (e.g. Thr^{41}). Then the 3D NOESY strips of the amino acid residues with the best matching crosspeaks are compared back to the 3D TOCSY strip of the already assigned residue (e.g. Thr^{41}) to find whether the 3D NOESY strip of the amino acid residue in question have short- and medium-range connectivities with the already assigned residue (e.g. Thr^{41}). The short- and medium-range connectivities in the 3D NOESY of the already assigned residue (e.g. Thr^{41}) are assigned with the most intense crosspeaks assigned to the previous (*i-1*) or next (*i+1*) amino acid in the sequence. All residues were sequentially assigned in a similar manner by mostly walking downstream (*i-1*), one residue at a time, of the amino acid sequence.

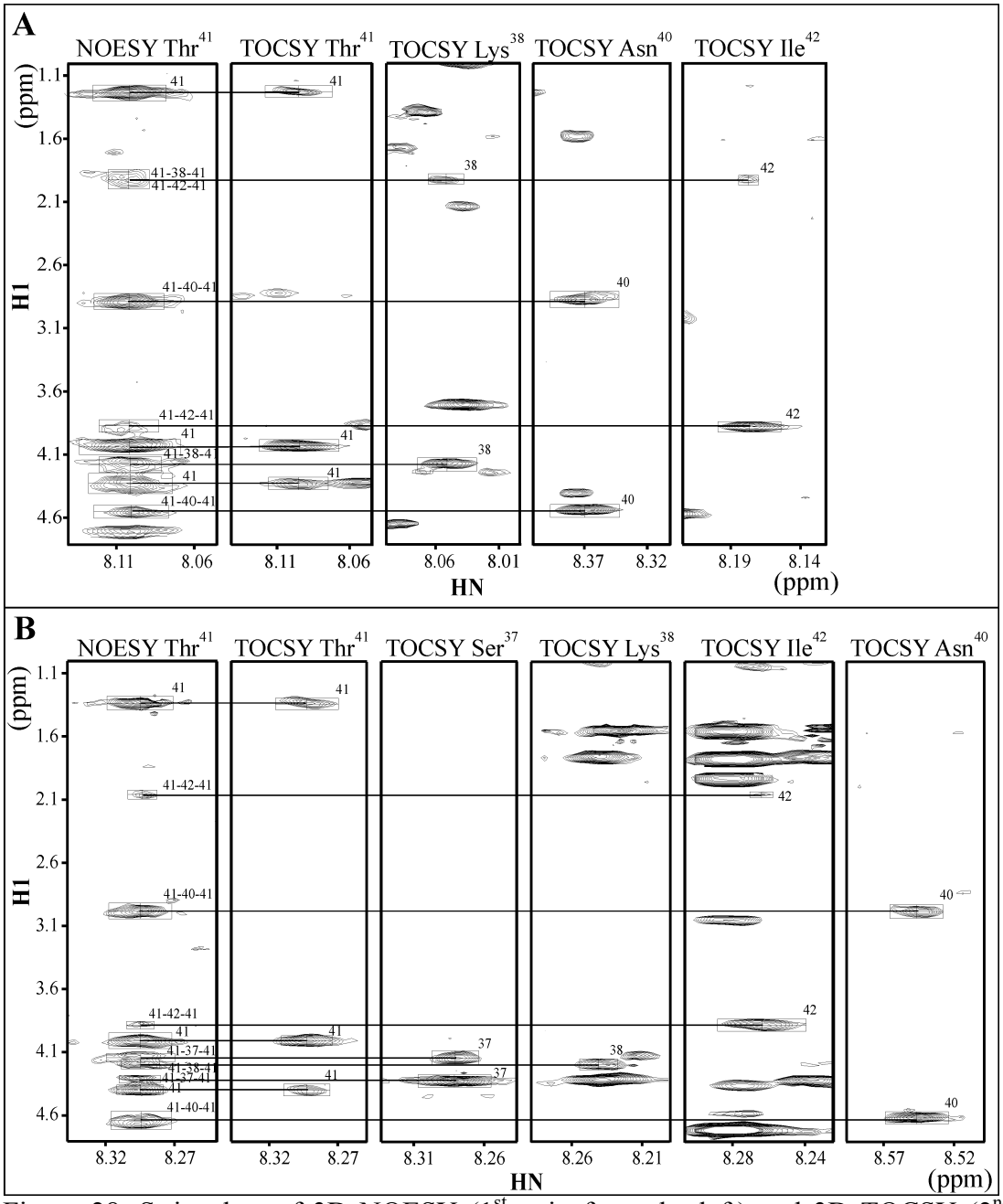

Figure 28. Strip plots of 3D NOESY (1st strip from the left) and 3D TOCSY (2nd strip) of Thr^{41} and 3D TOCSY of nearby residues (all other strips) showing short- and mediun-range connectivities with Thr^{41} of E3-M7-24-T40 in (A) TFE/H_2O (1:1) and (B) $CDCl_3/CD_3OH/H_2O$ (4:4:1). All unnumbered crosspeaks belong to residues other than the ones marked above.

Two dimensional TOCSY spectra (Figure 29) were also recorded and were used to confirm proline chemical shifts which were hypothesized based on analysis of

interresidue crosspeaks in ^{15}N NOESY-HSQC spectra as the chemical shifts of CHα and ring protons of proline residues cannot be directly observed in a 3D ^{15}N–edited TOCSY-HSQC experiment. Thus by using 3D TOCSY- and NOESY-HSQC and 2D TOCSY spectra and the assignment process mentioned above the backbone correlation crosspeaks in ^{1}H-^{15}N HSQC spectra of [^{15}N]-E3-M7-24-T40 peptide in TFE/H_2O (1:1) and $CDCl_3$/CD_3OH/H_2O (4:4:1) were finally assigned to each amino acid in the peptide sequence (Figure 30). Each assigned crosspeak is represented by an amino acid number, and the unmarked peaks in the upper right corner of the HSQC spectra belong to side chain amide group of glutamine and asparagine residues. A complete list of all of the chemical shifts of [^{15}N]-E3-M7-24-T40 in TFE/H_2O and $CDCl_3$/CD_3OH/H_2O is presented in the Appendix.

Comparing the nitrogen chemical shifts [^{15}N] as well as amide NH chemical shifts for the E3-M7-24-T40 peptide in both organic-aqueous media reveals that the chemical shifts in TFE/H_2O (1:1) were very similar to the chemical shifts obtained in $CDCl_3$/CD_3OH/H_2O [4:4:1] (Figure 31).

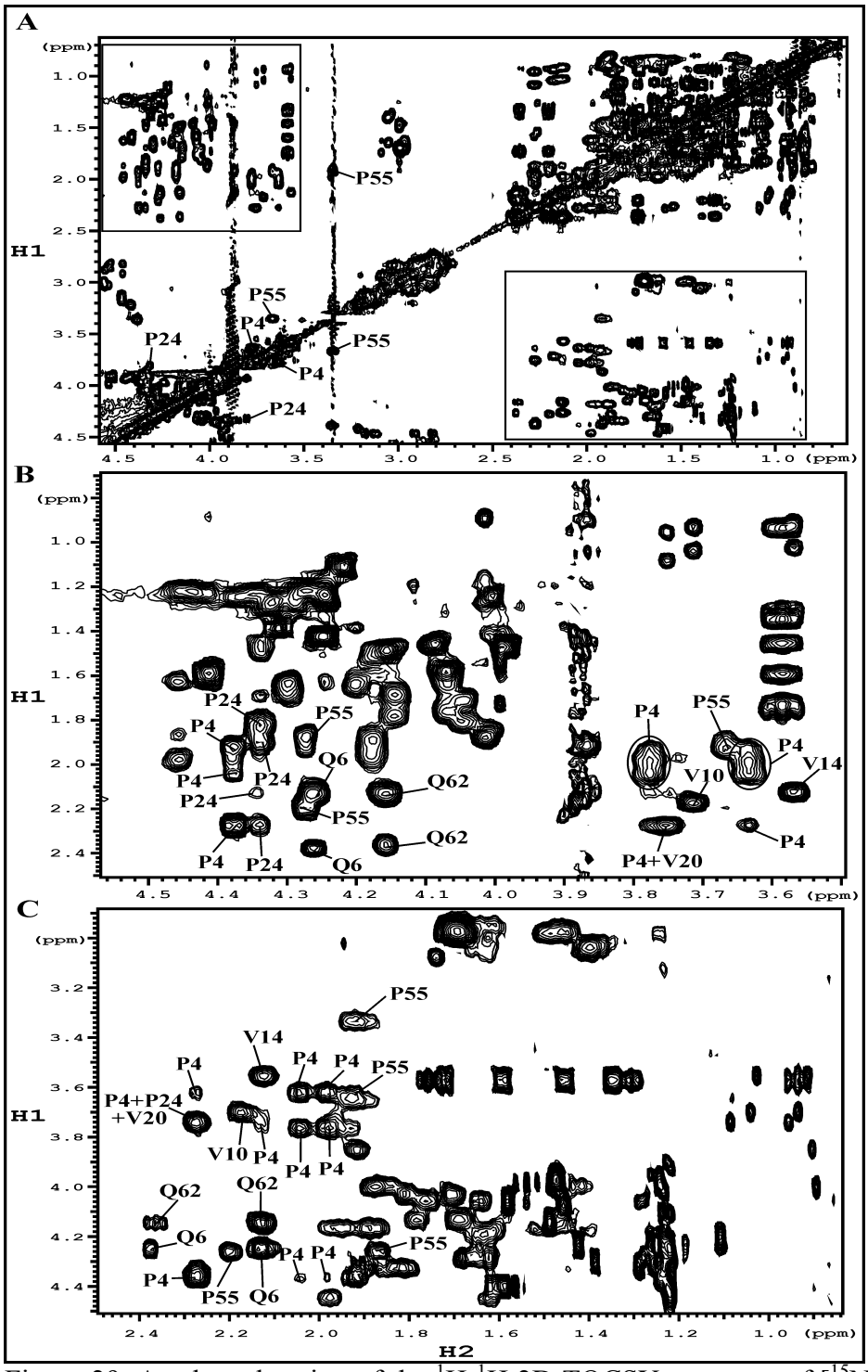

Figure 29. A selected region of the ^{1}H-^{1}H 2D TOCSY spectrum of [^{15}N]-labeled E3-M7-24-T40 peptide in TFE/H_2O (1:1) at 25 °C (A) showing connectivities between α, β, γ, and δ protons. (B) and (C) are enlarged regions of the left and right boxed inserts, respectively, in (A) and show crosspeaks corresponding to protons of proline residues 4, 24, and 55 and crosspeaks that are in the same region of the spectrum as the proline peaks, such as valine residues 10, 14, and 20 and glutamine residues 6 and 62.

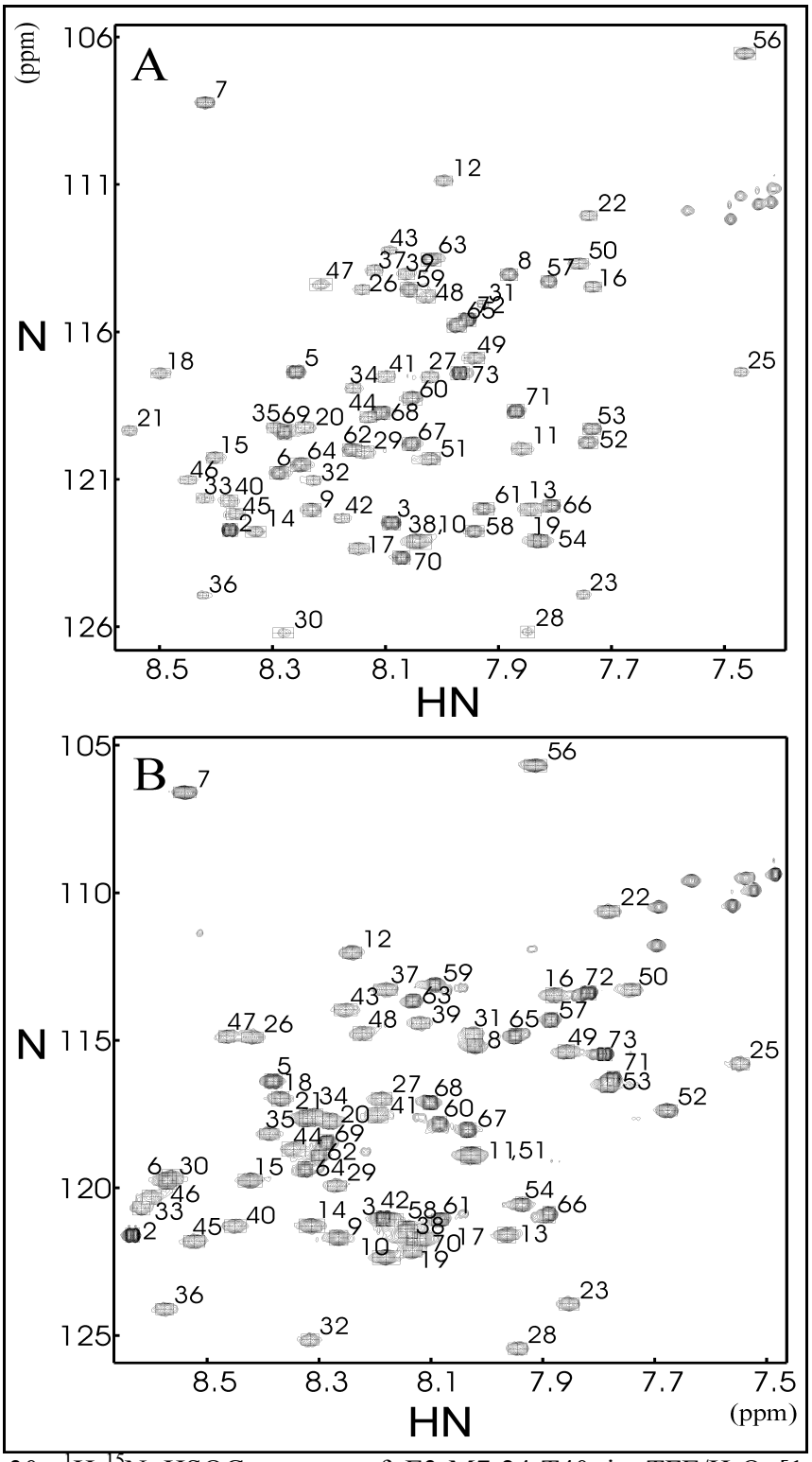

Figure 30. ^{1}H-^{15}N HSQC spectra of E3-M7-24-T40 in TFE/H_2O [1:1] (A) and $CDCl_3$/CD_3OH/H_2O [4:4:1] (B). These spectra are cropped spectra depicting the assigned backbone correlation peaks to each amino acid in the peptide sequence, represented by an amino acid number, and unmarked crosspeaks belonging to side chain amide group of glutamine and asparagine residues in the upper right corner.

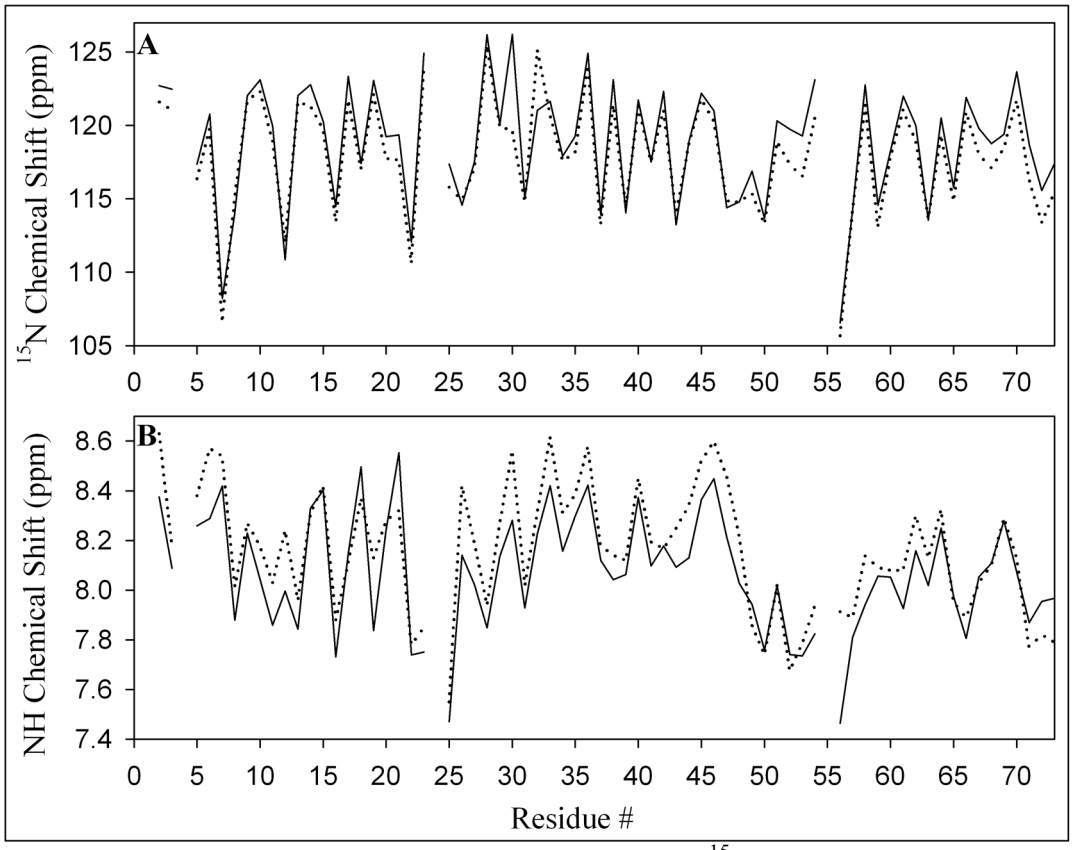

Figure 31. Comparison of nitrogen chemical shifts [^{15}N] and amide NH chemical shifts for the E3-M7-24-T40 peptide at 25 °C in both TFE/H_2O [1:1] (represented by a solid line) and $CDCl_3$/CD_3OH/H_2O [4:4:1] (represented by a dotted line). (A) ^{15}N chemical shifts and (B) amide NH chemical shifts.

The δCHα chemical shifts for E3-M7-24-T40 peptide in organic-aqueous media were compared to the δCHα random coil values (Wishart et al., 1992; Wishart & Sykes, 1994) and plotted versus position in the peptide (Figure 32A). A chemical shift index (CSI) value of +1 and –1 is assigned for deviation of ≥+0.1 ppm and ≤–0.1 ppm from the random coil value, respectively, and a CSI value of zero is assigned for deviation between +0.1 and –0.1 ppm from the random coil value. Using a minimum of three consecutive CSI values of +1 to define a β-strand, and a minimum of four (not necessarily consecutive) CSI values of –1 to define a helix as the criterion for

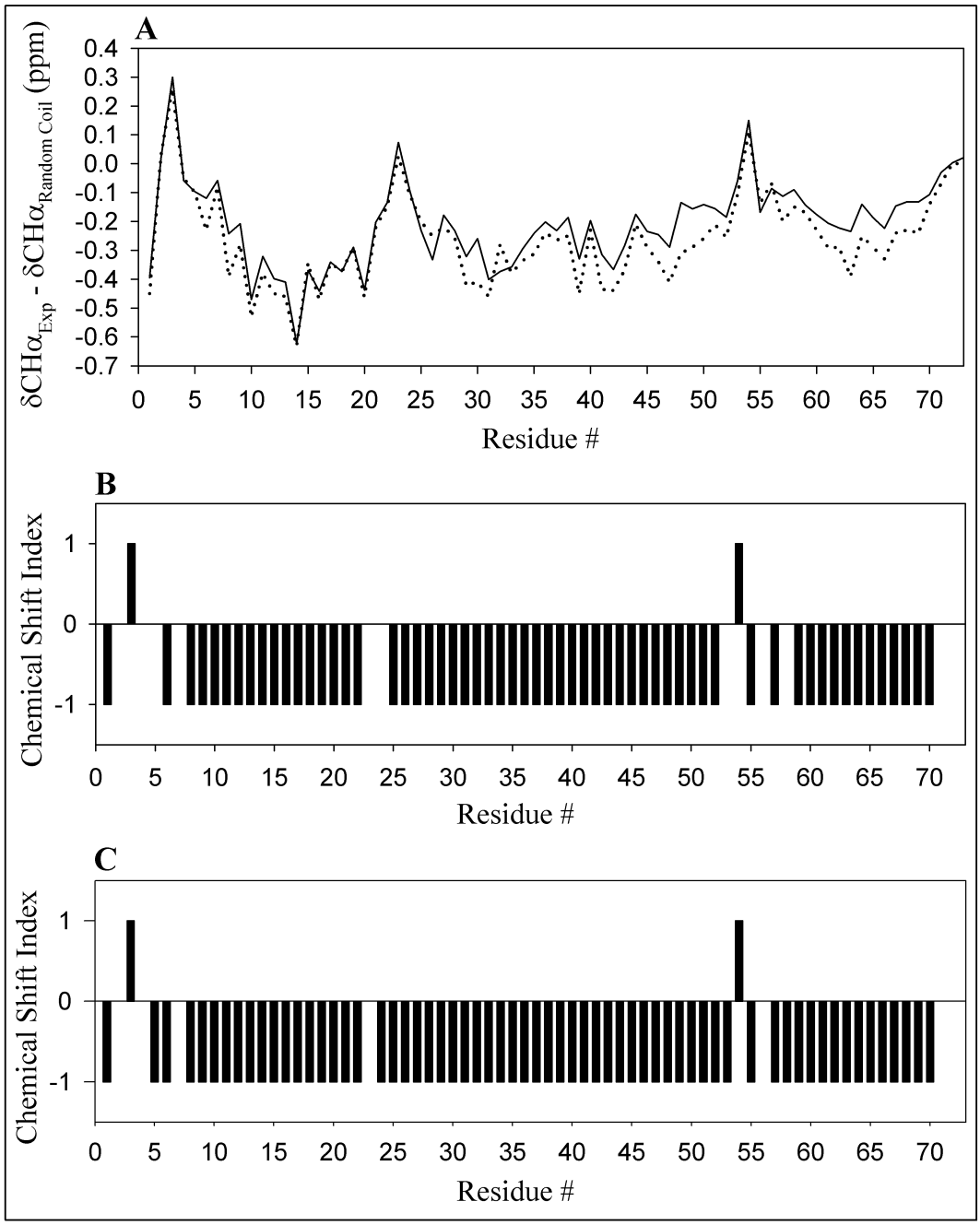

Figure 32. Comparison of chemical shift index (CSI) and δCHα values for the E3-M7-24-T40 peptide at 25 °C in both TFE/H_2O [1:1] and $CDCl_3$/CD_3OH/H_2O [4:4:1]. (A) The δCHα values plotted for the peptide in TFE/H_2O [1:1] (represented by a solid line) and $CDCl_3$/CD_3OH/H_2O [4:4:1] (represented by a dotted line) are the difference between the experimental CHα chemical shifts and the random coil values as per reference (Wishart & Sykes, 1994). (B) Plot of the CSI values of the peptide in TFE/H_2O [1:1]. (C) Plot of the CSI values of the peptide in $CDCl_3$/CD_3OH/H_2O [4:4:1]. The CSI values plotted are based on the δCHα values where a δCHα value of ≥0.1 is given a CSI value of 1.0, a δCHα value of ≤ –0.1 is given a CSI value of –1.0, and a δCHα value between 0.1 and –0.1 is given a CSI value of 0.

possibility of regular secondary structure (Wishart et al., 1992), residues 8-22, 25-52 and 59-70 have helical tendencies in both solvents (Figure 32B,C). If one considers the magnitude of the deviation as indicative of the stability of the helix, domain 8-22 forms the most stable helix followed by less stable 25-52 and 59-70. The three helical regions were interrupted by three short regions, each containing a proline, having δCHα values between 0.1 and –0.1, characteristic of a random structure. The disruption in the helical regions might be caused by the proline residues since prolines have been shown to disrupt helices (Chou & Fasman, 1974; Richardson & Richardson, 1988).

In both $CDCl_3/CD_3OH/H_2O$ and TFE/H_2O solvents, sequential and medium-range interproton distances were derived form NOESY data sets and confirmed the presence of 3 helical elements in E3-M7-24-T40 peptide as determined by the medium-range CHα-NH [$d_{\alpha N}(i,i+3)$] connectivity. Most of the residues in the peptide exhibited sequential NH-NH [$d_{NN}(i,i+1)$] crosspeaks with two notable breaks at residues 55-57 and 59-61 in the case of $CDCl_3/CD_3OH/H_2O$ (Figure 33A) and a notable break at residues 22-24 in the case of TFE/H_2O (Figure 33B). The medium range NOE connectivity pattern in $CDCl_3/CD_3OH/H_2O$ (Figure 31A) was similar to the one observed in TFE/H_2O (Figure 31B). Blocks of medium-range CHα-NH [$d_{\alpha N}(i,i+3)$] crosspeaks were observed for residues 9 through 19, 29-49 and 57-70 in $CDCl_3/CD_3OH/H_2O$ (Figure 33A) and for residues 7-19, 38-52 and 60-72 in TFE/H_2O (Figure 33B). These regions also exhibited numerous $d_{\alpha N}(i,i+2)$ and d_{NN} $(i,i+2)$ crosspeaks. Several $d_{\alpha N}(i,i+4)$ could be detected for residues in these same regions of E3-M7-24-T40. The gaps found in the helical regions are mostly due to

overlap in crosspeaks, which prevented the assignment of the long-range connectivity.

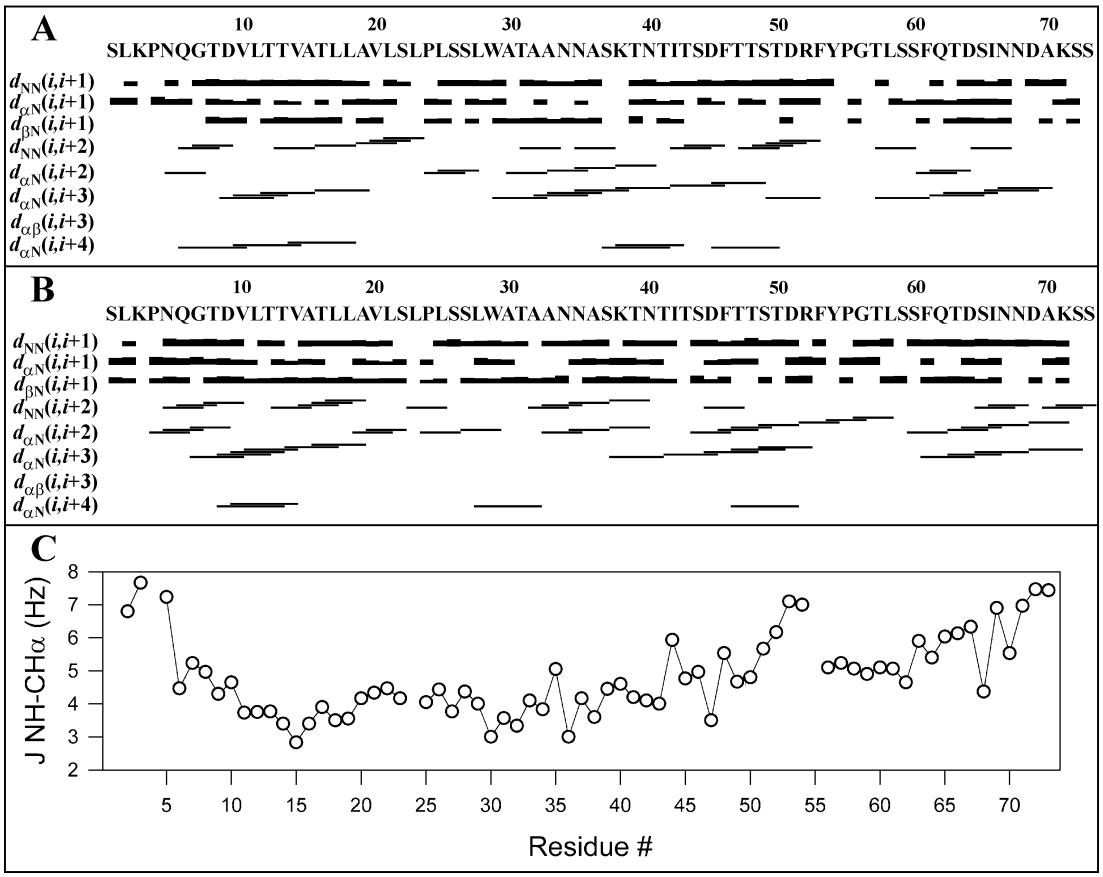

Figure 33. Summary of connectivities and coupling constants for E3-M7-24-T40 in TFE/H_2O (1:1) and $CDCl_3$/CD_3OH/H_2O (4:4:1). (A) Sequential and medium-range NOEs in $CDCl_3$/CD_3OH/H_2O (4:4:1). (B) Sequential and medium-range NOEs in TFE/H_2O (1:1). (C) J coupling values of NH to CHα ($^3J_{HN\alpha}$) in TFE/H_2O (1:1).

The relatively sharp peaks in the HSQC spectra of E3-M7-24-T40 in TFE/H_2O (1:1) and a 1.5 Hz/point digital resolution allowed the estimation of the J-coupling constants ($^3J_{HN\alpha}$) by measuring the distance in split peaks from HSQC spectra obtained at 25 °C, 35 °C, and 45 °C. For singlets estimation of maximum coupling constants was made by measuring peak width. Low values of coupling constants ($^3J_{HN\alpha}$) from 3 to 5 Hz were measured for 6-23 and 25-43 residues and most

residues between 56 and 65 had $^{3}J_{HN\alpha}$ values below 6Hz in TFE/H_2O (Figure 33C). The NOE connectivities and low coupling constants support the presence of three helical segments in the 73-residue peptide. This conclusion is in a good agreement with the NOE long-range connectivities and the chemical shift index analysis. As we proceeded toward the carboxyl terminus the $^{3}J_{HN\alpha}$ values rose to near 7Hz (residues 63 and higher). In addition, the loop region (residues 1-7) also showed higher $^{3}J_{HN\alpha}$ values (Figure 33C). In contrast, we were unable to measure J-coupling constants of E3-M7-24-T40 peptide in $CDCl_3$/CD_3OH/H_2O (4:4:1) because the linewidth of the HSQC spectra were not as good as the one in TFE/H_2O (1:1).

The distribution of distance constraints (total number of NOE constraints and NOE constraints per residue) for the E3-M7-24-T40 peptide in TFE/H_2O (1:1) and $CDCl_3$/CD_3OH/H_2O (4:4:1) were plotted as a function of their range and peptide sequence (Figure 34). As depicted in the plots of the total number of NOE constraints as a function of their range, a maximum long-range connectivity of $i,i\pm4$ and $i,i\pm5$ was obtained in TFE/H_2O (1:1) and $CDCl_3$/CD_3OH/H_2O (4:4:1), respectively (Figure 34A,B). The plots of the NOE constraints as a function of the peptide sequence depict intraresidue connectivity (i,i) represented by white bars, sequential connectivity ($i,i\pm1$) represented by grey bars, and medium- to long-range connectivity ($i,i\pm \geq 2$) represented by black bars (Figure 34C,D).

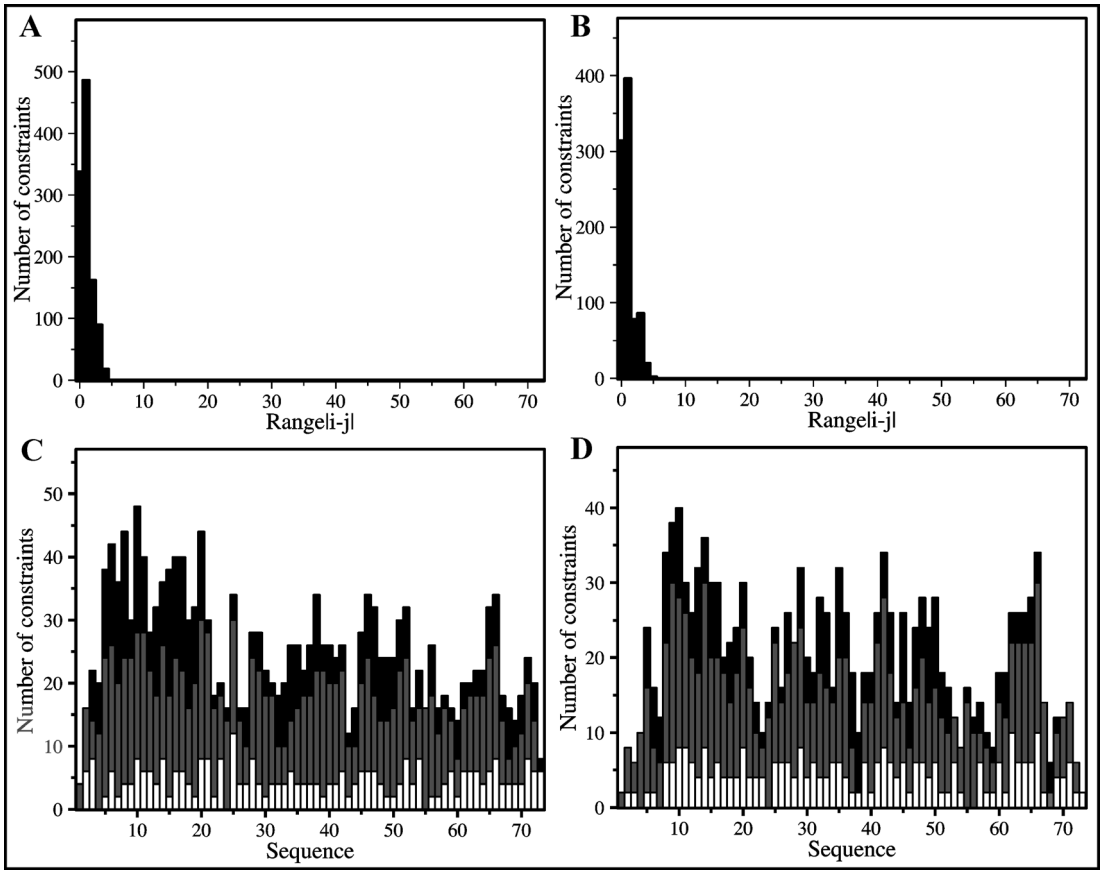

Figure 34. Distribution of distance constraints for E3-M7-24-T40 peptide in TFE/H_2O (1:1) and $CDCl_3$/CD_3OH/H_2O (4:4:1) as a function of their range and peptide sequence. (A) Total number of NOE constraints as a function of their range in TFE/H_2O (1:1). (B) Total number of NOE constraints as a function of their range in $CDCl_3$/CD_3OH/H_2O (4:4:1). (C) NOE constraints as a function of peptide sequence in TFE/H_2O (1:1). (D) NOE constraints as a function of peptide sequence in $CDCl_3$/CD_3OH/H_2O (4:4:1). In (C & D) white bars represent intraresidue connectivity (*i*,*i*), grey bars represent sequential connectivity (*i*,*i*±*1*), and black bars represent medium- to long-range connectivity (*i*,*i*± ≥*2*).

III.2.B.3. *Determination of the NMR Structure of E3-M7-24-T40 in TFE/H_2O and $CDCl_3$/CD_3OH/H_2O.* Molecular modeling was performed using the DYANA program (Güntert et al., 1997). A model of the structure of E3-M7-24-T40 in TFE/H_2O (1:1) [Figure 35] and $CDCl_3$/CD_3OH/H_2O (4:4:1) [Figure 36] at 25 °C was determined using a combination of NOE and dihedral angle constraints following the procedures described in the Materials and Methods section. A total of 547 NOE

distance constraints (169-intraresidue, 243 sequential and 135 medium- to long-range) and 60 Φ dihedral angle constraints were input into the structure calculation for the peptide in TFE/H_2O (Table 9). Modeling of the peptide structure in $CDCl_3/CD_3OH/H_2O$ was done using only NOE constraints. A total of 488 NOE distance constraints (157 intraresidues, 237 sequential, and 94 medium- to long-range) were input into the calculation for the peptide in $CDCl_3/CD_3OH/H_2O$ (Table 9). The dihedral angle constraints were not obtained and used in the calculation of the E3-M7-24-T40 in $CDCl_3/CD_3OH/H_2O$ (4:4:1) due to the lack of $^3J_{HN\alpha}$ constants, which were not measured in the case of $CDCl_3/CD_3OH/H_2O$ (4:4:1) because of the low number of split peaks in the HSQC spectrum.

Table 9

Number of NOE constraints that were input into the structure calculations performed by the DYANA program[a]

NOE connectivity	TFE/H_2O (1:1)	$CDCl_3/CD_3OH/H_2O$ (4:4:1)
Intraresidue (*i,i*)	169	157
Sequential (*i,i±1*)	243	237
(*i,i±2*)	82	40
(*i,i±3*)	44	43
(*i,i±4*)	9	10
(*i,i±5*)	_____	1
Sum of NOEs	547	488

[a] DYANA program (Güntert et al., 1997).

The model of E3-M7-24-T40 in both aqueous-organic solvents indicated 3 helical domains comprising residues 5-24, 30-52, and 55-70. When the 10 structures (20 in the case of $CDCl_3/CD_3OH/H_2O$) with the minimum NOE distance violations

were overlaid separately for each of these helical domains, region I (residues 5-24) had a backbone RMSD of 1.9 Å in TFE/H_2O and 1.71 Å in $CDCl_3$/CD_3OH/H_2O, region II (residues 30-52) had a backbone RMSD of 1.86 Å in TFE/H_2O and 4.32 Å in $CDCl_3$/CD_3OH/H_2O, and region III (residues 55-70) had a backbone RMSD of 1.5 Å in TFE/H_2O and 3.14 Å in $CDCl_3$/CD_3OH/H_2O (Figure 35, 36). The backbone RMSD calculated for the entire peptide was very large (~ 10 Å in TFE/H_2O

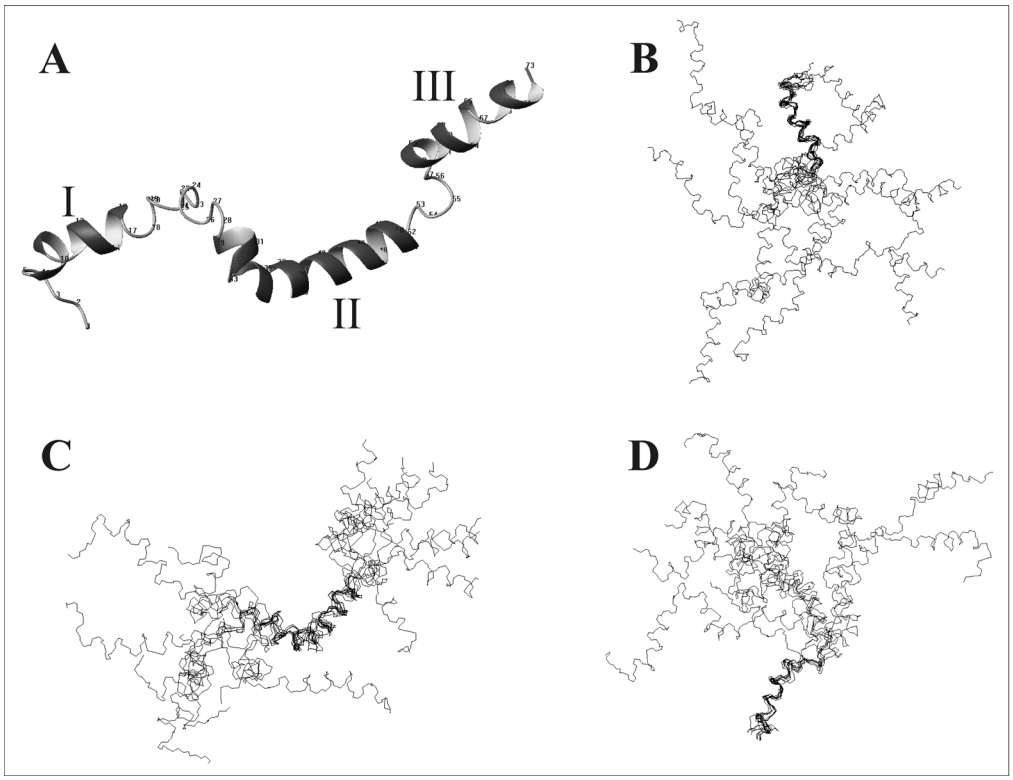

Figure 35. NMR based molecular model of E3-M7-24-T40 in TFE/H_2O (1:1), 25 °C. (A) Representative NMR-derived model of the peptide indicating 3 helical domains. (B) Overlay of region I (residues 5-24) of 10 minimum NOE distance violation structures with backbone RMSD=1.9 Å. (C) Overlay of region II (residues 30-52) with backbone RMSD=1.86 Å. (D) Overlay of region III (residues 55-70) with backbone RMSD=1.5 Å.

and ~ 11 Å in $CDCl_3$/CD_3OH/H_2O). The model obtained for the structure of E3-M7-24-T40 dissolved in $CDCl_3$/CD_3OH/H_2O (Figure 36) exhibited features similar to those found in TFE/H_2O (Figure 35 having helical residues near the N-terminus and

in the cytosolic tail. However, in general, the cytoplasmic section of the peptide was less ordered and the RMSD values for the two helical regions (30-52; 55-70) were relatively large (>4 Å and >3 Å, respectively). The low resolution structures obtained for E3-M7-24-T40 peptide in TFE/H_2O (1:1) and especially in $CDCl_3/CD_3OH/H_2O$ (4:4:1), depicting certain regions of the peptide as α-helix, could be due to the relatively low number of distance constraints used for the calculations, which are necessary to get a more refined structure. In the case of TFE/H_2O (1:1) if one

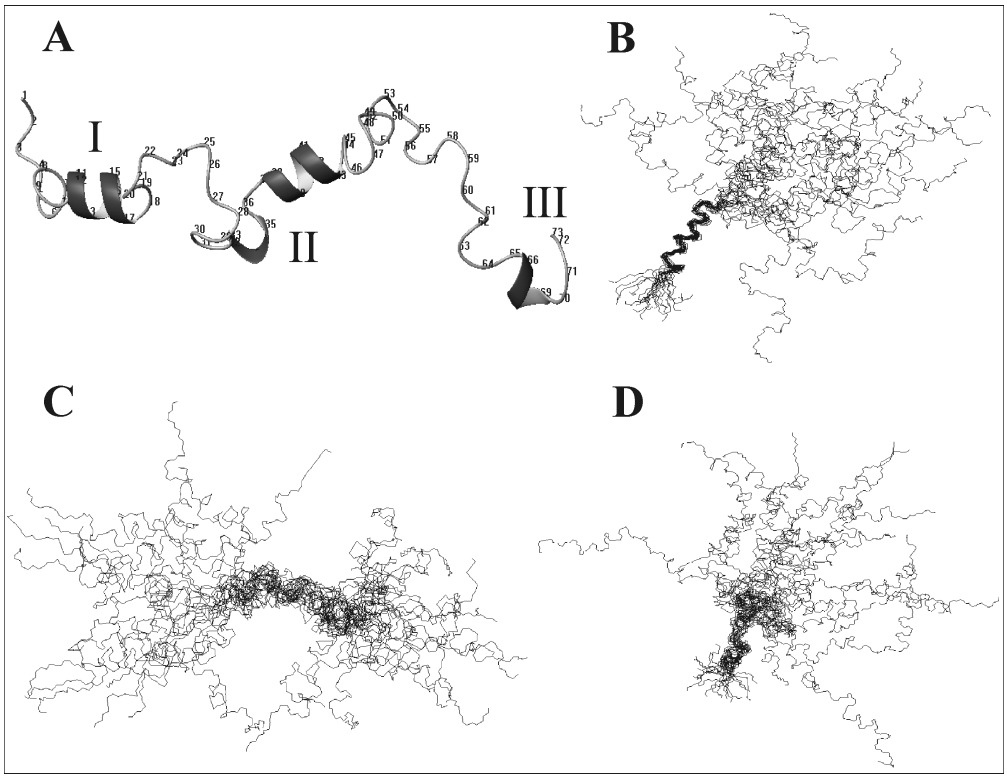

Figure 36. NMR based molecular model of E3-M7-24-T40 in $CDCl_3/CD_3OH/H_2O$ (4:4:1), 25 °C. (A) Representative NMR-derived model of the peptide indicating 3 helical domains. (B) Overlay of region I (residues 5-24) of 20 minimum NOE distance violation structures with backbone RMSD=1.71 Å. (C) Overlay of region II (residues 30-52) with backbone RMSD=4.32 Å. (D) Overlay of region III (residues 55-70) with backbone RMSD=3.14 Å.

considers 61 residues of the peptide, excluding both ends (residues 1-7 and 69-73), only 14 residues have (*i, i+3*) connectivities, a characteristic of an α-helix, 6 residues

have no (*i, i+3*) connectivities, 38 residues may have (*i, i+3*) connectivities but we were unable to confirm their presence and thus make definitive assignments because their positions are overlapping with supposedly stronger crosspeaks belonging to atoms of a residue itself (*i, i*) or to atoms of adjacent (*i, i-1*) or (*i, i+1*) residues. Similar difficulties were encountered in $CDCl_3/CD_3OH/H_2O$ (4:4:1).

Chapter IV: Discussion

IV. 1. Expression, isolation and purification of fusion proteins and their cleaved peptides corresponding to domains of Ste2p

The determination of high-resolution structures of membrane proteins remains an elusive goal. One approach to realize this goal utilizes fragments corresponding to different regions of a protein combined with spectroscopic analysis in membrane mimetic solvents. It is possible to question the use of peptide fragments to learn about the structural tendencies of an intact protein. Nevertheless, extensive analysis has indicated that localized interactions often determine the secondary structure of domains of proteins. Moreover, in the case of bacteriorhodopsin (Katragadda et al., 2001a) and rhodopsin (Yeagle et al., 2001) the structures of peptides corresponding to single transmembrane domains and loops were superimposed to assemble the entire protein and the good agreement obtained with crystal structures of the intact proteins would seem to validate the approach. Finally, both *in vitro* and *in vivo* reconstitution studies show that biologically inactive fragments of a GPCR reassemble to form biologically active receptor (Martin et al., 1999; Ridge et al., 1995; Marti, 1998). The above results suggest that, although one must be cautious in making absolute conclusions, it is reasonable to assume that structural information on a single domain peptide fragment (M6) and a multidomain peptide fragment (E3-M7-24-T40) corresponding to residues of Ste2p will be relevant to the intact GPCR.

Although CD has been extensively used in evaluations of membrane peptide structure, NMR is the preferred method to obtain high-resolution information at the

residue level. To carry out such NMR investigations on larger fragments of membrane peptides, especially in the presence of detergents, efficient methods to obtain isotopically labeled peptides are required and conditions must be optimized for measurement of well-resolved NMR spectra under membrane mimetic conditions. Isotopically labeled peptides can be easily synthesized by expressing plasmids coding for the peptides in an isotopically labeled minimal media using a bacterial expression system. To [^{15}N]-label the peptides [^{15}N]-labeled ammonium chloride ($^{15}NH_4Cl$) was used to introduce [^{15}N] into the fusion proteins. However, when the peptide is dissolved in detergent the increase in peak broadening and overlap in the NMR spectrum obscure its analysis. In order to overcome this problem and be able to assign the NMR spectrum in detergent [^{13}C/^{15}N] doubly labeled peptide was also biosynthesized using a similar method to the one used to biosynthesize [^{15}N]-labeled peptides except with the addition of [^{13}C]-uniformly labeled glucose.

In this thesis we report the successful biosynthesis of unlabeled and isotopically labeled fusion proteins containing a 33-residue peptide and a 73-residue multidomain peptide of Ste2p, a GPCR, with fair to good expression levels in the case of M6 fusion proteins and excellent expression levels in the case of M7 fusion protein (Figure 5). Isolation of 100 mg quantities of unlabeled and isotopically labeled fusion proteins with more than 95% homogeneity was accomplished (Figure 9). These fusion proteins were processed by CNBr to release the final receptor domains (Figure 12-15), which were isolated by reversed phase chromatography (Figure 16,17). We were able to isolate peptides corresponding primarily to one transmembrane domain [M6 and M6(P258L)] and corresponding to part of the cytosolic tail, a

transmembrane domain and an extracellular loop (E3-M7-24-T40) in 10 mg quantities. The 33-residue single transmembrane domain peptides were labeled with [^{15}N] whereas the 73-residue multidomain peptide was labeled with either [^{15}N] or [^{13}C/^{15}N]. Such quantities are more than sufficient for a complete structural analysis using NMR spectroscopy.

The biosynthesis of membrane peptides is complicated by their poor water solubility, tendency to aggregate, and toxicity to the host cell. Few reports have appeared for preparation of membrane peptides containing more than 30 to 40 residues. In one elegant study, peptides corresponding to two predicted transmembrane domains of the cystic fibrosis transmembrane conductance regulator were expressed, released from the thioredoxin carrier, and purified in milligram quantities (Therien et al., 2002). In this thesis report we compared the expression levels of E3-M7-24-T40 as driven by plasmids coding for TrpΔLE or thioredoxin fusion proteins. A plasmid coding for the TrpΔLE carrier was originally used to express bovine pancreatic trypsin inhibitor and very recently was used to express a single transmembrane domain of the CB2 receptor (Staley and Kim, 1994; Zheng et al., 2005). We observed that the TrpΔLE system gave higher expression of Ste2p fusions than the thioredoxin system. This was true when three different thioredoxin constructs were prepared, one coding for additional residues at the C-terminus and the other engineered to code for the identical peptide fused to TrpΔLE (Table 5). Both the level of expression and the percentage of the target protein appeared higher using the TrpΔLE vector (pREJ02M) as compared with the thioredoxin vectors (pKLS01 and its variants). The expression of the presumed target protein by the thioredoxin system

virtually disappeared in minimal medium whereas in minimal medium good to excellent expression of the Ste2p fusions were obtained using the pREJ vectors (Figure 7). Based on the above expression results the TrpΔLE expression system were chosen to express the M6 and E3-M7-24-T40 peptides.

An expression system using GB-1 fusion proteins in *E. coli* has also been used to biosynthesize a 26-residue peptide predicted to be a transmembrane region of the human sodium proton exchanger, and loss of expression of GB-1 fusion protein constructs was also noted in minimal medium (Lindhout et al., 2003). This difficulty was overcome by adding about 5% isotopically labeled rich medium to the culture. Our experiments indicate that some optimization will be required in determining which system to choose for the expression of membrane peptides and that the thioredoxin system cannot be considered to be a general vector for all membrane peptides as suggested previously (Therien et al., 2002). Since the expression levels of E3-M7-24-T40 fusions were consistently higher than those for the M6 fusions (Figure 5) it is likely that no one expression system will prove optimal for all membrane peptides.

The release of the membrane peptides from the TrpΔLE fusion proteins using CNBr was reasonably efficient. Interestingly, different conditions proved optimal for cleaving a single TM domain fusion peptide (Arevalo et al., 2003) and the multidomain peptide we examined. As judged by HPLC monitoring, maximal yields of E3-M7-24-T40 were obtained after 3.5-4.5 h of cleavage using 500 molar excess of CNBr (Figure 15) whereas release of the single transmembrane (M6) domain peptides required 24 h and the use of 5000 molar excess of CNBr (Figure 12,13).

When the cleavage of E3-M7-24-T40 proceeded for 24 h, very little product was isolated. In contrast when the fusion protein containing the M6 domain was cleaved for only 3 h, most of the fusion protein remained unprocessed. Thus, it is advisable to optimize the CNBr cleavage conditions when one seeks to produce isotopically labeled products. We found that the percent acid (50% or 70%) used for the CNBr cleavage had a small influence on the cleavage rate but did not affect the ultimate yield (Table 7). Similarly, we did not see significant improvements with sublimed CNBr. We do not sublime this reagent when newly purchased material is available because of the inherent danger of putting CNBr into the vapor phase.

Overall we obtained 30% to 42% yields of product with >95% homogeneity for the combination of the cleavage and purification steps. When fractions containing product of lower homogeneity were taken into account the yield was between 50% and 60%. Mass spectrometry analysis confirmed the molecular weight of the products. In the case of isotopically labeled M6 peptide percent labeling of the [^{15}N] isomer was calculated to be 98% (Table 8). Whereas in the case of E3-M7-24-T40 percent labeling of the [^{15}N] and [^{13}C/^{15}N] isomers was calculated to be 98% and 96.7%, respectively (Table 8). The above experiments indicated that approximately 10 liters of fermentation were required to get 10 mg quantities of the M6 peptides whereas 1-3 liters of culture sufficed to obtain 10 mg quantities of labeled E3-M7-24-T40 peptide. Thus to obtain [^{15}N]-labeled product required between $100 and $200 of [^{15}N]-labeled starting material for the E3-M7-24-T40 peptide and $400 for each M6 peptide. In the case of [^{13}C/^{15}N] double labeling we found that approximately a 25% decrease in yield occurred when 0.2% glucose was used compared with the 0.4%

glucose normally supplemented. However, this reduction was cost effective and allowed us to obtain the desired isotopically doubly labeled E3-M7-24-T40 for an expenditure of about $700 in [^{15}N]-labeled ammonium sulfate and [^{13}C]-labeled glucose.

IV. 2. Spectroscopic analyses

IV.2.A. *Circular Dichroism* -

Previous CD studies on the evaluation of the secondary structure of peptides corresponding to the 6th and 7th transmembrane domains of Ste2p indicated that in TFE/H_2O synthetic peptides corresponding to the sequences of putative transmembrane domains could assume helical structures (Xie et al., 2000; Naider et al., 2003). In the presence of SDS micelles and DMPC/DMPG vesicles a 6th transmembrane domain peptide M6-31 (KKQFDSFHILLI(Nle)SAQSLLVPSIIFILA YSLK) exhibited β-sheet like CD patterns (Xie et al., 2000). In contrast the 6th transmembrane peptides examined herein, M6 and M6(P258L) [KKQFDSFHILLIAS SQSLLVPSIIFILAYSLKK and KKQFDSFHILLIASSQSLLVLSIIFILAYSLKK], gave helical CD patterns in TFE/H_2O mixtures (Figure 19) and in the presence of SDS and PPG micelles (Figure 20). However, it was not possible to get concentrations of these peptides exceeding 100 μM even in the presence of more than a 500-fold molar excess of detergent. Both M6 and M6(P258L) peptides exhibited β-sheet or aggregation like CD patterns in the presence of DMPC/DMPG vesicles (Figure 21). Comparison of the sequence of M6-31 and M6 indicate very minor differences, in particular a change of Nle to Ala, Ala to Ser, and the addition of one

extra Lys residue at the carboxyl terminus in the latter peptide. Perhaps the extra Lys residue increases the solubility of the peptide and decreases the tendency to aggregate in detergent. However, because of the poor solubility of the M6 peptides in medium containing DMPC/DMPG vesicles, it seems that the extra Lys residue was not enough to solubilize the peptides in vesicles, thereby increasing aggregation. It is also possible that slight changes in the conditions used to prepare the samples affected our results. Nevertheless, our CD results on M6 and M6(P258L) indicate that NMR studies on the present construct are not feasible. Extremely long accumulations would be required to achieve acceptable signal to noise ratios. At the same time the CD results clearly show that studies on this important region of Ste2p will be possible if further increases in solubility of these M6 peptides can be achieved. This goal can be approached by increasing the number of lysine residues at the N- and C-termini, decreasing the length of the hydrophobic core or a combination of these strategies.

The CD studies on E3-M7-24-T40 indicated that both in TFE/H_2O environments and in the presence of various detergents and lipid vesicles this peptide maintained high helicity. Calculations using difference spectroscopy revealed that at ~400 μM concentration in the presence of a 200-fold molar excess of PPG the E3-M7-24 domain was 54% helical. This percent helicity is slightly lower than that calculated for M7-24 using similar approaches (Naider et al., 2003) and reflects the fact that the E3 loop residues are likely unstructured. This conclusion is validated by the direct comparison in DMPC/DMPG vesicles where the calculated difference spectrum for E3-M7-24 indicated a lower helicity than that calculated for M7-24 (Figure 22). Thus CD distinguished the contribution from the transmembrane core

and the E3 loop residues. Most importantly the CD studies provided clear evidence that there is no significant difference in CD spectra of this peptide measured at low (50 mM) and high (500 mM) concentrations (Figure 22). Therefore, using isotopically labeled E3-M7-24-T40 high-resolution NMR studies could be conducted in both organic-aqueous media and the presence of various detergents allowing us to compare structural findings under different membrane-mimetic conditions.

IV.2.B. ***NMR Spectroscopy*** **-**

HSQC spectra of [^{15}N]-labeled peptides are useful to obtain a preliminary assessment concerning the presence of secondary or tertiary structure in a peptide/protein, information about the behavior of a peptide in solution, and its suitability for a further NMR investigation. ^{1}H-^{15}N HSQC spectra of [^{15}N]-E3-M7-24-T40 were measured in organic-aqueous and detergent media. Spectra measured on 0.5 mM concentrations of [^{15}N]-E3-M7-24-T40 in organic-aqueous media (Figure 23) and 0.2 mM or 0.5 mM peptide in the presence of 400-fold molar excess of detergent (Figure 25) were well-resolved and relatively well-dispersed consistent with the existence of secondary structure. In contrast, using a 20-fold molar excess of detergent at peptide concentration of 0.5 mM poorly resolved spectra with extensive line broadening were obtained (Figure 24). No significant improvement in these detergent spectra was obtained using a TROSY pulse sequence that has been recommended for large protein and protein/membrane complexes (Figure 24). Opella and coworkers previously showed that the peptide/detergent ratio is the critical parameter in getting well-resolved spectra in detergents (McDonnell et al., 1993). In

their study poor resolution and in some cases, especially with SDS, peak doublings were observed at low detergent concentrations. In our study, we have not found extra peaks, although the peak assignments for E3-M7-24-T40 in the various detergents are still in progress. When the peptide concentration was increased to nearly 0.5 mM with a concomitant increase in detergent concentration highly resolved spectra were still observed (Figure 25E). However, in contrast to the low concentration samples, which were stable for several weeks, these samples were not stable with time.

A recent report screened 25 detergents for suitability in determining structural information on membrane proteins and concluded that PPG was most fit as judged using five proteins (Krueger-Koplin et al., 2004). Our results indicated that for the E3-M7-24-T40 domain of Ste2p, DHPC and PPG were superior to SDS and DPC as judged by line-width and peak separation; unfortunately, in these detergents sample stability was insufficient for long NMR experiments. As indicated in the results, we were also unable to prepare a stable peptide solution in detergent at peptide concentrations high enough to perform 3D NMR measurements using our 600 MHz NMR instrument. Instead, NMR experiments on the E3-M7-24-T40 peptide in detergent were carried out on an 800 MHz spectrometer equipped with a cryoprobe using 0.2 mM [$^{13}C/^{15}N$]-labeled peptide in the presence of DPC micelles. NMR analysis and structural determination of the E3-M7-24-T40 peptide in DPC detergent are being processed by professor Oliver Zerbe at the University of Zurich, Switzerland.

Based on the fact that stable solutions with 0.5 mM peptide concentrations could be readily obtained in TFE/H_2O (1:1) and in $CDCl_3$/CD_3OH/H_2O (4:4:1), and

that similar CD spectra (Naider et al., 2003) and similar overall peak spreading in ^{15}N-HSQC spectra were observed in all media examined (Figures 23,25), a complete NMR analysis on E3-M7-24-T40 in organic/water media was undertaken. We chose 50% aqueous TFE because the CD results indicated very similar secondary structures in both 75% and 50% TFE (Figure 22A) and we wanted to minimize the concentration of the organic co-solvent. Chloroform/methanol/water has also been used as a membrane mimetic solvent and the EmrE membrane transport protein could be reconstituted in biologically active form from this medium (Schwaiger et al., 1998). Therefore, we also carried out a detailed NMR analysis in chloroform/methanol/water.

Initial insights into the structural tendencies of E3-M7-24-T40 came from analysis of the δCHα values for the E3-M7-24-T40 peptide in TFE/H_2O (1:1) and in $CDCl_3$/CD_3OH/H_2O (4:4:1) at 25 °C (Figure 32). One might question the use of random coil values obtained in water (Wishart et al., 1992) in the determination of chemical shift indices in organic aqueous medium. In comparing the chemical shifts in TFE/water (see Appendix I) with those in chloroform/methanol/water (see Appendix II) we observed that despite the differences in these solvents the CHα shifts are very close, and in ~80% of the residues differed by less than 0.1 ppm. This suggests that the local conformation of the residue is more determining for the CHα chemical shift than the solvent. Moreover, even in a protein the local dielectric constant of residues in the same helix varies significantly depending on whether these residues are buried or exposed. Others have used chemical shift indexing to assess structure of membrane peptides in organic aqueous media and in detergents (Franzoni

et al., 1999; Déméné et al., 2003; Zhang et al., 2002). We believe that provided that the conclusions reached are in agreement with those from other NMR parameters chemical shift indexing provides a useful way to analyze secondary structural tendencies of domains of long peptides.

When analyzing the CSI values in TFE/H_2O and $CDCl_3$/CD_3OH/H_2O, we observed three regions of uninterrupted chemical shift indices of –1 (Figure 32). Using the absolute value of δCHα from the random coil shift and the criterion that a helical region should have at least 4 consecutive residues with chemical shift indices of -1 (Wishart et al., 1992), residues 8-22, 25-52, and 59-70 in TFE/H_2O and residues 8-22, 24-53, and 57-70 in $CDCl_3$/CD_3OH/H_2O have helical tendencies with the first region being most helical, the second less helical and the third least helical. Even in the 25-52 range of residues on the average the amino acids at the beginning have greater δCHα deviations, indicative of a more stable helix. Based on the CHα chemical shift data, 49 of 73 residues (67%) of the peptide are helical. This approach overestimates the percent helicity compared to that determined from the CD data (43%) in TFE/H_2O (1:1).

Using the medium-range CHα-NH [$d_{\alpha N}(i,i+3)$] NOE connectvities as the criterion, three segments of E3-M7-24-T40 have helical tendencies: 9-19, 29-49, and 57-70 in $CDCl_3$/CD_3OH/H_2O and 7-19, 38-52, and 60-72 in TFE/H_2O solutions (Figure 33). The gaps found in the helical regions mostly arise by overlap in crosspeaks, which prevented the assignment of the long-range connectivity (Figure 33,34). In the case of TFE/H_2O (1:1) only 14 of the 61 central residues, excluding both ends (residues 1-7 and 69-73), have (*i, i+3*) connectivities, a characteristic of an

α-helix. Thirty-eight additional residues may have (*i, i+3*) connectivities but this could not be confirmed because of the overlap with strong crosspeaks assigned to (*i, i*) or (*i, i+1*) connectivities.

In TFE/H_2O $J_{NH\text{-}CH\alpha}$ coupling constants ($^3J_{HN\alpha}$) of 3 Hz to 5 Hz were also found for many residues in these same regions with the $^3J_{HN\alpha}$ values drifting up to ~6-7 Hz near the carboxyl terminus (Figure 33C). The increase in $^3J_{HN\alpha}$ values near the amino terminus (the loop region) and the carboxyl terminus (the T40 tail), which have great flexibility, is somewhat expected and elucidated by the fact that the $^3J_{HN\alpha}$ is influenced by the mobility of the peptide having larger $^3J_{HN\alpha}$ values for regions of larger mobility. Thus based on three independent criteria – chemical shift indices, Overhauser connectivities and J coupling values - for E3-M7-24-T40 in TFE/H_2O and based on chemical shift indices and NOE connectivities for E3-M7-24-T40 in $CDCl_3/CD_3OH/H_2O$ we can conclude that the putative transmembrane region (residues 10-33) is helical with a break at the proline-24 residue, that the region following this domain also has significant helical tendencies and that the residues near the end of the peptide have a small but clear tendency to form helices. The disruption of the putative transmembrane helix is probably caused by the proline residue. Since proline has a highly constrained backbone and limited conformational flexibility, it cannot adopt the standard values of ϕ and Ψ angles and the tight turn of the standard α helix and is thus considered a 'helix breaker.' Moreover, the inability of the proline to hydrogen bond because of its doubly substituted (a secondary amide) amide nitrogen would destabilize the α helix even more.

The helical tendency in the cytosolic tail was confirmed by the NMR derived model of the E3-M7-24-T40 peptide that indicated 3 helical domains [residues 5-24, 30-52, and 55-70] (Figure 35,36). The backbone RMSD calculated for the entire peptide was very large (~10 Å), probably because of significant segmental mobility in the complete 73-residue peptide in organic-aqueous solvents. Moreover, this peptide does not assume only one 3D structure in these membrane mimetic solvents. However, when 10 structures modeled for the TFE/H_2O solution with the minimum NOE distance violations were overlaid separately for each of the three helical domains much smaller backbone RMSD values were calculated for these regions of the multidomain peptide (Figure 35). The poorer convergence observed in the model derived from $CDCl_3$/CD_3OH/H_2O data (Figure 35) is likely due to the availability of fewer long-range constraints for use in the calculation. Nevertheless, the similar outcomes from studies in two quite different media suggest that the essential structural features may be an inherent characteristic of this domain of Ste2p.

The conclusion that the multidomain E3-M7-24-T40 receptor peptide has helices with differential stabilities is supported by the CD studies at different temperatures (Figure 22B). About one-third of the helical residues appear to randomize as the temperature is raised to 110 ºC using the Peltier thermoelectric heating/cooling cell holder. The remaining residues likely correspond to the transmembrane region of this peptide and form an extraordinarily stable secondary structure.

A previous CD investigation from our laboratory concluded that a 40-residue peptide corresponding to Ste2p[300-339] had a partial helical structure in TFE/H_2O

(3:1) but was primarily disordered in TFE/H_2O (1:1) (Naider et al., 2003). Similar CD results were obtained on a 23-residue peptide corresponding to Ste2p[350-372] (Reddy et al., 1994). A NMR investigation revealed that the same 23-residue peptide was disordered in H_2O (Arshava et al., 1998). In contrast to the previous CD and NMR data (Reddy et al., 1994; Arshava et al., 1998; Naider et al., 2003) of the tail fragments alone, the present NMR results on the multidomain fragment show that many residues in the tail of Ste2p from position 300 to 339 have helical tendencies when attached to the transmembrane domain. Our hypothesis is that in the context of the transmembrane domain there is some constraint on the conformational freedom of the tail residues resulting in an increased propensity to assume helical conformations. If this is true, the subtraction method previously used (Naider et al., 2003) and currently utilized in this thesis to determine the helicity of the transmembrane domain must be re-evaluated. More insights into this question will come after the structure of E3-M7-24-T40 in detergents is determined.

It is known that the cytosolic tail of Ste2p plays an essential role in desensitization of this GPCR by endocytosis (Hicke et al., 1998). Many studies have shown that GPCR endocytosis requires phosphorylation by kinases and additional protein-protein interactions (Gainetdinov et al., 2004). Since protein-protein interactions often involve helical surfaces (Eilers et al., 2002; Gimpelev et al., 2004), it is reasonable to hypothesize that the helical tendencies found for domains of the cytosolic tail of Ste2p reflect the biologically relevant structure of these regions of the receptor. It is also noteworthy that the X-ray structure of rhodopsin indicated that residues following the seventh transmembrane domain of this GPCR are helical

(Palczewski et al., 2000). This X-ray structure was determined for crystals in a mixed detergent/lipid medium. Our results represent the first NMR data suggesting that a cytosolic tail region of Ste2p has some preferred secondary structure.

The relevance of structures determined in aqueous TFE to the biological state of a protein has been discussed almost since the introduction of this solvent in conformational analysis of peptides (Tamburro et al., 1968; Goodman et al., 1971). Many studies on peptides and protein fragments indicate a reasonably good correlation between the TFE-induced structure of the peptides and the protein structure (Sonnichsen et al., 1992). Our experience with transmembrane peptides from Ste2p is that with the exception of peptides correlating to TM6 domains peptides correlating to other TM domains including TM7 in aqueous-TFE solutions resulted in CD spectra that correlated well with those measured in detergents or lipid bilayers. It is significant that although the HSQC spectra of E3-M7-24-T40 in TFE/H_2O and $CDCl_3$/CD_3OH/H_2O appear different (Figure 23), systematic comparison of nitrogen-15 (^{15}N) chemical shifts and $\delta CH\alpha$ values in both solvents shows that they were almost identical (Figure 31; chemical shift Tables in Appendix). The amide NH proton chemical shifts in both solvents reveal only minor differences (Figure 31). These similarities allow us to conclude that the conformation of the E3-M7-24-T40 peptide in TFE/H_2O and $CDCl_3$/CD_3OH/H_2O will be very close. Unlike TFE/H_2O mixtures, $CDCl_3$/CD_3OH/H_2O has not been considered to be a helix inducing medium. Indeed a high-resolution structure of the *E. coli* EmrE transport protein in $CDCl_3$/CD_3OH/H_2O (6:6:1) has been taken to represent the native structure of this molecule (Schwaiger et al., 1998). Therefore, we believe that the structural

preferences for the 73-residue domain of Ste2p that was determined may be relevant to the functional state of this GPCR. The ability to measure well-dispersed and highly resolved spectra of E3-M7-24-T40 in detergents (Figure 25) and organic-aqueous media (Figure 23), and the ability to model its structure should permit comparison of high-resolution structures of this multidomain peptide under a broad spectrum of conditions. Such structures will provide further insights into the biological relevance of studies in organic-aqueous media.

Chapter V: Conclusions and Perspectives

In conclusion, the successful biosynthesis of isotopically labeled 33-residue peptides [M6 and M6(P258L)] and an isotopically labeled 73-residue multidomain peptide (E3-M7-24-T40) from a GPCR (Ste2p) on a 10-mg scale has been achieved. The results show that biosynthesis of peptide fragments of membrane receptors needs to be optimized to obtain efficient synthesis and that no one expression system or CNBr release condition will likely be useful with all receptor domains.

Circular dichroism analysis of M6, M6(P258L), and E3-M7-24-T40 was performed in TFE/H_2O mixtures, and in the presence of detergent micelles and lipid bilayers. The M6 and M6(P258L) peptides could not be studied at concentrations higher than 100 μM in the presence of detergent micelles and did not integrate into bilayers. In contrast, the E3-M7-24-T40 peptide integrated into PPG and DPC micelles at concentrations (200 μM to 500 μM) suitable for NMR investigations and into bilayers.

The 73-residue multidomain peptide we studied exhibited well-dispersed HSQC spectra in several membrane mimetic solvents, including a variety of detergents at high detergent/peptide ratios. The NMR results in TFE/H_2O (1:1) and in $CDCl_3$/CD_3OH/H_2O (4:4:1) indicate that residues in both the seventh transmembrane domain and the cytosolic tail region of Ste2p may have helical tendencies. We suggest that these helical tendencies may play a role in protein-protein interactions involved in the desensitization of this GPCR.

The sixth transmembrane peptides examined herein, M6 and M6(P258L) [KKQFDSFHILLIASSQSLLVPSIIFILAYSLKK and KKQFDSFHILLIASSQSLLV LSIIFILAYSLKK], were somewhat difficult to biosynthesize and exhibited lower expression than the E3-M7-24-T40 peptide.

We also encountered some problems when studying the M6 peptides using CD spectroscopy. Although the M6 peptides exhibited helical structures in an aqueous-organic solvent (TFE/H_2O) their ellipticities were slightly lower than the ellipticity of the E3-M7-24-T40 in the same solvent. The lower ellipticities of M6 peptides probably resulted from the poor solubility of these peptides in media containing significant percentages of water, and some aggregation may occur even in TFE/H_2O mixtures. This problem was even more evident in the presence of detergent micelles where the ellipticity was significantly lowered compared to values expected for a helical peptide containing 33 residues. It was not possible to get concentrations of these peptides exceeding 100 μM even in the presence of more than a 500-fold molar excess of detergent. In addition, both M6 and M6(P258L) peptides exhibited β-sheet or aggregation-like CD patterns in the presence of DMPC/DMPG vesicles. Although the extra Lys residues seem to increase the solubility of the M6 peptides and decrease their tendency to aggregate in detergent, the extra Lys residues were not sufficient to solubilize the peptides in vesicles. NMR studies on this important region of Ste2p will be possible only if further increases in solubility of these M6 peptides can be achieved.

Several approaches can be attempted to increase the solubility of the M6 peptides. One approach is to use shorter peptide fragments containing only a portion

of the M6 domain. Then the structures of these shorter fragments can be overlapped using a strategy developed by Yeagle et al. (2001). The problem with this approach is there is always the risk that end effects will prevent good overlap of the structures determined for the different parts of the transmembrane domain. Another approach to increase the solubility of the M6 domain is increase the number of lysine residues at the N- and C-termini and/or decrease the length of the hydrophobic core. It is preferred to maintain the number of added Lys residues to a minimum in order to avoid possible interference of these lysine residues with the structure of the peptide. In addition, it is possible that increasing the number of Lys residues might interfere with the peptide insertion in micelles and vesicles, leading to their aggregation, which will hinder biophysical and structural studies. Finally, recent work (Lin et al., 2005) and *de novo* calculations (Gregory Nikiforovich, personal communication) indicate that the hydrophobic core we chose for M6 is too long and a 20 to 22 residue core is probably more representative of the sixth transmembrane domain.

Based on the achievement of well-resolved HSQC NMR spectra of the [^{15}N]-labeled E3-M7-24-T40 peptide in different detergent media, the E3-M7-24-T40 peptide was biosynthesized and labeled with ^{13}C/^{15}N to study its 3D structure in the presence of DPC detergent. DPC was our detergent of choice for NMR studies because solutions of the peptide in this detergent had the best stability. Since we were unable to prepare a stable peptide solution in detergent at peptide concentrations high enough to perform 3D NMR measurements using our 600 MHz NMR instrument, NMR experiments on 0.2 mM [^{13}C/^{15}N]-labeled E3-M7-24-T40 peptide in DPC micelles were carried out on an 800 MHz spectrometer equipped with a cryoprobe.

NMR analysis and structural determination of the E3-M7-24-T40 peptide in DPC detergent are being processed. Determining the NMR structure of the E3-M7-24-T40 peptide in detergent might tell us a great deal about the native structure of the peptide and might strengthen the NMR structural modeling obtained for the peptide in aqueous-organic media.

The [^{15}N]-labeled E3-M7-24-T40 peptide is also being studied in oriented lipid bilayers using solid-state NMR as part of a collaborative effort with Dr. Stanley Opella. No successful results have been obtained as of yet due to difficulties in sample preparation. One possible reason for the difficulties in sample preparations is the tail. Having 40 residues of the C-terminal tail might be too long and hydrophilic to insert or orient in lipid bilayers. If this is the case a [^{15}N]-labeled M7 peptide should be biosynthesized without the tail and studied using solid-state NMR.

List of publications as a result of this thesis work

Estephan, R., Englander, J., Arshava, B., Samples, K., Becker, J. M., and Naider, F. Biosynthesis and NMR analysis of a 73-residue domain of a *Saccharomyces cerevisiae* G protein-coupled receptor, *Biochemistry*, in press.

Estephan, R., Englander, J., Arshava, B., Becker, J. M., and Naider, F. NMR analysis of a multi-domain peptide of the *Saccharomyces cerevisiae* alpha factor receptor, Proceedings of the Nineteenth American Peptide Symposium, in press.

Estephan, R., Englander, J., Arshava, B., Becker, J. M., and Naider, F. Biosynthesis and biophysical studies of fragments of a G protein-coupled receptor: approaches to study the structure of membrane proteins, Proceedings of the Third International and Twenty eighth European Peptide Symposium, in press.

Naider, F., **Estephan, R.**, Englander, J., Suresh Babu, V. V., Arevalo, E., Samples, K., and Becker, J. M. (2004) Sexual conjugation in yeast: A paradigm to study G protein-coupled receptor domain structure, *Biopolymers 76*, 119-128.

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.

Estephan, R., Arevalo, E., Becker, J. M., and Naider, F. (2003) Biosynthesis of the sixth transmembrane domain of the *Saccharomyces cerevisiae* alpha-factor receptor, Proceedings of the Eighteenth American Peptide Symposium, edited by Michael Chorev and Tomi K. Sawyer, American Peptide Society, pp. 754-755.

Appendix

Appendix I

Chemical shift assignments of backbone and side chain protons of the [^{15}N]-labeled E3-M7-24-T40 peptide in TFE/H_2O (1:1) at 25 °C								
Sequence #	AA[a]	CSI[b]	N	NH	CHα	CHβ	CHγ	CHδ
1	Ser	-1			4.105	4.003		
2	Leu	0	122.72	8.375	4.401	1.590	1.479	0.877/ 1.026
3	Lys	1	122.48	8.089	4.660	1.849	1.464/ 1.322	1.698
4	Pro	0			4.381	2.273	2.04/ 1.935	3.645/ 3.778
5	Asn	0	117.36	8.259	4.653	2.858		
6	Gln	-1	120.78	8.288	4.250	2.128	2.363	
7	Gly	0	108.22	8.419	3.911			
8	Thr	-1	114.05	7.880	4.107	4.242	1.235	
9	Asp	-1	122.04	8.230	4.552	2.972		
10	Val	-1	123.11	8.042	3.710	2.164	0.935/ 1.036	
11	Leu	-1	119.97	7.859	4.059	1.638/ 1.752	1.562	
12	Thr	-1	110.87	7.996	3.952	4.240	1.258	
13	Thr	-1	122.03	7.843	3.940	4.450	1.211	
14	Val	-1	122.78	8.330	3.557	2.111	0.902/ 1.030	
15	Ala	-1	120.27	8.402	3.981	1.471		
16	Thr	-1	114.48	7.732	3.908	4.412	1.215	
17	Leu	-1	123.35	8.147	4.039	1.794	1.703	

Appendix I (Continued)

18	Leu	-1	117.41	8.497	4.007	1.854		
19	Ala	-1	123.07	7.837	4.061	1.584		
20	Val	-1	119.25	8.243	3.742	2.270	0.948/ 1.090	
21	Leu	-1	119.35	8.553	4.177	1.892		
22	Ser	-1	112.05	7.739	4.365	3.977/ 4.060		
23	Leu	0	124.92	7.750	4.453	1.978	1.640	
24	Pro	0			4.349	2.260/ 2.127	1.860/ 1.967	3.763/ 4.003
25	Leu	-1	117.37	7.471	4.149	1.693		
26	Ser	-1	114.56	8.141	4.167	4.006/ 4.060		
27	Ser	-1	117.54	8.021	4.322	3.882/ 3.965		
28	Leu	-1	126.18	7.848	4.147	1.683		
29	Trp	-1	120.07	8.137	4.378	3.350		
30	Ala	-1	126.21	8.281	4.091	1.454		
31	Thr	-1	115.05	7.928	3.949	4.360	1.240	
32	Ala	-1	121.03	8.227	3.976	1.476		
33	Ala	-1	121.64	8.420	3.993	1.237		
34	Asn	-1	117.91	8.157	4.456	2.822/ 2.903		
35	Asn	-1	119.26	8.298	4.509	2.826		
36	Ala	-1	124.93	8.423	4.148	1.486		
37	Ser	-1	113.91	8.119	4.268	4.031		
38	Lys	-1	123.11	8.042	4.174	1.963		
39	Thr	-1	114.05	8.063	4.020	4.270	1.225	

Appendix I (Continued)

40	Asn	-1	121.74	8.372	4.553	2.866		
41	Thr	-1	117.51	8.098	4.035	4.334	1.226	
42	Ile	-1	122.32	8.177	3.863	1.907	0.898	
43	Thr	-1	113.24	8.092	4.066	4.279	1.276	
44	Ser	-1	118.89	8.130	4.325	3.954/ 4.024		
45	Asp	-1	122.18	8.365	4.525	2.828/ 3.012		
46	Phe	-1	121.02	8.449	4.414	3.203		
47	Thr	-1	114.39	8.214	4.061	4.313	1.289	
48	Thr	-1	114.80	8.028	4.215	4.320	1.252	
49	Ser	-1	116.89	7.943	4.344	3.924		
50	Thr	-1	113.68	7.758	4.208	4.208	1.104	
51	Asp	-1	120.31	8.019	4.605	2.904		
52	Arg	-1	119.76	7.741	4.195	1.636	1.397	
53	Phe	0	119.29	7.735	4.600	2.926/ 3.075		
54	Tyr	1	123.10	7.825	4.750	2.920/ 3.001		
55	Pro	-1			4.271	2.200	1.891/ 1/965	3.355/ 3.682
56	Gly	0	106.56	7.464	3.884			
57	Thr	-1	114.30	7.811	4.237	4.237	1.224	
58	Leu	0	122.76	7.942	4.290	1.635	1.635	0.857
59	Ser	-1	114.57	8.057	4.357	3.885/ 3.955		
60	Ser	-1	118.26	8.052	4.324	3.868		
61	Phe	-1	122.00	7.926	4.454	3.139		

Appendix I (Continued)

62	Gln	-1	120.01	8.159	4.147	2.125	2.359	
63	Thr	-1	113.55	8.019	4.115	4.261	1.221	
64	Asp	-1	120.51	8.250	4.620	2.916		
65	Ser	-1	115.77	7.974	4.312	3.805/ 3.902		
66	Ile	-1	121.90	7.806	4.006	1.881	1.547/ 1.168	0.893
67	Asn	-1	119.80	8.054	4.603	2.766		
68	Asn	-1	118.76	8.108	4.618	2.803		
69	Asp	-1	119.40	8.279	4.627	2.918		
70	Ala	-1	123.66	8.072	4.243	1.414		
71	Lys	0	118.69	7.869	4.329	1.813/ 1.904	1.470	1.688
72	Ser	0	115.59	7.955	4.503	3.797/ 3.905		
73	Ser	0	117.41	7.968	4.521	3.901/ 3.974		

[a] AA is amino acid in the sequence.

[b] CSI is the chemical shift index (Wishart et al., 1992; Wishart and Sykes, 1994).

Appendix II

Chemical shift assignments of backbone and side chain protons of the [^{15}N]-labeled E3-M7-24-T40 peptide in $CDCl_3/CD_3OH/H_2O$ (4:4:1) at 25 °C								
Sequence #	AA[a]	CSI[b]	N	NH	CHα	CHβ	CHγ	CHδ
1	Ser	-1			4.050	3.94		
2	Leu	1	121.60	8.63	4.410	1.590		0.90
3	Lys	1	121.03	8.18	4.620	1.84		1.68
4	Pro	0			4.390	2.27/ 2.09	1.96	3.67/ 3.82
5	Asn	-1	116.37	8.38	4.650	2.87		
6	Gln	-1	119.75	8.57	4.140	2.13	2.40	
7	Gly	0	106.60	8.54	3.890			
8	Thr	-1	115.18	8.015	3.960	4.21	1.25	
9	Asp	-1	121.70	8.27	4.480	3.07/ 2.89		
10	Val	-1	122.36	8.177	3.650	2.25	0.99/ 1.12	
11	Leu	-1	118.90	8.03	4.000	1.83/ 1.66	1.66	0.94
12	Thr	-1	112.04	8.24	3.900	4.23	1.27	
13	Thr	-1	121.59	7.96	3.890	4.43	1.275	
14	Val	-1	121.30	8.31	3.550	2.17	0.99/ 1.095	
15	Ala	-1	119.75	8.42	4.005	1.545		
16	Thr	-1	113.47	7.88	3.880	4.38	1.270	
17	Leu	-1	121.74	8.11	4.035	1.85	1.74	0.935
18	Leu	-1	117.05	8.37	4.006	1.89	1.52	0.90
19	Ala	-1	122.13	8.13	4.060	1.63		

Appendix II (Continued)

20	Val	-1	117.74	8.28	3.720	2.30	1.140/ 1.045	
21	Leu	0	117.64	8.32	4.160	1.92	1.60	0.92
22	Ser	-1	110.63	7.78	4.350	4.035		
23	Leu	1	124.00	7.85	4.405	2.07	1.71	0.995
24	Pro	-1			4.335	1.93/ 2.37	1.69	3.75/ 3.97
25	Leu	0	115.80	7.55	4.180	1.83	1.66	0.94
26	Ser	-1	114.90	8.42	4.250	4.05		
27	Ser	-1	117.06	8.19	4.280	3.95/ 4.05		
28	Leu	0	125.46	7.94	4.125	1.79	1.92	0.93
29	Trp	-1	119.93	8.267	4.280	3.37/ 3.42		
30	Ala	-1	119.66	8.562	3.940	1.568		
31	Thr	-1	114.78	8.02	3.890	4.34	1.270	
32	Ala	-1	125.14	8.315	4.070	1.495		
33	Ala	-1	120.70	8.617	3.970	1.275		
34	Asn	-1	117.58	8.31	4.415	2.843		
35	Asn	-1	118.15	8.385	4.437	2.90		
36	Ala	-1	124.10	8.578	4.110	1.525		
37	Ser	-1	113.27	8.18	4.235	4.055		
38	Lys	-1	121.40	8.14	4.110	1.96		
39	Thr	-1	114.42	8.12	3.900	4.216	1.243	
40	Asn	-1	121.3	8.45	4.520	2.890		
41	Thr	-1	117.54	8.194	3.915	4.306	1.241	

Appendix II (Continued)

42	Ile	-1	121.03	8.167	3.790	1.98	1.71/ 0.86	0.94
43	Thr	-1	113.97	8.250	3.980	4.270	1.290	
44	Ser	-1	118.70	8.340	4.290	4.020		
45	Asp	-1	121.80	8.520	4.470	2.82/ 3.12		
46	Phe	-1	120.30	8.600	4.320	3.265		
47	Thr	-1	114.87	8.460	3.940	4.330	1.316	
48	Thr	-1	114.77	8.220	4.040	4.280	1.270	
49	Ser	-1	115.39	7.860	4.210	3.950		
50	Thr	-1	113.29	7.740	4.090	4.190	1.140	
51	Asp	-1	118.90	8.020	4.550	2.920		
52	Arg	-1	117.37	7.676	4.125	1.65	1.424	
53	Phe	-1	116.50	7.785	4.555	2.933/ 3.157		
54	Tyr	1	120.54	7.937	4.713	3.048		
55	Pro	-1			4.303	2.23/ 2.03	1.93	3.436/ 3.78
56	Gly	0	105.67	7.913	3.902			
57	Thr	-1	114.30	7.89	4.15	4.240	1.225	
58	Leu	0	121.40	8.14	4.230	1.66	1.66	0.90
59	Ser	-1	113.10	8.09	4.330	3.920		
60	Ser	-1	117.82	8.08	4.270	3.860/ 3.920		
61	Phe	-1	121.08	8.08	4.370	3.190		
62	Gln	-1	118.90	8.30	4.080	2.14/ 2.22	2.33/ 2.41	
63	Thr	-1	113.66	8.13	3.960	4.230	1.220	

Appendix II (Continued)

64	Asp	-1	119.36	8.324	4.510	2.850/ 2.930		
65	Ser	-1	114.85	7.95	4.210	3.78/ 3.90		
66	Ile	0	120.91	7.89	3.900	1.90	1.63/ 1.19	0.92
67	Asn	-1	118.03	8.03	4.510	2.750		
68	Asn	-1	117.11	8.10	4.520	2.790		
69	Asp	-1	118.47	8.29	4.520	2.880/ 2.960		
70	Ala	-1	121.70	8.13	4.210	1.450		
71	Lys	0	116.30	7.77	4.29	1.84/ 1.90	1.50	1.67
72	Ser	0	113.40	7.82	4.500	3.900		
73	Ser	0	115.47	7.79	4.500	3.920		

[a] AA is amino acid in the sequence.

[b] CSI is the chemical shift index (Wishart et al., 1992; Wishart and Sykes, 1994).

Bibliography

Abel, M. G., Lee, B. K., Naider, F., and Becker, J. M. (1998) Mutations affecting ligand specificity of the G-protein-coupled receptor for the *Saccharomyces cerevisiae* tridecapeptide pheromone, *Biochim. Biophys. Acta 1448*, 12-26.

Akal-Strader, A., Khare, S., Xu, D., Naider, F., and Becker, J. M. (2002) Residues in the first extracellular loop of a G protein-coupled receptor play a role in signal transduction, *J. Biol. Chem. 277*, 30581-30590.

Albert, A. D., and Yeagle, P. L. (2000) Domain approach to three-dimensional structure of rhodopsin using high-resolution nuclear magnetic resonance, *Methods Enzymol. 315*, 107-115.

Altman, J.D., Henner, D., Nilsson, B., Anderson, S., Kuntz, I.D. (1991) Intracellular expression of BPTI fusion proteins and single column cleavage/affinity purification by chymotrypsin, *Protein Eng. 4*, 593-600.

Angelova, K., Narayan, P., Simon, J. P., and Puett, D. (2000) Functional role of transmembrane helix 7 in the activation of the heptahelical lutropin receptor, *Mol. Endocrinol. 14*, 459-471.

Archer, E., Maigret, B., Escrieut, C., Pradayrol, L., and Fourmy, D. (2003) Rhodopsin crystal: new template yielding realistic models of G-protein-coupled receptors?, *Trends Pharmacol. Sci. 24*, 36-40.

Arevalo, E. (2002) Biosynthesis, purification and biophysical analysis of domains of the *Saccharomyces cerevisiae* alpha-factor receptor, *Doctoral Dissertation, City University of New York,* pp 41-52.

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.

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.

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.

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

Baldwin, J. M., Schertler, G. F., and Unger, V. M. (1997) An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors, *J. Mol. Biol. 12*, 144-164.

Bardwell, L., Cook, J. G., Inouye, C. G., and Thorner, J. (1994) Signal propagation and regulation in the mating pheromone response pathway of the yeast *Saccharomyces cerevisiae*, *Dev. Biol. 166*, 363–379.

Bargmann, C. I., Hung, M. C., and Weinberg, R. A. (1986) Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185, *Cell 45*, 649-657.

Biebermann, H., Schoneberg, T., Schulz, A., Krause, G., Gruters, A., Schultz, G., and Gudermann, T. (1998) A conserved tyrosine residue (Y601) in transmembrane domain 5 of the human thyrotropin receptor serves as a molecular switch to determine G-protein coupling, *FASEB J. 12*, 1461-1471.

Blumer, K., and Thorner, J. (1991) Receptor-G protein signaling in yeast, *Annu. Rev. Physiol*. *53*, 37-57.

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

Bockaert, J., and Pin, J. P. (1999) Molecular tinkering of G protein-coupled receptors: an evolutionary success, *EMBO J. 18*, 1723-1729.

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

Böhm, S. K., Grady, E. F., and Bunnett, N. W. (1997) Regulatory mechanisms that modulate signaling by G-protein-coupled receptors, *Biochem. J. 322*, 1-18.

Booth, P. J. (2003) The trials and tribulations of membrane protein folding in vitro, *Biochim. Biophys. Acta 1610*, 51-56.

Bourne, H. R. (1997) How receptors talk to trimeric G proteins, *Curr. Opin. Cell Biol. 9*, 134-142.

Bourne, H. R., and Meng, E. C. (2000) Rhodopsin sees the light, *Science 289*, 733-734.

Braunschweiler, L., and Ernst, R. (1983) Coherence transfer by isotopic mixing: application to proton correlation spectroscopy, *J. Magn. Reson. 53*, 521-528.

Burger, K., Gimpl, G., and Fahrenholz, F. (2000) Regulation of receptor function by cholesterol, *Cell. Mol. Life Sci. 57*, 1577-1592.

Bystrov, V. F. (1976) Spin-spin coupling and the conformational states of peptide systems, *Prog. Nucl. Magn. Reson. Spectrosc. 10*, 41-82.

Callaghan, R., Berridge, G., Ferry, D. R., and Higgins, C. F. (1997) The functional purification of P-glycoprotein is dependent on maintenance of a lipid-protein interface, *Biochim. Biophys. Acta 1328*, 109-124.

Cartwright, C. P., and Tipper, D. J. (1991) In vivo topological analysis of Ste2, a yeast plasma membrane protein, by using beta-lactamase gene fusions, *Mol. Cell Biol. 11*, 2620-2628.

Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., and Skelton, N. J. (1996) *Protein NMR spectroscopy: Principles and practice, Chapter 7, Part 2: Heteronuclear-edited NMR spectroscopy,* pp 447-468*, and Chapter 8: Sequential assignments and structure calculations*, pp 532-556, Academic Press, San Diego, California.

Celic, A., Martin, N. P., Son, C. D., Becker, J. M., Naider, F., and Dumont M. E. (2003) Sequences in the intracellular loops of the yeast pheromone receptor Ste2p required for G protein activation, *Biochemistry 42*, 3004-3017.

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

Chapot, M. P., Eshdat, Y., Marullo, S., Guillet, J. G., Charbit, A., Strosberg, A. D. and Delavier-Klutchko, C. (1990) Localization and characterization of three different beta-adrenergic receptors expressed in *Escherichia coli*, *Eur. J. Biochem. 187*, 137-144.

Chen, Y. H., Yang, J. T., and Chau, K. H. (1974) Determination of the helix and beta form of proteins in aqueous solution by circular dichroism, *Biochemistry 13*, 3350-3359.

Chin, C. N., von Heijne, G., and de Gier, J. W. (2002) Membrane proteins: shaping up, *Trends Biochem. Sci. 27*, 231-234.

Choma, C., Gratkowski, H., Lear, J. D., and DeGrado, W. F. (2000) Asparagine-mediated self-association of a model transmembrane helix, *Nat. Struct. Biol. 7*, 161-166.

Chopra, A., Yeagle, P. L., Alderfer, J. A., and Albert, A. D. (2000) Solution structure of the sixth transmembrane helix of the G-protein-coupled receptor, rhodopsin, *Biochim. Biophys. Acta 1463*, 1-5.

Chou, P. Y., and Fasman, G. D. (1974) Prediction of protein conformation, *Biochemnistry 13*, 222-245.

Clark, C. D., Palzkill, T., and Botstein, D. (1994) Systematic mutagenesis of the yeast mating pheromone receptor third intracellular loop, *J. Biol. Chem. 269*, 8831-8841.

Cramer, W. A., Engelman, D. M., Von Heijne, G., and Rees, D. C. (1992) Forces involved in the assembly and stabilization of membrane proteins, *FASEB J. 6*, 3397-3402.

Crowe, M. L., Perry, B. N., and Connerton, I. F. (2000) Golf complements a GPA1 null mutation in *Saccharomyces cerevisiae* and functionally couples to the STE2 pheromone receptor, *J. Recept. Signal Transduction Res. 20*, 61-73.

Curran, A. R., Templer, R. H., and Booth, P. J. (1999) Modulation of folding and assembly of the membrane protein bacteriorhodopsin by intermolecular forces within the lipid bilayer, *Biochemistry 38*, 9328-9336.

Deber, C. M., Liu, L. P., and Wang, C. (1999) Perspective: Peptides as mimics of transmembrane segments in proteins, *J. Pept. Res*. *54*, 200-205.

DeGrado, W. F., Gratkowski, H., and Lear, J. D. (2003) How do helix-helix interactions help determine the folds of membrane proteins? Perspectives from the study of homo-oligomeric helical bundles, *Protein Sci. 12*, 647-665.

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

Déméné, H., Granier, S., Muller, D., Guillon, G., Dufour, M.-N., Delsuc, M.-A., Hibert, M., Pascal, R., and Mendre, C. (2003) Active peptidic mimics of the second intracellular loop of the V_{1A} vasopressin receptor are structurally related to the second intracellular rhodopsin loop: A combined ^{1}H NMR and biochemistry study, *Biochemistry 42*, 8204-8213.

Ding, F., Xie, H., Arshava, B., Becker, J. M., and Naider, F. (2001) ATR-FTIR study of the structure and orientation of transmembrane domains of the *Saccharomyces cerevisiae* α-mating factor receptor in phospholipids, *Biochemistry 40*, 8945-8954.

Dohlman, H. G. (2002) G proteins and pheromone signaling, *Annu. Rev. Physiol. 64*, 129-159.

Dohlman, H. G., and Thorner, J. W. (2001) Regulation of G protein-initiated signal transduction in yeast: paradigms and principles, *Annu. Rev. Biochem*. *70*, 703-754.

Dohlman, H. G., Thorner, J., Caron, M. G., and Lefkowitz, R. J. (1991) Model systems for the study of seven-transmembrane-segment receptors, *Annu. Rev. Biochem. 60*, 653-688.

Dosil, M., Schandel, K. A., Gupta, E., Jenness, D. D., and Konopka, J. B. (2000) The C terminus of the *Saccharomyces cerevisiae* alpha-factor receptor contributes to the formation of preactivation complexes with its cognate G protein, *Mol. Cell. Biol. 20*, 5321-5329.

Dube, P., and Konopka, J. B. (1998) Identification of polar region in transmembrane domain 6 that regulates the function of the G protein-coupled α-factor receptor, *Mol. Cell. Biol. 18*, 7205-7215.

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

Duntze, W., Betz, R., and Nientiedt, M. (1994) Pheromones in yeasts. In: Wessel, J. G. H., Meinhardt, F. (Eds.). Growth, differentiation and sexuality, The Mycota, Vol. I., pp 381-399, Springer, Heidelberg.

Edwards, A. M., Arrowsmith, C. H., Christendat, D., Dharamsi, A., Friesen, J. D., Greenblatt, J. F., and Vedadi, M. (2000) Protein production: feeding the crystallographers and NMR spectroscopists, *Nat. Struct. Biol. 7 Suppl.*, 970-972.

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.

Eilers, M., Patel, A. B., Liu, W., and Smith, S. O. (2002) Comparison of helix interactions in membrane and soluble alpha-bundle proteins, *Biophys. J. 82*, 2720-2736.

Eilers, M., Shekar, S. C., Shieh, T., Smith, S. O., and Fleming, P. J. (2000) Internal packing of helical membrane proteins, *Proc. Natl. Acad. Sci. U.S.A. 97*, 5796-5801.

Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L., and Khorana, H. G. (1996) Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin, *Science 274*, 768-770.

Feng, Y., and Davis, N. G. (2000) Feedback phosphorylation of the yeast a-factor receptor requires activation of the downstream signaling pathway from G protein through mitogen-activated protein kinase, *Mol. Cell. Biol. 20*, 563-574.

Fiermonte, G., Walker, J. E., and Palmieri, F. (1993) Abundant bacterial expression and reconstitution of an intrinsic membrane-transport protein from bovine mitochondria, *Biochem. J. 294*, 293-299.

Franzoni, L., Nicastro, G., Pertinhez, T. A., Oliveira, E., Nakaie, C. R., Paiva, A. C. M., Schreier, S., and Spisni, A. (1999) Structure of two fragments of the third cytoplasmic loop of the rat angiotensin II AT_{1A} receptor, *J. Biol. Chem. 274*, 227-235.

Fraser, C. M., Lee, N. H., Pellegrino, S. M., and Kerlavage, A. R. (1994) Molecular properties and regulation of G-protein-coupled receptors, *Prog. Nucleic Acids Res. Mol. Biol*. *49*, 113-156.

Gainetdinov, R. R., Premont, R. T., Bohn, L. M., Lefkowitz, R. J., and Caron, M. G. (2004) Desensitization of G protein-coupled receptors and neuronal functions, *Annu. Rev. Neurosci. 27*, 107-144.

Garavito, R. M., Picot, D., and Loll, P. J. (1995) The 3.1 Å X-ray crystal structure of the integral membrane enzyme prostaglandin H_2 synthase-1, *Adv. Prostaglandin Thromboxane Leukot. Res. 23*, 99-103.

Gether, U. (2000) Uncovering molecular mechanisms involved in activation of G protein-coupled receptors, *Endocr. Rev. 21*, 90-1.

Gimpelev, M., Forrest, L. R., Murray, D., and Honig, B. (2004) Helical packing patterns in membrane and soluble proteins, *Biophys. J. 87*, 4075-4086.

Goodman, M., Naider, F., and Toniolo, C. (1971) Circular dichroism studies of isoleucine oligopeptides in solution, *Biopolymers 10*, 1719-1730.

Goulding C. W., and Perry, L. J. (2003) Protein production in *Escherichia coli* for structural studies by X-ray crystallography, *J. Struct. Biol. 142*, 133-143.

Gratkowski, H., Lear, J. D., and DeGrado, W. F. (2001) Polar side chains drive the association of model transmembrane peptides, *Proc. Natl. Acad. Sci. U. S. A. 98*, 880-885.

Grisshammer, R., and Tate, C. G. (1995) Overexpression of integral membrane proteins for structural studies, *Q. Rev. Biophys. 28*, 315-422.

Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA, *J. Mol. Biol. 273*, 283-298.

Gurrath, M. (2001) Peptide-binding G protein-coupled receptors: New opportunities for drug design, *Curr. Med. Chem. 8*, 1605-1648.

Herzyk, P., and Hubbard, R. E. (1995) Automated method for modeling seven-helix transmembrane receptors from experimental data, *Biophys. J. 69*, 2419-2442.

Herzyk, P., and Hubbard, R. E. (1998) Combined biophysical and biochemical information confirms arrangement of transmembrane helices visible from the three-dimensional map of frog rhodopsin, *J. Mol. Biol*. *281*, 741-754.

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

Hicke, L., Zanolari, B., and Riezman, H. (1998) Cytoplasmic tail phosphorylation of the alpha-factor receptor is required for its ubiquitination and internalization, *J. Cell Biol*. *141*, 349-358.

Hines, J., Heerding, J. N., Fluharty, S. J., and Yee, D. K. (2001) Identification of angiotensin II type 2 (AT2) receptor domains mediating high-affinity CGP 42112A binding and receptor activation, *J. Pharmacol. Exp. Ther. 298*, 665-673.

Hirschman, J., De Zutter, G., Simonds, W., and Jennes, D. (1997) The G beta gamma complex of the yeast pheromone response pathway. Subcellular fractionation and protein-protein interactions, *J. Biol. Chem. 272*, 240-248.

Horn, F., Vriend, G., and Cohen, F. E. (2001) Collecting and harvesting biological data: the GPCRDB and NucleaRDB information systems, *Nucleic Acids Res*. *29*, 346-349.

Huang, P., Li, J., Chen, C., Visiers, I., Weinstein, H., and Liu-Chen, L. Y. (2001) Functional role of a conserved motif in TM6 of the rat μ opioid receptor: Constitutively active and inactive receptors result from substitutions of Thr6.34(279) with Lys and Asp, *Biochemistry 40*, 13501-13509.

Hunt, J. F., Earnest, T. N., Bousché, O., Kalghatgi, K., Reilly, K., Horváth, C., Rothschild, K. J., and Engelman, D. M. (1997a) A biophysical study of integral membrane protein folding, *Biochemistry 36*, 15156-15176.

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

Javadpour, M. M., Eilers, M., Groesbeek, M., and Smith, S. O. (1999) Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association, *Biophys. J. 77*, 1609-1618.

Jeansonne, N. E. (1994) Yeast as a model system for mammalian seven-transmembrane segment receptors, *Proc. Soc. Exp. Biol. Med. 206*, 35-44.

Jeener, J., Meier, B., Bachmann, P., and Ernst, R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy, *J. Chem. Phys. 71*, 4546-4553.

Ji, T. H, Grossmann, M., and Ji, I. (1998) G protein-coupled receptors. I. Diversity of receptor-ligand interactions, *J. Biol. Chem. 273*, 17299-17302.

Johnson, B. A., and Blevins, R. A. (1994) A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603-614.

Kallal, L., and Kurjan, J. (1997) Analysis of the receptor binding domain of Gpa1p, the G(alpha) subunit involved in the yeast pheromone response pathway, *Mol. Cell. Biol. 17*, 2897-2907.

Karnik, S. S., Gogonea, C., Patil, S., Saad, Y., and Takezako, T. (2003) Activation of G-protein-coupled receptors: A common molecular mechanism, *Trends Endocrinol. Metab*. *14*, 431-437.

Karplus, M. (1959) Contact electron-spin coupling of nuclear magnetic moments, *J. Phys. Chem. 30*, 11-15.

Katragadda, M., Alderfer, J. L., and Yeagle, P. L. (2001a) Assembly of a polytopic membrane protein structure from the solution structures of overlapping peptide fragments of bacteriorhodopsin, *Biophys. J*. *81*, 1029-1036.

Katragadda, M., Chopra, A., Bennett, M., Alderfer, J. L., Yeagle, P. L., Albert, A. D. (2001b) Structures of the transmembrane helices of the G-protein coupled receptor, rhodopsin, *J. Pept. Res*. *58*, 79-89.

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.

Kiefer, H. (2003) In vitro folding of alpha-helical membrane proteins, *Biochim. Biophys. Acta 1610*, 57-62.

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

Kiefer, H., Vogel, R., and Maier K. (2000) Bacterial expression of G-protein-coupled receptors: prediction of expression levels from sequence, *Recept. Channels 7*, 109-119.

King, K., Dohlman, H. G., Thorner, J., Caron, M. G., and Lefkowitz, R. J. (1990) Control of yeast mating signal transduction by a mammalian beta 2-adrenergic receptor and Gs alpha subunit, *Science 250*, 121-123.

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

Konopka, J. B., Margarit, S. M., and Dube, P. (1996) Mutation of Pro-258 in transmembrane domain 6 constitutively activates the G protein-coupled alpha-factor receptor, *Proc. Natl. Acad. Sci. U.S.A. 93*, 6764-6769.

Konopka, J. B., and Jenness, D. (1991) Genetic fine-structural analysis of the *Saccharomyces cerevisiae* alpha-pheromone receptor, *Cell Reg. 2*, 439-452.

Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51-55.

Kostenis, E., Conklin, B., and Wess, J. (1997) Molecular basis of receptor/G protein coupling selectivity studied by coexpression of wild type and mutant m2 muscarinic receptors with mutant Gαq subunits, *Biochemistry 36*, 1487-1495.

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

Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection, *Proc. Natl. Acad. Sci. U.S.A. 82*, 488-492.

Kunkel, T. A., Roberts, J., and Zakour, R. (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection, *Methods Enzymol. 154*, 367-382.

Kurjan, J. (1993) The pheromone response pathway in *Saccharomyces cerevisiae*, *Annu. Rev. Genet. 27*, 147-179.

LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F., and McCoy, J. M. (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm, *Biotechnology 11*, 187-193.

Lawson, Z., and Wheatley, M. (2004) The third extracellular loop of G-protein-coupled receptors: more than just a linker between two important transmembrane helices, *Biochem. Soc. Trans. 32*, 1048-1050.

Lazarova, T., Stoeber, K., and Robinson, C. R. (2004) Characterization of peptides corresponding to the seven transmembrane domains of the human adenosine A_2a receptor, *Biochemistry 43*, 12945-12954.

Leberer, E., Thomas, D. Y., and Whiteway, M. (1997) Pheromone signaling and polarized morphogenesis in yeast, *Curr. Opin. Genet. Dev. 7*, 59-66.

Lee, A. G. (1998) How lipids interact with an intrinsic membrane protein: the case of the calcium pump, *Biochim. Biophys. Acta 1376*, 381-390.

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

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

Levi, V., Rossi, J. P., Echarte, M. M., Castello, P. R., and Gonzalez Flecha, F. L. (2000) Thermal stability of the plasma membrane calcium pump. Quantitative analysis of its dependence on lipid-protein interactions, *J. Membr. Biol. 173*, 215-225.

Li, S. C., and Deber, C. M. (1993) Peptide environment specifies conformation. Helicity of hydrophobic segments compared in aqueous, organic, and membrane environments, *J. Biol. Chem. 268*, 22975-22978.

Li, S. C., Goto, N. K., Williams, K. A., and Deber, C. (1996) α-Helical, but not β-sheet, propensity of proline is determined by peptide environment, *Proc. Natl. Acad. Sci. U.S.A. 93*, 6676-6681.

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

Lin, J. C., Duell, K., Saracino, M., and Konopka, J. B. (2005) Identification of residues that contribute to receptor activation through the analysis of compensatory mutations in the G protein-coupled α-factor receptor, *Biochemistry 44*, 1278-1287.

Lin, J. C., Parrish, W., Eilers, M., Smith, S. O., and Konopka, J. B. (2003) Aromatic residues at the extracellular ends of transmembrane domains 5 and 6 promote ligand activation of the G protein-coupled alpha-factor receptor, *Biochemistry 42*, 293-301.

Lin, Z., Shenker, A., and Pearlstein, R. (1997) A model of the lutropin/choriogonadotropin receptor: insights into the structural and functional effects of constitutively activating mutations, *Protein Eng. 10*, 501-510.

Lindhout, D. A., Thiessen, A., Schieve, D., and Sykes, B. D. (2003) High-yield expression of isotopically labeled peptides for use in NMR studies, *Protein Sci. 12*, 1786-1791.

Liri, T., Farfel, Z., and Bourne, H. (1998) G-protein diseases furnish a model for the turn-on switch, *Nature 394*, 35-38.

Liu, J., Schöneberg, T., Rhee, M., and Wess, J. (1995) Mutational analysis of the relative orientation of transmembrane helices I and VII in G protein-coupled receptors, *J. Biol. Chem. 270*, 19532-19539.

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

Lund, S., Orlowski, S., de Foresta, B., Champeil, P., le Maire, M., and Moller, J. V. (1989) Detergent structure and associated lipid as determinants in the stabilization of solubilized Ca2+-ATPase from sarcoplasmic reticulum, *J. Biol. Chem. 264*, 4907-4915.

Lundstrom K. (2004) Structural genomics on membrane proteins: mini review, *Comb. Chem. High Throughput Screening 7*, 431-439.

Machamer, C. E., Grim, M. G., Esquela, A., Chung, S. W., Rolls, M., Ryan, K., and Swift, A. M. (1993) Retention of a cis Golgi protein requires polar residues on one face of a predicted alpha-helix in the transmembrane domain, *Mol. Biol. Cell 4*, 695-704.

Manolios, N., Bonifacino, J. S., and Klausner, R. D. (1990) Transmembrane helical interactions and the assembly of the T cell receptor complex, *Science 249*, 274-277.

Marassi, F., and Opella, S. (1998). NMR structural studies of membrane proteins, *Curr. Opin. Struct. Biol. 8*, 640-648.

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

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

Marsh, D. (1990) Lipid-protein interactions in membranes, *FEBS Lett. 268*, 371-375.

Marsh, L. (1992) Substitutions in the hydrophobic core of the alpha-factor receptor of *Saccharomyces cerevisiae* permit response to *Saccharomyces kluyveri* alpha-factor and to antagonist, *Mol. Cell. Biol. 12*, 3959-3966.

Marshall, G. R. (2001) Peptide interactions with G-protein coupled receptors, *Biopolymers 60*, 246-277.

Marti, T. (1998) Refolding of bacteriorhodopsin from expressed polypeptide fragments, *J. Biol. Chem*. *273*, 9312-9322.

Martin, N. P., Celic, A., and Dumont, M. E. (2002) Assembly of G protein-coupled receptors from fragments: identification of functional receptors with discontinuities in each of the loops connecting transmembrane segments, *J. Mol. Biol. 317*, 765-788.

Martin, N. P., Leavitt, L. M., Sommers, C. M., and Dumont, M. E. (1999) Mutagenic mapping of helical structures in the transmembrane segments of the yeast alpha-factor receptor, *Biochemistry 38*, 682-695.

Marullo, S., Delavier-Klutchko, C., Eshdat, Y., Strosberg, A. D., and Emorine, L. (1988) Human beta 2-adrenergic receptors expressed in *Escherichia coli* membranes retain their pharmacological properties, *Proc. Natl. Acad. Sci. U.S.A. 85*, 7551-7555.

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

Mierke, D. F., and Giragossian, C. (2001) Peptide hormone binding to G-protein-coupled receptors: structural characterization via NMR techniques, *Med. Res. Rev. 21*, 450-471.

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

Mosberg, H. I. (1999) Complementarity of delta opioid ligand pharmacophore and receptor models, *Biopolymers 51*, 426-439.

Naider, F., Arshava, B., Ding, F. X., Arevalo, E., and Becker, J. M. (2001) Peptide fragments as models to study the structure of a G-protein coupled receptor: the alpha-factor receptor of *Saccharomyces cerevisiae*, *Biopolymers 60*, 334-350.

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.

Naider, F., Estephan, R., Englander, J., Suresh babu, V. V., Arevalo, E., Samples, K., and Becker, J. M. (2004) Sexual conjugation in yeast: A paradigm to study G-protein-coupled receptor domain structure, *Biopolymers 76*, 119-128.

Nakayama, N., Miyajima, A., Arai, K., and Matsumoto, K. (1985) Nucleotide sequences of *STE2* and *STE3*, cell type-specific sterile genes from *Saccharomyces cerevisiae*, *EMBO J. 4*, 2643-2648.

Neumann, S., Krause, G., Chey, S., and Paschke, R. (2001) A free carboxylate oxygen in the side chain of position 674 in transmembrane domain 7 is necessary for TSH receptor activation, *Mol. Endocrinol. 15*, 1294-1305.

Opella, S. (1997) NMR and membrane proteins, *Nature Struct. Biol. 4*, 845-848.

Oren, Z., Ramesh, J., Avrahami, D., Suryaprakash, N., Shai, Y., and Jelinek, R. (2002) Structures and mode of membrane interaction of a short α helical lytic peptide and its diastereomer determine by NMR, FTIR, and fluorescence spectroscopy, *Eur. J. Biochem. 269*, 3869-3880.

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.

Pardi, A., Billeter, M., and Wüthrich, K. (1984) Calibration of the angular dependence of the amide proton-C-alpha proton coupling constants, ^{3}JHNα, in a globular protein, *J. Mol. Biol. 180*, 741-751.

Parrish, W., Eilers, M., Ying, W., and Konopka, J. B. (2002) The cytoplasmic end of transmembrane domain 3 regulates the activity of the *Saccharomyces cerevisiae* G-protein-coupled alpha-factor receptor, *Genetics 160*, 429-443.

Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997) Attenuated T-2 relaxation by mutual cancellation of DD coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12366-12371.

Picot, D., Loll, P. J., and Garavito, R. M. (1994) The X-ray crystal structure of the membrane protein prostaglandin H_2 synthase-1, *Nature 367*, 243-249.

Pogozheva, I. D., Lomize, A. L., and Mosberg, H. I. (1997) The transmembrane 7-alpha-bundle of rhodopsin: distance geometry calculations with hydrogen bonding constraints, *Biophys. J. 72*, 1963-1985.

Popot, J. L., and Engelman, D. M. (1990) Membrane protein folding and oligomerization: the two-stage model, *Biochemistry 29*, 4031-4037.

Popot, J. L., and Engelman, D. M. (2000) Helical membrane protein folding, stability, and evolution, *Annu. Rev. Biochem. 69*, 881-922.

Price, L. A., Kajkowski, E. M., Hadcock, J. R., Ozenberger, B. A., and Pausch, M. H. (1995) Functional coupling of a mammalian somatostatin receptor to the yeast pheromone response pathway, *Mol. Cell Biol. 15*, 6188-6195.

Price, L. A., Strnad, J., Pausch, M. H., and Hadcock, J. R. (1996) Pharmacological characterization of the rat A2a adenosine receptor functionally coupled to the yeast pheromone response pathway, *Mol. Pharmacol. 50*, 829-837.

Pumiglia, K., LeVine, H., Haske, T., Habib, T., Jove, R., and Decker, S. (1995) A direct interaction between G-protein beta gamma subunits and the Raf-1 protein kinase, *J. Biol. Chem. 270*, 14251-14254.

Rashid, A. J., O'Dowd, B. F., and George, S. R. (2004) Minireview: Diversity and complexity of signaling through peptidergic G protein-coupled receptors, *Endocrinology 145*, 2645-2652.

Rastogi, V. K, and Girvin, M. E. (1999) Structural changes linked to proton translocation by subunit c of the ATP synthase, *Nature 402*, 263-268.

Reddy, A. P., Tallon, M. A., Becker, J. M., and Naider, F. (1994) Biophysical studies on fragments of the α-factor receptor protein, *Biopolymers 34*, 679-689.

Reiersen, H., and Rees, A. R. (2001) The hunchback and its neighbours: proline as an environmental modulator, *Trends Biochem. Sci. 26*, 679-684.

Richardson, J. S., and Richardson, D. C. (1988) Amino acid preferences for specific locations at the ends of alpha helices, *Science 240*, 1648-1652.

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

Ridge, K. D., Lee, S. S., and Yao, L. L. (1995) In vivo assembly of rhodopsin from expressed polypeptide fragments, *Proc. Natl. Acad. Sci. U.S.A. 92*, 3204-3208.

Ridge, K. D., Ngo, T., Lee, S. S., and Abdulaev, N. (1999) Folding and assembly in rhodopsin. Effect of mutations in the sixth transmembrane helix on the conformation of the third cytoplasmic loop, *J. Biol. Chem. 274*, 21437-21442.

Rohrer, J., Benedetti, H., Zanolari, B., and Riezman, H. (1993) Identification of a novel sequence mediating regulated endocytosis of the G protein-coupled alpha-pheromone receptor in yeast, *Mol. Biol. Cell 4*, 511-521.

Russ, W. P., and Engelman, D. M. (2000) The GxxxG motif: a framework for transmembrane helix-helix association, *J. Mol. Biol. 296*, 911-919.

Sansom, M. S., and Weinstein, H. (2000) Hinges, swivels and switches: the role of prolines in signalling via transmembrane alpha-helices, *Trends Pharmacol. Sci*. *21*, 445-451.

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

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

Schwaiger, M., Lebendiker, M., Yerushalmi, H., Coles, M., Gröger, A., Schwarz, C., Schuldiner, S., and Kessler, H. (1998) NMR investigation of the multidrug transporter EmrE, an integral membrane protein, *Eur. J. Biochem. 254*, 610-619.

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

Sen, M., and Marsh, L. (1994) Noncontiguous domains of the alpha-factor receptor of yeasts confer ligand specificity, *J. Biol. Chem. 269*, 968-973.

Sen, M., Shah, A., and Marsh, L. (1997) Two types of alpha-factor receptor determinants for pheromone specificity in the mating-incompatible yeasts *S. cerevisiae and S. kluyveri*, *Curr. Genet. 31*, 235-240.

Senes, A., Gerstein, M., and Engelman, D. M. (2000) Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with β-branched residues at neighboring positions, *J. Mol. Biol*. *296*, 921-936.

Shah, A., and Marsh, L. (1996) Role of Sst2 in modulating G protein-coupled receptor signaling, *Biochem. Biophys. Res. Commun. 226*, 242-246.

Shichida, Y., and Imai, H. (1998) Visual pigment: G-protein-coupled receptor for light signals, *Cell. Mol. Life Sci. 54*, 1299-1315.

Sommers, C. M., and Dumont, M. E. (1997) Genetic interaction among the transmembrane segments of the G protein coupled receptor encoded by the yeast *STE2* gene, *J. Mol. Biol*. *266*, 559-575.

Sommers, C. M., Martin, N. P., Akal-Strader, A., Becker, J. M., Naider, F., and Dumont, M. E. (2000) A limited spectrum of mutations causes constitutive activation of the yeast alpha-factor receptor, *Biochemistry 39*, 6898-6909.

Sonnichsen, F. D., Van Eyk, J. E., Hodges, R. S., and Sykes, B. D. (1992) Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide, *Biochemistry 31*, 8790-8798.

Spiegel, A. M., and Weinstein, L. S. (2004) Inherited diseases involving g proteins and g protein-coupled receptors, *Annu. Rev. Med. 55*, 27-39.

Sprague, G. F., and Thorner, J. W. (1992) The Molecular and cellular biology of the yeast *Saccharomyces* gene expression, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

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

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

Stefan, C. J., Overton, M. C., and Blumer, K. J. (1998) Mechanism governing the activation and trafficking of yeast G protein-coupled receptors, *Mol. Biol. Cell 9*, 885-899.

Stevens, T. J., and Arkin, I. T. (1999) Are membrane proteins "inside-out" proteins?, *Proteins 36*, 135-143.

Strosberg, A. D., and Marullo, S. (1992) Functional expression of receptors in microorganisms, *Trends Pharmacol. Sci. 13*, 95-98.

Sui, H., Han, B. G., Lee, J. K., Walian, P., and Jap, B. K. (2001) Structural basis of water-specific transport through the AQP1 water channel, *Nature 414*, 872-878.

Suryanarayana, S., von Zastrow, M., and Kobilka, B. (1992) Identification of intramolecular interactions in adrenergic receptors, *J. Biol. Chem. 267*, 21991-21994.

Tamburro, A. M., Scatturin, A., Rocchi, R., Marchiori, F., Borin, G., and Scoffone, E. (1968) Conformational-transitions of bovine pancreatic ribonuclease S-peptide, *FEBS Lett. 1*, 298-300.

Tan, P. K., Davis, N. G., Sprague, G. F., and Payne, G. S. (1993) Clathrin facilitates the internalization of seven transmembrane segment receptors for mating pheromones in yeast, *J. Cell Biol. 123*, 1707-1716.

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

Uhlen, M., Moks, T. (1990) Gene fusions for purposes of expression: an introduction, *Methods Enzymol. 185*, 129-143.

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

Vinogradova, O., Sonnichsen, F., Sanders, C. R. 2nd (1998) On choosing a detergent for solution NMR studies of membrane proteins, *J. Biomol. NMR 11*, 381-386.

Walker, J., and Gaastra, W. (1988) *Methods in Molecular Biology, New Protein Techniques,* 3rd ed., pp 427-440, Humana Press Inc., Clifton, New Jersey.

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

Watson, S., and Arkinstall, S. (1994) *The G-protein linked receptor facts book*, Academic Press, London.

Weiner, J. L., Guttierez-Steil, C., and Blumer, K. J. (1993) Disruption of receptor-G protein coupling in yeast promotes the function of an SST2-dependent adaptation pathway, *J. Biol. Chem. 268*, 8070-8077.

Wess, J. (1997) G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition, *FASEB J. 11*, 346-354.

White, S. H., and Wimley, W. C. (1999) Membrane protein folding and stability: physical principles, *Annu. Rev. Biophys. Biomol. Struct. 28*, 319-365.

Wigley, W. C., Vijayakumar, S., Jones, J. D., Slaughter, C., and Thomas, P. J. (1998) Transmembrane domain of cystic fibrosis transmembrane conductance regulator: Design, characterization, and secondary structure of synthetic peptides m1-m6, *Biochemistry 37*, 844-853.

Williams, K. A., and Deber, C. M. (1991) Proline residues in transmembrane helices: structural or dynamic role?, *Biochemistry 30*, 8919-8923.

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

Wishart, D. S., and Sykes, B. D. (1994) Chemical shifts as a tool for structure determination, *Methods Enzymol. 239*, 363-392.

Wu, C. S., Ikeda, K., and Yang, J. T. (1981) Ordered conformation of polypeptides and proteins in acidic dodecyl sulfate solution, *Biochemistry 20*, 566-570.

Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, *Chapter 2: NMR of amino acid residues and mononucleotides*, pp 13-25, John Wiley and Sons, Inc., New York.

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.

Yeagle, P. L., Choi, G., and Albert, A. D. (2001) Studies on the structure of the G-protein-coupled receptor rhodopsin including the putative G-protein binding site in unactivated and activated forms, *Biochemistry 40*, 11932-11937.

Yeagle, P. L., Danis, C., Choi, G., Alderfer, J. L., and Albert, A. D. (2000a) Three dimensional structure of the seventh transmembrane helical domain of the G-protein receptor, rhodopsin, *Mol. Vis. 6*, 125-131.

Yeagle, P. L., Salloum, A., Chopra, A., Bhawsar, N., Ali, L., Kuzmanovski, G., Alderfer, J. L., and Albert, A. D. (2000b) Structures of the intradiskal loops and amino terminus of the G-protein receptor, rhodopsin, *J. Pept. Res. 55*, 455-465.

Yohannan, S., Faham, S., Yang D., Whitelegge, J. P., and Bowie, J. U. (2004) The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors, *Proc. Natl. Acad. Sci. U. S. A. 101*, 959-963.

Zhang, L., DeHaven, R. N., and Goodman, M. (2002) NMR and modeling studies of a synthetic extracellular loop II of the *k* opioid receptor in a DPC micelle, *Biochemistry 41*, 61-68.

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

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.

Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T., and Engelman, D. M. (2000) Interhelical hydrogen bonding drives strong interactions in membrane proteins, *Nature Struct. Biol. 7*, 154-160.