INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

Biosynthesis, Purification and Biophysical Analysis of Domains of the *Saccharomyces cerevisiae* Alpha-factor Receptor

by

ENRIQUE ARÉVALO

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.

2002

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.

May 28, 2002
Date

Fred Naider
Chair of Examining Committee

July 12th, 2002
Date

Lesley Davenport
Executive Officer

Ruth Stark
Robert Birge
Robert Britton
Supervisory Committee

The City University of New York

Abstract

BIOSYNTHESIS, PURIFICATION AND BIOPHYSICAL ANALYSIS OF DOMAINS OF THE *Saccharomyces cerevisiae* ALPHA-FACTOR RECEPTOR

By

Enrique Arevalo

Mentor: Professor Fred Naider

The α-factor receptor (Ste2p) is an essential protein for the sexual conjugation of the yeast *Saccharomyces cerevisiae*. It belongs to the family of G protein-coupled receptors, and it is believed to share a common structural motif with the other members of the family; seven transmembrane segments, that are α-helices within the membrane, connected by alternating extracellular and intracellular loops.

In the current thesis project, studies were performed on the M5I3M6 region that corresponds to transmembrane domains 5 and 6 and the connecting third intracellular loop to complement the biophysical studies carried out on single transmembrane regions and the intracellular and extracellular loops of Ste2p (Reddy et al., 1994; Arshava et al., 1998; Xie et al., 2000; Ding et al., 2001). Molecular biology techniques were used to express the M5I3M6 and M6 segments of Ste2p fused to a leader protein in *E. coli*. These polypeptides were isolated and purified in 10-100 mg quantities. The fusion proteins were characterized by gel electrophoresis, mass spectrometry and amino acid analysis. Their secondary structures were investigated in membrane mimetic environments including trifluoroethanol:water mixtures, SDS micelles, and lipid vesicles using circular dichroism. The structures of M5I3M6 and M6 linked to a solubilizing leader peptide were partial helices.

The CD studies initiated by Xie and others on single transmembrane domains of Ste2p were extended by applying nuclear magnetic resonance (NMR) spectroscopy to investigate the M6 region. NMR experiments were performed in $CDCl_3$:CD_3OH:H_2O (4:4:1) to obtain a high resolution structure of the sixth transmembrane region. Different models were calculated for a wildtype peptide sequence and for an analog containing a Pro to Leu mutation that causes Ste2p to become constitutively active. The helical structure found in these membrane-spanning peptides is consistent with the findings published by Xie et al. (2000) and provides biophysical verification of hydropathy predictions on intact receptor.

This thesis is dedicated to my dear parents Jorge and Zoilita, my siblings Otto, Iracema, Patricio and Joaquin, and my beautiful wife Veronica.

Acknowledgments

I would like to thank my mentor Dr. Fred Naider for his constant and tireless support. He proposed to me this challenging thesis project and along the process I was able to learn innumerable techniques and scientific details that taught me to think and work as a scientist. I want to thank him for his permanent encouragement during the difficult phases of this project and his valuable friendship.

I also would like to thank the committee members for their precious time and their insightful comments along this research project. I want to thank Dr. Charlotte Russell, for her charming introduction to biochemistry during my first steps of the doctoral studies. I want to thank immensely Dr. Probal Banerjee for his help and suggestions during the molecular biology stage of my thesis. To Dr. Robert Bittman, Dr. Peter Lipke, and Dr. Ruth Stark, I express my deep appreciation for the valuable orientation I obtained during their lectures.

The research environment nurtures the soul of the researcher. I was fortunate to participate in a research laboratory, including the Chemistry Department of the College of Staten Island, where the surroundings could not be more propitious to work. I want to thank to all members of Dr. Naider's laboratory, especially to Dr. Boris Arshava, for his enormous help throughout my thesis work. His infinite patience and his clever suggestions were always very highly appreciated. I want to personally thank Dr. Sanjay Khare for his friendship and suggestions during my introduction to HPLC purification. I also want to thank Dr. Fa-Xiang Ding for sharing with me his expertise in FT-IR spectroscopy. I want to thank David Schreiber for contributing to my research with his

friendship and sense of humor, which is sometimes a necessary ingredient in a work environment.

I am grateful to Dr. Hsin Wang, the NMR manager of the College of Staten Island, for his valuable support in setting some NMR experiments and for teaching me the NMRView software during the NMR analyses. I also want to thank Dr. Xiaomin Yang for her contribution in my understanding of the molecular modeling software.

My gratitude also goes to Jennifer Madeo with whom we started the project of expression of the double domain of Ste2p in bacterial system. Her keen mind was always a stimulating force at the beginning of my thesis work.

I am thankful to the OMRDD Fellowship in the CSI/IBR center for Developmental Neuroscience for the financial support during the last three years of my thesis and for giving me the opportunity to broaden my knowledge in the fascinating field of neuroscience.

I am certainly grateful to the support, constant care and love, of every member of my family. My eternal gratitude to my parents Jorge and Zoilita for teaching me to always strive beyond my own capacities. I am thankful to my brothers and sister for their moral support whenever I needed.

Finally, all my gratitude goes to a woman that is the center and the reason of my life, my wife Veronica. Veronica provided me with the strength, infinite support necessary during this long journey. To her all my love and eternal gratitude.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Å	Angstrom
BPTI	bovine pancreatic trypsin inhibitor
BR	bacteriorhodopsin
BSA	bovine serum albumin
CCC	covalently closed circular
CD	circular dichroism
CFTR	cystic fibrosis transmembrane conductance regulator
CH_3OH	methanol
CNBr	cyanogen bromide
CNS	crystallography & NMR System
DMPC	1,2-dimyristoyl-*sn*-glycero-3-phosphocholine
DMPG	1,2-dimyristoyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)]
EDTA	ethylenediaminetetraacetic
ESI-MS	electrospray ionization mass spectrometry
FT-IR	Fourier transformed infrared
GPCR	G protein-coupled receptor
GuHCl	guanidinium hydrochloride
HFIP	1,1,1,3,3,3-hexafluoro-2-propanol
HPLC	high-performance liquid chromatography
Ile	isoleucine
IPTG	isopropyl-β-D-thiogalactopyranoside
^3J	spin-spin coupling constant

K_D	dissociation constant
KHz	kilohertz
Ki	affinity binding constant
LB	Luria-Bertani
LUVETs	large unilamellar vesicles
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser enhanced spectroscopy
ONPG	o-nitrophenyl-β-D-galactopyranoside
PAGE	polyacrylamide gel elctrophoresis
PBS	phosphate-buffer saiine
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride
ppm	parts per million
R^2	R squared
RPM	revolutions per minute
SD	synthetic drop-out
SDS	sodium dodecylsulfate
TBS	Tris-buffered saline
TBST	Tris-buffered saline + Tween-20
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
TOCSY	total correlation spectroscopy

URA	uracil
YPD	yeast extract-peptone-dextrose

LIST OF TABLES

LIST OF FIGURES

Chapter I: Introduction

Intercellular signaling

Plasma membranes of all cells contain a variety of proteins: peripheral proteins, integral proteins, glycoproteins, and lipoproteins. One class of integral membrane proteins is G protein-coupled receptors (GPCRs), which have specific binding sites for extracellular signaling molecules. The outer surface of a cell is in permanent contact with its environment, other cells, extracellular fluids, solutes, nutrients, hormones, pheromones, neurotransmitters, and antigens. When any of those extracellular molecules binds to its specific GPCR, the receptor protein transmits the signal to the transducer G-protein that carries the message internally. In yeast, sexual conjugation (mating) is a process that involves intercellular signaling to initiate conjugation and it is an important step in the development of this eukaryotic cell.

Saccharomyces cerevisiae: A model for intercellular signal transduction

Saccharomyces cerevisiae (yeast) is a very good model for intercellular signal transduction, because it does not have the complexity of other eukaryotic cells and a wealth of biochemical and genetic information is already available on this cell type. Yeast exists as two different haploid cell types (**a** cells and α cells), which conjugate to

form diploid cells (**a**/α cells). Conjugation is initiated in response to the reciprocal mating pheromone (*i.e.*, α cells release α factor and **a** cells have receptors for the α factor and vice versa). An early event triggered by the yeast pheromones is the biosynthesis of cell surface proteins (agglutinins). This is followed by growth arrest of both haploids in the G_1 phase of the cell division cycle (Sprague et al., 1992).

Using genetic techniques, several proteins have been identified to be important in the signal transduction process. Some of these are specific for one haploid and some are specific for the diploid (Table 1).

Table 1*
Mating Type Specific Genes in *S. cerevisiae*

		Expression in cell type		
Gene set	Genes	**a**	α	**a**/α
α-specific genes (α*sg*)	*STE3, MFα1, MFα2*	–	+	–
a-specific genes (**a***sg*)	*STE2, STE6, MFA1, MFA2, BAR1*	+	–	–
haploid-specific genes (hsg)	*STE4, STE5, STE12, STE18, MATα1, MATα2, FUS1, GPA1, Ty1, RME1*	+	+	–
diploid-specific genes (*ssg*)	*SGA1, SPO11, SPO13, SPS1, SPS2, SPS3, SPS4, SPS100*	–	–	+

*Taken from Herskowitz I., Rine J., Strathern J., The Molecular and Cellular Biology of the Yeast *Saccharomyces* Gene Expression, Cold Spring Harbor Lab. Press, 1992, page

586.

It has been demonstrated that *STE2* and *STE3* gene products alone are responsible for the transmission of the signal from their respective pheromone to the interior of the respective cell (Blumer et al., 1988; Blumer et al., 1991). In addition to an increase in the tendency of the cells to agglutinate and arrest at G_1, the pheromones also induce morphogenic changes, where the cells present directional growth toward each other (projection formation) resulting in cell fusion and nuclear migration to the center of the fused cell. Ultimately a diploid cell with a fused nucleus is formed. Depending on the nutrient characteristics of the medium the zygote can either commence meiosis and sporulation to generate α and **a** cell types under low nutrient conditions, or undergo asexual reproduction to form daughter diploid buds in media containing a sufficient amount of nutrients (Sprague et al., 1992) (Figure 1).

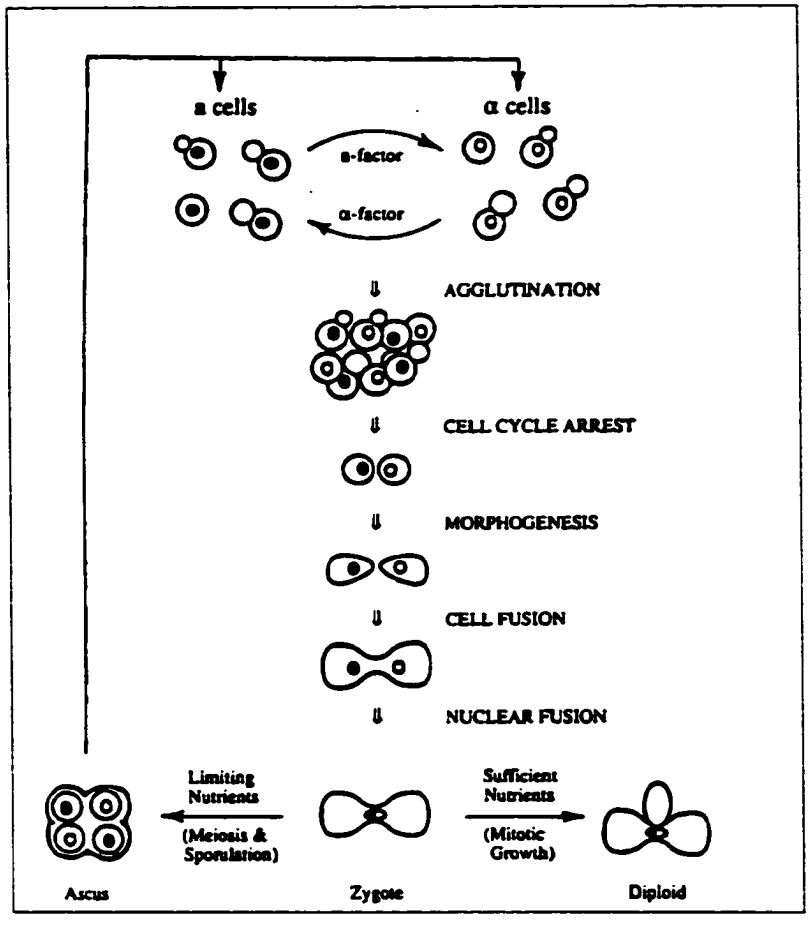

Figure 1: Representation of the mating process in *S. cerevisiae**

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

G protein coupled-receptors

STE2 and *STE3* genes code for the α-factor receptor and the a-factor receptor in a cells and α cells, respectively (Nakayama et al., 1985). These two receptors belong to a family of GPCRs found in mammals and other organisms. GPCRs mediate different kinds of signals: light (rhodopsin), odorants, peptide hormones and neurotransmitters

(β-adrenergic receptors). They amplify these extracellular signals by activating an effector system (G proteins) that produces an intracellular signal (Dohlman et al., 1991; Dohlman et al., 2001).

Determination of the structure of transmembrane proteins has proven to be intrinsically difficult for several technical reasons such as obtaining X-ray grade crystals and the large size and slow motions of the protein-lipid complex in NMR experiments. For these reasons fewer structures of transmembrane proteins have been determined by X-ray crystallography in comparison to globular proteins. Several cases where success has been achieved include the photoreaction center (*Halobacteria halobium*) (Kataoka et al., 1981), bacterial and mitochondrial cytochrome c oxidases (Soulimane et al., 2000), mitochondrial bc1 complex (Iwata et al., 1998), alpha-hemolysin (Song et al., 1996), and the potassium channel (Doyle et al., 1998). The structure of bacteriorhodopsin has also been determined at 1.55 Å (Luecke et al., 1999). This proton pump, although not a GPCR, has been extensively used as a model for the GPCRs because of its homology to rhodopsin. Recently the structure of the first GPCR, rhodopsin, has been elucidated by X-ray crystallography. The structure has been determined at 2.8 Å resolution (Palczewski et al., 2000).

The structures of bacteriorhodopsin and rhodopsin superimpose fairly well with the theoretical models built using hydropathy analysis and the assumption that the membrane spanning domains are helical (Baldwin et al., 1997; Herzyk et al., 1998; Bourne et al., 2000). These structures result in the working model for all GPCRs

wherein there are seven transmembrane domains and the amino and carboxy-terminal are located extracelluarly and in the cytoplasm, respectively (Dohlman et al., 1991; Cramer et al., 1992; Baldwin, 1993, Herzyk et al., 1995; Baldwin et al. 1997; Herzyk et al., 1998; Luecke et al., 1999; Palczewski et al., 2000). Using the currently accepted topology of all model GPCRs, and the hydrophobicity index based on the primary sequence of Ste2p a cartoon depicting the general location of the principal domains is depicted in Figure 2 (Reddy et al., 1994).

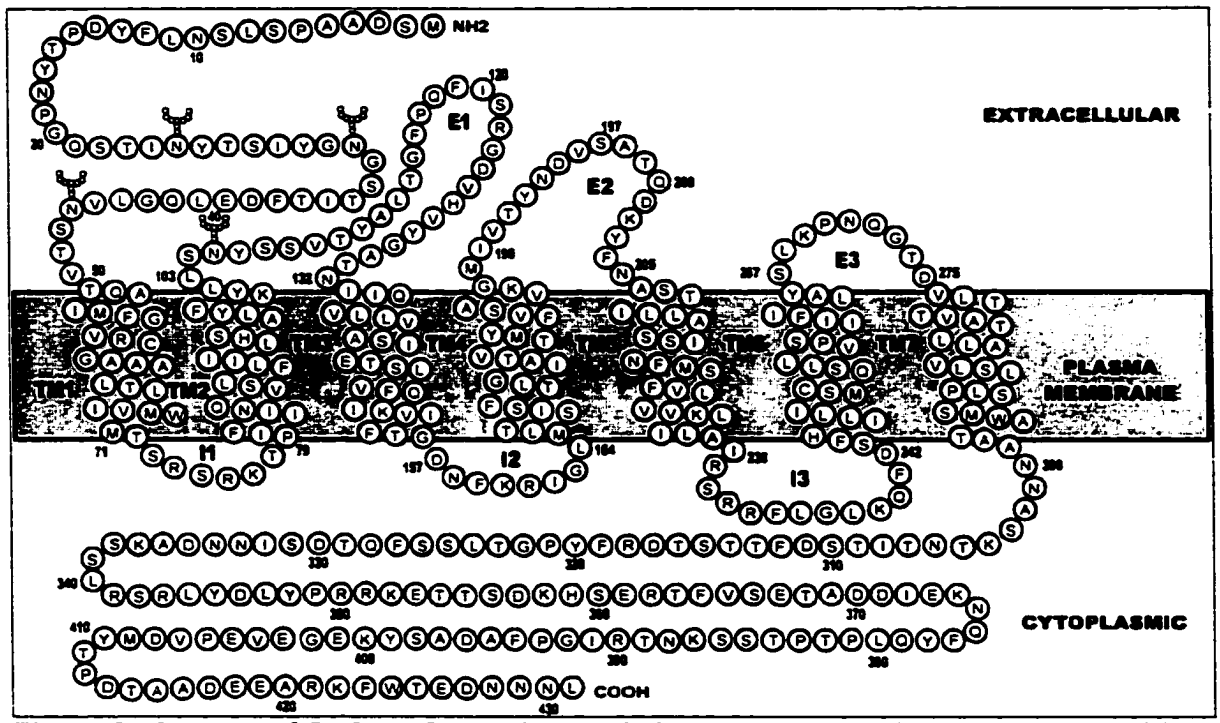

Figure 2: Cartoon of Ste2p a G protein-coupled receptor. Domains are represented by E, extracellular loops; I, intracellular loops and TM, transmembrane domains.

This overall arrangement agrees fairly well with topology models derived by Sommers et al. (1997) and Dube et al. (2000). Differences arise in the position of certain amino acids at the membrane interface of the different transmembrane domains.

For instance, Sommers et al. (1997) position proline 79 at the intracellular interface of transmembrane 2 whereas Reddy and others place threonine 78 in that location and Dube et al. (2000) place isoleucine 80 at the membrane interface. This minor difference results in a discrepancy in the placement of a tyrosine residue in the different models. Specifically, (Tyr^{101}) at the extracellular side of transmembrane 2 is in the membrane interface in certain models whereas another hydrophobic residues (Leu^{102}) is at the membrane interface in others. It has been documented that tryptophan and tyrosine residues prefer to be located in the membrane interface in order to satisfy both Van der Waals interactions of the aromatic group and the polar interactions of the functional polar groups of their side chains (Yau et al., 1998; De Planque et al., 1999). The same type of discrepancy of one or two residues at the membrane interface is observed in most of the transmembrane domains. The biggest difference between Reddy's and Dube's topology is located in M4, where 6 residues considered to be in the transmembrane region by Reddy are located in the second extracellular loop (E2) in Dube's structure. Therefore, Dube's E2 domain is 5 residues longer than Reddy's proposal. The discrepancy in number of residues embedded within the membrane carries with it the consequence that the transmembrane domain will be forced to have a tilt angle with respect to the plane of the membrane in order to avoid exposing some hydrophobic residues. Therefore, one proposal predicts some tilt in a specific transmembrane segment whereas the other predicts almost a perpendicular positioning of the membrane spanning segment. The transmembrane domain tilt also changes the

possible contact points among inter helical side chains.

The budding yeast *Saccharomyces kluyveri* is closely related to *S. cerevisiae*. It makes an α-factor peptide (k-α-f) that, like the α-factor for *S. cerevisiae* (c-α-f) is 13 amino acids residues long but it differs in sequence at five amino acid positions. The k-α-f receptor has been cloned and sequenced and is 50% identical to *STE2* at the amino acid level (Marsh, L. 1992). But, there is no sequence similarity of *STE3* to either *STE2* or to the k-α-f receptor. Using chimeric receptors of *S. cerevisiae* and *S. kluyveri* (Sen et al., 1994), it was determined that the regions relevant for ligand specificity include the first and third extracellular loop and all transmembrane domains except for the fourth and fifth. Although an early study concluded that the N-terminal domain of Ste2p does not seem to be required for ligand specificity (Sen et al., 1994), more recent results suggest that residues 47-49 do indeed contain an important determinant of binding specificity (Sen et al., 1997). Another report reinforces the scheme that the first and sixth transmembrane domains of the yeast receptor play an important role in ligand binding, ligand specificity and receptor activation (Abel et al., 1998). Dube and Konopka (1998) identified possible contact sites between polar residues Gln^{253} from transmembrane 6 and Ser^{288} and Ser^{292} from transmembrane 7 that are important for regulation of the function of the Ste2p receptor.

The third intracellular loop plays a key role in the interaction with the G protein (Blumer et al., 1991; Clark et al., 1994; Bukusoglu et al., 1996). This putative cytoplasmic loop appears to have an amphipathic character that favors formation of a

helical structure at the membrane interface (Blumer et al., 1991). The amphipathic character of the third intracellular loop has also been observed in other GPCRs like the m2 muscarinic receptor (Kostenis et al., 1997) and the platelet-activating factor receptor (Carlson et al., 1998). This implied secondary structure would explain the interaction of this short sequence with both the membrane and a protein (in this case a G protein). However, the only physical evidence available about the structure of the I3 loop indicates that it behaves as a random coil in water (Arshava et al., 1998). It is possible that under the influence of phospholipid head groups it behaves as predicted by the helical wheel, or that the anchoring transmembrane regions (M5 & M6) impose a structure on this internal loop.

The hydropathy analysis of the yeast Ste2p predicts the presence of a proline (258) in the center of the sixth transmembrane region (Figure 2). The prediction corroborates the fact that proline is present in the sixth transmembrane region in more than 90% of all GPCRs (Baldwin, 1993; Stefan et al., 1998). A molecular biological screen used to identify constitutively active receptors uncovered mutations at Pro^{258} of Ste2p. Receptor with a Pro258Leu mutation exhibited 45% activation in the absence of pheromone compared to the maximum level of activation observed in wild-type cells stimulated with α-factor. A double mutation of Pro258Leu and Ser259Leu further activated the signal transduction to over 90% of the maximum level (Konopka et al., 1996). Another report that studied the mechanism involved in the activation and trafficking of GPCRs by analyzing constitutively active receptors that had substitutions

of the conserved proline in the sixth transmembrane segment of Ste2p and Ste3p, concluded that this constrained amino acid plays an important role in the inactive conformation of the receptors (Stefan et al., 1998). Although proline is known to be helix breaker in globular proteins, it has been found that in the membrane environment a proline residue can be supportive of the helical structure and even more, it can stabilize the helical conformation at higher temperatures (Li et al., 1996). These results suggest the importance of the structure of the receptor in the transduction of signal. Change in conformation of the GPCR upon interaction with its pheromone is still subject to demonstration and such information will aid in understanding the mechanism of signal transduction by Ste2p and possibly of all members of this family.

The fact that the third intracellular loop is perfectly conserved in the α-factor receptor of both *Saccharomyces cerevisiae* and *Saccharomyces kluyveri*, would indicate that this loop is not associated with ligand specificity but serves only to couple the G protein (Sen et al., 1994). It was also observed that the removal of the G protein from its coupled receptor in different cell lines resulted in a decreased affinity of the receptor for its ligand (Shah et al., 1996). This finding indicates that the coupling of the G protein and the third intracellular loop of the receptor may induce a conformation of the receptor which is optimum for binding, thereby increasing the affinity for the ligand. These results support the working hypothesis that the pheromone receptor acts as a negative regulator of the heterotrimeric G protein in the sense that the receptor sequesters a limiting pool of G proteins in the inactive state (Stefan et al., 1998). Dosil

and others (1998) suggested the same possibility after working with dominant negative mutants that compete with wild type receptors for the G proteins. In a recent report Dosil et al. (2000) included the C-terminus of the yeast receptor in the receptor/G protein complex that forms prior to pheromone activation.

Mutational analyses on adrenergic receptors (members of the GPCR family) reveal that there is an important interaction between the first and seventh transmembrane regions (Suryanarayana et al., 1992). Another study on the muscarinic acetylcholine receptor found that besides the importance of M1 and M7 interaction for the proper folding of the receptor, the seventh helix must be rotated approximately 30° counterclockwise with respect to the orthogonal axis of the membrane plane (Liu et al., 1995) (contrary to what Baldwin predicted). Further insights into the orientation of M1 and M7 come from the observation that a number of residues near the extramembraneous face of M7 of several GPCRs (e.g., Tyr^{426} of the m2 muscarinic receptors) have been shown to be essential for high ligand affinity (Liu et al., 1995). The overall arrangement of the GPCR transmembrane helices seems to be in a counterclockwise manner as viewed from the extracellular side of the membrane (Baldwin, 1993; Liu et al., 1995; Wess, 1997). This arrangement concurs with the recent arrangement found in rhodopsin (Palczewski et al., 2000). In the case of Ste2p, Dube et al. (2000) reported a specific interaction between residues Val^{223} and Leu^{247} located in transmembrane region fifth and sixth, respectively. This interaction helps to orient those domains with respect to each other.

Mutations on the N-terminal domain of the *S. cerevisiae* α-factor receptor showed lowered mating efficiency, whereas mutations in the transmembrane domains were defective for mating (Konopka et al., 1991). However, this study indicated that the lowered mating efficiency of the N-terminal mutants could be due to a reduction of the number of receptors on the cell surface. The main conclusion is that the transmembrane domains of GPCRs are essential for both pheromone binding and/or signal transduction (Konopka et al., 1991).

It was reported earlier that truncation of the C-terminus of Ste2p or mutations of residues in this portion of the receptor does not affect pheromone binding (Konopka et al., 1988), and, that the C-terminus does not interact with the G proteins (Konopka et al., 1996). These two results are in conflict with the recent data presented by Dosil and others (2000). At present, the accepted functions of the C-terminus of Ste2p are to regulate receptor number and adaptation to α-factor. One possible mechanism, for pheromone adaptation, seems to be endocytosis (Hicke et al., 1998), because it was found that the receptor undergoes rapid endocytosis following exposure to the α-factor. After internalization, new receptors are localized to the cell surface and are able to bind alpha factor. However, they "remain refractory to the pheromone action" (Dohlman et al., 1991). One possible reason is that newly synthesized receptors are still uncoupled from the intracellular effector(s) (Dohlman et al., 1991). The mechanism of internalization involves phosphorylation of any or all of the potential phosphorylating sites (Ser or Thr) within the C-terminal domain, particularly serine residues in the

SINNDAKSS region of the receptor tail (Hicke et al., 1998). Ubiquitination of lysine residues in the receptor C-terminus is also part of the receptor uptake process triggered upon binding to the external ligand (Hicke et al., 1998). The process of desensitization upon exposure to the pheromone involves additional intracellular gene products such as Sst2p, Cdc36p and Cdc39p (Dohlman et al., 1991; Blumer et al., 1991; Dohlman et al., 1996). Some studies found that Sst2p acts to stabilize the complex of G protein with the inactivated receptor (Shah et al., 1996), but the function of Sst2p is independent of the interaction with the C-terminus (Konopka et al., 1988), and this protein (Sst2p) is one of the key regulators for desensitization (Dohlman et al., 1996). Sst2p belongs to the family of regulators of G protein signaling (RGS) (Dohlman & Thorner, 1997) and it becomes phosphorylated on Ser^{539} upon stimulation of the Ste2p receptor with the α-factor ligand (Runyan, et al., 1999). Interestingly Runyan and others (1999) suggested that the phosphorylation of Sst2p in yeast could be carried out by MAP kinase in a pheromone dependent manner. It seems that Sst2p down regulates G protein signaling by binding the Gα subunit and serving as a GTPase activating protein (Dohlman & Thorner, 1997).

Based on the genetic results outlined above and the known structures of bacteriorhodopsin and rhodopsin, a picture of some structural features of the receptor is emerging. Residues important for binding and signal transduction have been identified and some possible contact points within Ste2p have been discerned. Although rhodopsin is an integral membrane protein and belongs to the GPCR family, its

activation differs from that of GPCRs, such as Ste2p, that responds to ligands as extracellular signals. Furthermore, even point mutation of a GPCR in one amino acid may affect several receptor characteristics including signaling, ligand binding, and ligand affinity, receptor interaction with the G protein, and receptor desensitization simultaneously. Therefore, in order to determine the molecular details of Ste2p structure and function more direct biophysical and biochemical studies are necessary.

G proteins

G proteins belong to a family of heterotrimeric proteins composed of three subunits α, β, and γ. In yeast, *GPA1* [also named *SCG1*, *CDC*, and *DAC1* (Dohlman et al., 1991)] and *GPA2* genes were identified to encode proteins of 472 and 449 residues, respectively, that were potentially related to some of the subunits of a mammalian G protein. Of the two potential G protein subunits, only *GPA1* encodes for a protein that is essential for the signal transduction, because *gpa2* deletion mutants show no variation in response to the pheromone (Blumer et al., 1991). *GPA1* is expressed only in the two haploid cell types, whereas, *GPA2* gene is expressed in all three cell types and it is not involved in signal transduction (Nakafuku et al., 1988, Blumer et al., 1991). It was found that *GPA1* deletion mutants are constitutively activated, thereby resembling cells responding to the peptide mating factor. The *GPA1* gene product exhibits 45% homology to vertebrate G_{α}, and it is believed that it mediates the interaction of pheromone and receptor through binding and hydrolysis of GTP. Thus Gpa1p is the

putative yeast Gα subunit.

Two other gene products were found to be important in the normal function of Gpa1p coupled to the receptor. *STE4* and *STE18* gene products are homologues to β and γ subunits of vertebrate G proteins. They encode 432 and 110-residue proteins, respectively. Deletion or mutation of these genes causes sterility and lack of response to the pheromone (Blumer et al., 1991). Research on mammalian G_α subunits suggests that they can interact with the yeast βγ complex to prevent activation of the pheromone signal, but they alone cannot associate with Ste2p to transduce signal (Blumer et al., 1991). In the same fashion as the mammalian homologues, the yeast βγ subunits complex the G_α protein and they are released from the trimeric complex upon binding of GTP (Liri et al., 1998). It was originally proposed that the βγ subunits anchor the α subunit to the membrane, facilitating the interaction with the receptor (Blumer, et al., 1991). Conversely, other results suggest that Gpa1p may help to anchor $G_{\beta\gamma}$ to the plasma membrane (Hirschman et al., 1997).

A current model of G protein action based on results from membrane fractionation and genetic analysis of yeast concludes that the $G_{\beta\gamma}$ complex plays a positive role in the pheromone response pathway, that it is negatively regulated by the G_α subunit, and positively regulated by GPCR. Stable association of the $G_{\beta\gamma}$ complex with plasma membrane requires G_α but it does not require the receptor (Hirschman et al., 1997). However, a review (Wess, 1997) establishes that the βγ complex of the mammalian cells directly contacts the m2 muscarinic receptor protein besides

contacting the G_α subunit. Moreover, the specificity of the G_α subunit for the receptor results from the interaction of the three components of the heterotrimer with the membrane protein. This might not necessarily be the case in the yeast pheromone receptor. Following GPCR activation by the pheromone, dissociation of G_α from $G_{\beta\gamma}$ releases this complex causing transfer of the signal to the interior of the cell. The internalization of the α-factor receptor as a consequence of the desensitization process described earlier, does not involve Ste4p redistribution within the cell (Hirschman et al., 1997). However, G_β becomes phosphorylated in a pheromone-dependent manner (Cole & Reed, 1991). *In vitro* experiments raised the possibility that binding of $G_{\beta\gamma}$ to Raf-1 (a serine-threonine protein kinase) could play a role in the mitogenic signal pathway, and this interaction would follow the release of the complex from the membrane bound state (Pumiglia et al., 1995).

X-ray crystallographic structures have been solved for the heterotrimeric $G_{i\alpha1}\beta\gamma$ (rat $G_{i\alpha1}$ and bovine βγ heterodimer) (Wall et al., 1995) and for the transducin $G_\alpha\beta\gamma$ from rod cells (Lambright et al., 1996). They indicate that G_α subunits are composed of two domains: the core domain and the α-helical domain. The guanine nucleotide is bound within the cleft between the core and the α-helical domains (Neer, 1995; Sprang, 1997). The C-terminus of G_α is believed to interact with the receptor and the conformation of this region of Gpa1p changes upon interaction with the **a**- and α-factor receptors (Kallal et al., 1997). Considering the fact that the same gene is expressed in both haploid cell types and performs the same function it is likely that similar protein-

protein interactions occur between Gpa1p and Ste2p or Ste3p.

Biophysical studies of membrane proteins

As mentioned before, the difficulty in performing biophysical studies of large membrane proteins in lipids or detergent environment is associated with their high molecular weight, their low levels of solubility and irreversible association (aggregation), that leads to sample heterogeneity at the concentrations required for NMR studies. These technical difficulties hinder progress in the structural determination of large integral membrane proteins. One approach to overcome this problem is to study the structure of isolated domains of large membrane proteins in order to infer the tertiary structure from the information gained on the different domains. This approach has been applied to study transmembrane domains of the cystic fibrosis transmembrane conductance regulator (CFTR) (Wigley et al., 1998). Hunt et al. (1997) studied biophysically the structure of seven transmembrane domains of bacteriorhodopsin (BR) and obtained important information about the tendency of each membrane segment to form a particular secondary structure. The secondary structure of individual transmembrane domains of the nicotinic acetylcholine receptor was also reported (Corbin et al., 1998). The sarcoplasmic reticulum Ca^{+2}-ATPase contains 10 putative transmembrane segments. Conformational studies of the sixth transmembrane domain that is believed to be associated with the ATPase calcium binding sites was reported by Soulié and others (1999). The structure of individual transmembrane

regions of rhodopsin has also been reported in the literature. The structures of the sixth transmembrane helix (Chopra et al., 2000) and of the seventh membrane-spanning region (Yeagle et al., 2000a) were published prior to the report of the crystal structure of the full GPCR. The secondary structure of the N terminus and 3 interhelical loops of rhodopsin were also studied by NMR (Yeagle et al., 2000b) but no comparison can be made with the crystal structure results because of the lack of structural details from the crystallographic data. Single domains of the Ste2p receptor were also studied biophysically (Reddy et al., 1994; Arshava et al., 1998, Xie et al., 2000; Ding et al., 2001b). Other examples of a polytopic transmembrane proteins studied using isolated domains are the elongation factor Tu (Cetin et al., 1998) and the pore-forming domain of *Bacillus thuringiensis* insecticidal CryIIIA δ-endotoxin (Gazit & Shai, 1995).

In vivo assembly of rhodopsin from expressed polypeptides (Ridge et al., 1995; Ridge et al., 1996; Ridge et al., 1999), refolding of bacteriorhodopsin from expressed fragments (Luneberg et al., 1998; Marti, 1998) and *in vivo* assembly of the Ste2p receptor from expressed regions (Martin et al., 1999) all provide some support for the biophysical studies of isolated domains. In these cases the reconstituted fragments all assemble into functional proteins despite the fact that they lack the covalent bond in some interconnecting loops. The results from these studies indicate that the conformation of some transmembrane domains is specified by the interactions with other helices and that some stable helices serve as templates to induce the final conformation of the integral membrane protein (Luneberg et al., 1998).

Solid-state NMR and solution NMR analyses of small proteins in membranes and micelles are becoming more important due to the possibility of isotopic labeling, high expression of the proteins or fragments of a protein using molecular biology techniques, and implementation of different membrane mimetic lipids and detergents as the environment for the study (Sanders et al., 2000; Baleja, 2001). Complete resolution of the solid-state NMR spectrum of uniformly ^{15}N-labeled of the major coat protein of fd bacteriophage in phospholipid bilayers, was accomplished by Marassi et al., (1997). *E. coli* diacylglycerol kinase, a 121 amino acid protein, was used to develop methods to study an integral membrane protein by solution NMR (Vinogradova et al., 1997). The Pf1 filamentous bacteriophage coat protein, a 46-residue protein, was also characterized by multidimensional solution NMR experiments in sodium dodecylsulfate (SDS) micelles (Schiksni et al., 1986). M13, another bacteriophage coat protein, was studied using multinuclear multidimensional NMR when complexed with SDS detergent micelles as well (van de Ven et al., 1993). The 23-residue antibiotic, magainin, was studied by Gesell and others (1997) in SDS micelles, dodecylphosphocholine micelles, and trifluoroethanol/water solutions. Myristoyllysophosphatidylglycerol (MPG) was used to study the structure of the bacteriophage Ike major coat protein by solution NMR (Williams et al., 1996). Solution NMR experiments were conducted in isotropic solutions of phospholipid bicelles to study the structure of mastoparan *vespula lewissi* (Vold et al., 1997). Structure determination by NMR using organic solvents as membrane mimetics has also been widely carried out. For example, the multidrug

transporter EmrE was studied in chloroform/methanol/water (6:6:1) (Schwaiger et al., 1998). A very similar solvent system chloroform/methanol/water (4:4:1) was used in the structure determination of subunit c of the F_1F_0 ATP synthase (Rastogi et al., 1999). Other membrane proteins that were studied in organic solvents are the N-terminus of bacteriorhodopsin (Pervushin et al., 1994) and the integral membrane colicin E1 immunity protein (Taylor et al., 2000). Czerski and others (2000) published a method of studying hydrogen-deuterium exchange of membrane proteins in lipid bilayer where the exchange is carried out in the lipid-protein complex in the presence of H_2O and/or D_2O followed by lyophilization and resuspension of the exposed protein in organic solvents in order to measure the extend of proton-deuterium exchange. Several structural studies using solid-state NMR were carried out by Opella and coworkers (Opella, 1997; Marassi & Opella, 1998). These investigations conclusively demonstrate that high-resolution information on membrane peptides and proteins can be obtained using solid-state NMR spectroscopy.

Ste2p as a receptor model

The α-factor receptor (Ste2p) of *S. cerevisiae* represents a good model to study the structure and function of the GPCR family. It is the only GPCR in the haploid "**a**" cells from yeast. Biological assays have been accurately developed and characterized. A number of experiments have been performed on the α-factor that includes synthesis of α-factor analogues (Breslav et al., 1997), and studies on the structure in solution of

the pheromone and its analogues in search for the biologically active conformation (Gounarides et al., 1994; Rao et al., 1995; Yang et al., 1995; Xue et al., 1996; Marepalli et al., 1996). A receptor binding assay using radiolabeled α-factor and analogues has been developed (Jiang et al., 1995; Breslav et al., 1996, Liu et al., 2000). Overexpression and purification of the complete α-factor receptor (David et al., 1997) is readily carried out providing a source of mg quantities of this protein. Biophysical studies on fragments of the yeast GPCR are a major topic of investigations in our laboratory (Reddy et al., 1994; Xie et al., 2000; Ding et al., 2001b). Fluorescence studies of the binding environment of the ligand/receptor complex (Ding et al., 2001a) and assessment of receptor affinity of pheromone analogs (David et al., 1997; Raths et al., 1988) are all in progress.

Objective

In order to address the question of the overall arrangement of the membrane domains and the specific interaction of this receptor and its ligand we have expressed, isolated and investigated the solution conformation of the M5I3M6 and M6 regions of Ste2p. The M5I3M6 region corresponds to transmembrane sequences M5 and M6 joined by the third intracellular loop I3. M6 represents only the sixth transmembrane domain. The significance of these regions is derived from experimental evidence indicating the involvement of the third intracellular loop in the interaction with the G protein (Clark et al., 1994) and mutagenesis data indicating that the substitution of

Pro258 leads to permanent activation of the receptor without the presence of the activating pheromone. Elucidation of the structure of this region of the membrane protein is part of a larger project to solve the structure of the receptor in order to understand the mechanism by which an extracellular signal transduces the information to a G-protein. It is our belief that this goal can only be accomplished by establishing an efficient expression system for preparing isotopically labeled receptor domains.

In vivo assembly of polypeptide fragments of rhodopsin was carried out successfully and represents a good approach to study the mechanism of protein folding, and assembly of a transmembrane protein (Ridge et al., 1995; Ridge et al., 1996; Luneberg et al., 1998; Marti et al., 1999; Martin et al., 1999; Ridge et al., 1999). We propose to express M5I3M6 and M6 in *E. coli* and study their structure in lipid-like environments. Biophysical studies of the longer Ste2p segment in a membrane environment would give the first physical evidence of the structure and interaction of different domains of this GPCR in a lipid environment. The specific aims of this work are as follows:

1. Prepare a construct containing the M5I3M6 and M6 domains as fusion proteins, optimize expression and develop procedures to purify these Ste2p segments.
2. Characterize M5I3M6 fusion protein using chromatographic and biochemical techniques.
3. Study the conformation of M5I3M6 and M6 fusion proteins in membrane mimetic solvents to determine their 2° structures.

4. Carry out CNBr cleavage on the M6 membrane fusion protein to release it from the leader expression peptide.
5. Develop conditions for isotopic labeling of M6 with ^{15}N; isolate and purify ^{15}N-M6 polypeptide.
6. Determine the 3D structure of M6 by NMR spectroscopy.

Chapter II: Materials and Methods

Vectors

The yeast expression vector pNED1 was obtained from Ned David (Berkeley, CA) (David et al., 1997). The parent plasmid pMD194 containing *STE2* gene used to generate all mutants was obtained from Mark Dumont (Martin et al., 1999). The cloning vector pRSET B was purchased from Invitrogen (Carlsbad, CA). The plasmid pMMHa was obtained as a gift from Peter Kim (Staley et al., 1994).

Primers

Generation of mutants was carried out using the single stranded DNA template method (Kunkel et al., 1987). The plasmid that contains the wild type *STE2* gene used as a template was pMD194 (Martin et al., 1999). Primers designed to mutagenize Met^{218} and Met^{250} are listed in Table 2.

Table 2

Sequence of Primers Used in Mutation of Ste2p

Primer name	Primer Sequence	Plasmid Generated	Mutation of amino acid
ON300	5'GATTGACATGACAATATGAGTAAAATATGG3'	pMD600	Met218Leu
ON301	5'CCAGGACAAATGACAAAAAGTTTATTGAGG3'	pMD601	Met250Leu
ON302	5'CCAGGACAAATGACAAAAAGTTTATTGAGG3'	pMD602	Met250Ala
ON300 and ON301	5'GATTGACATGACAATATGAGTAAAATATGG3' 5'CCAGGACAAATGACAAAAAGTTTATTGAGG3'	pMD603	Met218Leu and Met250Leu
ON300 and ON302	5'GATTGACATGACAATATGAGTAAAATATGG3' 5'CCAGGACAAATGACAAAAAGTTTATTGAGG3'	pMD604	Met218Leu and Met250Ala

Primers used to amplify *STE2* regions (M5I3M6 or M6) are listed in Table 3:

Table 3

Primers Used to Amplify Coding Sequences for Single and Double Transmembrane Domains of *STE2*

Primer name	Primer sequence	Region of Ste2p amplified
M5I3M6f	5'TGGAAGCTTTGGSTGSGTGCCACCCAAGATA AATACTTCAAT3'	M5I3M6
M5I3M6r	5'ATATTGGATCCTCAGACATCTGTTCCCTGGT TTGG3'	M5I3M6
M6f	5'TGGAAGCTTTGGATGAAAAAGCAGTTCGATA GTTTCCATATTTTACTC3'	M6
M6r	5'ATATTGGATCCTCACTTTTTCAAACTGTATGC GAGGATGAATATTAT3'	M6
Trpf	5'AGCTTTGGATGAGTTGAG3'	-
Trpr	5'GATCCTCAACTCATCCAA3'	-

Strains

The original yeast strain transformed with the plasmid DNA harboring the wild type and mutant receptors was A232 (MATa *ste2-Δcryl*^{R} *ade2-1 his4-580 lys2*$_{oc}$ *tyr1*$_{oc}$ *SUP4-3*^{ts} *leu2 ura3 bar1-1 FUS1::p[FUS1-lacZ TRP1]*). A232 was obtained as a courtesy from M. Dumont (Martin et al., 1999). *E. coli* expression strains: BL21(D3) and BL21(DE3)pLysS, were purchased as competent cells from Promega (WI). *E. coli* propagation cells: DH5α and XL-1 Blue were purchased from Gibco BRL Life

Technologies (Grand Island, NY) and Stratagene (La Jolla, CA), respectively.

Chemicals

Polymerase chain reaction reagents, dNTP (dATP, dGTP, dCTP, and dTTP), Taq polymerase, restriction enzymes, and ligase were obtained from Promega (Madison, WI). Agarose and low melting point agarose were purchased from GIBCO BRL (Grand Island, NY). The anti-Xpress mouse monoclonal IgG_1 antibody, isopropyl-β-D-thiogalactopyranoside (IPTG), M13 phage that contain T7 RNA polymerase, and the immobilized metal affinity column for protein purification (ProBond resin that contains $NiCl_2$) were all purchased from Invitrogen (Carlsbad, CA). Guanidinium hydrochloride (GuHCl), the detergents sodium deoxycholic acid, igepal CA60, sodium N-laurylsarcosine, and SDS, and most of the buffer chemicals (trizma-HCl, sodium phosphate, sodium chloride, potassium chloride, magnesium chloride) were purchased from Sigma (St. Louis, MO). Media ingredients such as Luria Broth base and yeast extract were purchased from Gibco BRL Life Technologies, peptone and agar were purchased from Difco laboratories (Detroit, MI) and tryptone from Sigma (St. Louis, MO). SDS-PAGE reagents (Coomassie brilliant blue G-250, acrylamide) were purchased from BioRad (Hercules, CA). Glycerol, bromophenol blue, ammonium persulfate, N,N,N',N'-tetramethylethylenediamine (TEMED), and β-mercaptoethanol were purchased from Sigma (St. Louis, MO). Luminol reagents for the development of immunoblots or His tags were purchased from Amersham Life Science

(Buckinghamshire, England), secondary IgG mouse antibody horseradish peroxidase labeled and nitrocellulose membrane 0.45 μm pore size were purchased from Sigma (St. Louis, MO). India His Probe was obtained from Pierce (Rockford, IL). Phospholipids used to prepare the lipid vesicles were 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and 1,2-dimyristoyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)] (DMPG) purchased from Avanti Polar Lipids (Alabaster, AL).

Media, buffers and solvent systems

Liquid Luria-Bertani medium (LB medium) was prepared mixing 10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl in 1 L of H_2O. Plates of LB medium were prepared in the same way as the liquid with the addition of 15 g of agar. 2 x YT medium consisted of 16 g of tryptone, 10 g of yeast extract, and 5 g of NaCl in 1 L of H_2O. SOB medium for glycerol stocks was prepared as follows: 20 g of tryptone, 5 g of yeast extract, 0.5 g of NaCl, and 186.0 mg of KCl were mixed in 1 L of H_2O. All media were brought to pH 7.0 and sterilized by autoclaving. After autoclaving, the following filter-sterilized antibiotics were added to the cool media: chloramphenicol (final concentration 35 μg/mL) and ampicillin (final concentration 100 or 200 μg/mL).

Synthetic drop-out uracil deficient media (SD–URA) was prepared by mixing 4 g of yeast nitrogen base (without amino acids), 12 g of glucose and 0.4 g of synthetic drop-out mix of amino acids in 600 mL distilled water. The clear solution was autoclaved at 121 °C for 20 min. The synthetic drop-out mix of amino acids deficient in

uracil was prepared by mixing 2 g of each the following components: adenine hemisulfate, arginine HCl, histidine HCl, isoleucine, lysine HCl, methionine, serine, threonine and tyrosine; 4 g of leucine, 3 g of each of phenylalanine and tryptophan and 9 g of valine. Yeast extract-peptone-dextrose media (YPD) was prepared by dissolving 10 g of yeast extract, 20 g of peptone, 20 g of dextrose in 1 L of water followed by autoclaving at 121 °C for 20 min. Z buffer was made of 60 mM $Na_2HPO_4.7H_2O$, 40 mM $NaH_2PO_4.H_2O$, 10 mM KCl, and 1 mM $MgSO_4.7H_2O$ at pH 7.0.

Lysis buffer was made of 50 mM Tris-HCl, 1 mM EDTA, 1 mM PMSF, and 300 µg/ml lysozyme at pH 8.7. Inclusion bodies washing buffer was composed of 1% igepal Ca-60 and 1% deoxycholic acid in lysis buffer.

Tris-buffered saline (TBS buffer) contained 20 mM Tris-HCl, 500 mM NaCl, pH 7.5. Tris-buffered saline + Tween 20 (TBST) was made of TBS + 0.05% Tween-20 (w/v). Blocking buffer was made of TBS + 3% bovine serum albumin (BSA) (w/v). Dilution buffer was prepared by mixing TBST + 1% BSA (w/v). Phosphate buffer saline (PBS) contained 10 mM sodium phosphate, 138 mM NaCl, and 2.7 mM KCl, pH 7.4. TE buffer (Tris- HCl and EDTA) was made of 8.3 mM Tris-HCl and 40 µM EDTA. Guanidinium buffer was made of 6 M GuHCl, 20 mM sodium phosphate, 500 mM sodium chloride, pH 7.8. Denaturing binding buffer was made of 8 M urea, 20 mM sodium phosphate, 500 mM sodium chloride, pH 7.8. Denaturing washing buffer pH 6.0, denaturing washing buffer pH 5.3, and denaturing elution buffer pH 4 were prepared by adjusting the pH of the denaturing binding buffer to 6.0, 5.3 and 4.0,

respectively. The alternative elution buffer was made of 8M urea, 20 mM sodium phosphate, 500 mM sodium chloride, pH 8.2, varying the imidazole concentration from 50, 200, 350 to 500 mM. Lyophilization buffer was made of 5% ammonium bicarbonate or 5% acetic acid. Solvents used for HPLC purification were 2-propanol:1-butanol (2:1) or acetonitrile containing 0.1% trifluoroacetic acid (TFA). Buffers used to resuspend lipid vesicles were 0.1 mM potassium phosphate, pH 4, and 10 mM ammonium acetate, pH 4. The organic solvent systems used for NMR were 50% d_3-TFE:H_2O and $CDCl_3$:CD_3OH:H_2O (4:4:1)

Generation of Mutant Receptors

Mutation of the transmembrane methionines (Met^{218} and Met^{250}) of Ste2p was accomplished by site directed mutagenesis (Kunkel et al., 1987). Single-stranded and uracil-containing DNA that contains the *STE2* wild type gene was hybridized to the primer or primers listed in Table 2 and the annealed oligonucleotides were converted into biologically active covalently closed circular (CCC) DNA using T7 DNA polymerase and T4 DNA ligase for extension and ligation, respectively. After generation of double stranded DNA *in vitro*, the DNA was used to transfect XL1-Blue *E. coli* strain in order to inactivate the DNA strand containing uracil and generate new double-stranded DNA from the mutated CCC and finally propagate it. Manual sequencing was performed in order to confirm the correct DNA sequence and the proper location of the mutation within each plasmid. The plasmids carrying the proper

mutations were used to transform the yeast strain aforementioned. Transformed yeast strains are summarized in Table 4.

Table 4

Designation of Plasmids and Strains Carrying Mutant *STE2*

Mutation	Plasmid	Strain generated
Met218Leu	pMD600	A1314
Met250Leu	pMD601	A1315
Met250Ala	pMD602	A1316
Met218Leu & Met250Leu	pMD603	A1317
Met218Leu & Met250Ala	pMD604	A1318
Wild Type	PMD194	A1319

Liquid assay for expression of *FUS1-LacZ* expression

The parent yeast strained used to generate the different mutated strains contained the *FUS1* gene fused to the *LacZ* gene. The assays were performed simultaneously on three independent transformants. The cells were grown overnight in selective medium, uracil-deficient dropout medium (SD-URA) to an OD_{600} of approximately 1.0. The cultures were diluted with YPD medium to an OD_{600} of 0.04. Cells were incubated again until the OD_{600} reached 0.1. The desired amount of α-factor was added to each set of cells and the cells were allowed to grow for 2 h in the presence of the pheromone. Cells were transferred to an ice bath and the OD_{600} was determined for each incubated

culture. Approximately 2.5 x 10^6 cells were transferred to microcentrifuge tubes and harvested by centrifugation. The cell pellet was resuspended in 500 µL of Z buffer containing 0.27 µL/mL of β-mercaptoethanol and 0.004% SDS (w/v final concentration). Thirty microliters of $CHCl_3$ were added to the resuspended cells, and the mixture was vortexed and incubated at 28 °C for 5 min. The colorimetric assay was started by adding 200 µL of 4 mg/mL o-nitrophenyl-β-D-galactopyranoside (ONPG) resuspended in Z buffer containing 0.27 µL/mL of β-mercaptoethanol. The samples were incubated at 28 °C for 20-30 min until there was clear development of the yellow color. The reaction was stopped by addition of 500 µL of 1 M Na_2CO_3, the reaction mixture was centrifuged for 5 min, and the optical density at 420 nm was measured. The activity of the β-galactosidase enzyme was calculated as a ratio of the spectrophotometrically determined amount of substrate cleaved (OD_{420}) and the amount of cells harvested relative to the time of the assay.

Growth arrest

MATa yeast cells are incubated overnight up to stationary phase in selective synthetic dropout liquid media (SD-URA). Approximately 3 x 10^6 cells were harvested and resuspended in 1.5 mL of sterile water. The cell suspension was mixed with 2 mL of SD–URA containing 2% agar, briefly vortexed, and poured over plates containing the same selective media. The agar was allowed to solidify and the indicated amounts of α-factor were spotted onto sterile filter disks and the wet filter paper placed in different

sections of the plate. The diameter of the halo was measured after incubation of plates for 2 days at 30 °C. The diameter of the circular region of growth inhibition around the site of pheromone addition was plotted against the log of the mass of pheromone added.

Competition binding assay

Binding studies were performed by growing the cells harboring the different plasmids overnight at 30 °C. Cells (1 x 10^7 per mL) were harvested by centrifugation at 5,000 x g at 4 °C. The pelleted cells were washed with cold YM-1 medium and resuspended to 4 x 10^7 cells/mL. Tritiated α-factor and various concentrations of non-labeled pheromone (140 μL) were added to a 560 μL cell suspension so that the final concentration of radioactive peptide was 6 nM (20 Ci/mmole). After a 30-min incubation of the cells with the pheromone, triplicate samples of 200 μL were filtered and washed over glass fiber filtermats using the Standard Cell Harvester and placed in scintillation vials for radioactivity counting. The affinity binding constants (K_i) were calculated using the equation previously derived (Linden, 1982).

Cloning and protein expression

The polymerase chain reaction (PCR) and cloning into expression vectors (pRSET B or pMMHa) were carried out following literature procedures (Sambrook et al., 1989). Protein expression using the plasmid pRSETB was done following the manufacturer's instructions (Invitrogen). Protein expression using pEAP plasmids were

performed similarly to the expressions previously published (Staley et al., 1994). Western Blot and dot blot analysis was performed following literature procedures (Walker et al., 1988). His-tag detection was performed following the manufacturer's procedures (Pierce). A summary of the plasmids generated and the name of the bacterial strains transformed with the respective plasmid appears in Table 5.

Table 5

Designation of Plasmids and Strains Coding for Ste2p Domains

DNA Sequence for expression	Parent Plasmid	Name of expression plasmid	Strain name
TrpΔLE-M5I3M6 (wild type)	pRSETB	pJenMa01	EB002
TrpΔLE-M5I3M6 (wild type)	pMMHa	pEAP01	EB007
TrpΔLE-M5I3M6 (Met218L & Met250Ala)	pMMHa	pEAP02	EB008
TrpΔLE-M5I3M6 (Met218L & Met250Leu)	pMMHa	pEAP03	EB009
TrpΔLE-M6 (Met250Ala)	pMMHa	pSW02	EB011
TrpΔLE	pMMHa	pEAP04	EB012

Expression of TrpΔLE-M5I3M6 and TrpΔLE-M6

The amplified recombinant plasmids were isolated and transformed into *E. coli* strain BL21(DE3)pLysS generating the strains listed in Table 5. In order to achieve expression a single colony from LB plate was inoculated into 5 mL of LB media containing ampicillin and chloramphenicol and allowed to grow overnight. Two milliliters of the overnight culture was then diluted into 250 mL of the same media and cells were allowed to grow to late log phase (OD_{600} ~ 0.7-0.8). Cells were induced with 1 mM IPTG, incubated for 4-5 h, and harvested by centrifugation.

Isolation of Inclusion Bodies and Protein Purification

Expression cells were harvested by centrifugation at 8000 RPM for 20 min at 4 °C. The cell pellet is resuspended in lysis buffer at a ratio of 5 mL per gram of wet cells. The following procedures were done at 4 °C. Cell suspensions were sonicated for 5 to 10 min. The lysate solution was centrifuged at 19,000 RPM for 25 min. The resulting pellet was resuspended and sonicated in 3 mL of inclusion bodies washing buffer. The solution was centrifuged at the same conditions as before. The washing step was repeated once or twice depending on the amount of cell debris observable by eye. Finally, the inclusion bodies pellet was washed with 3 mL of PBS. Inclusion bodies were resuspended by sonication in 3 mL of guanidinium buffer and centrifuged as specified above.

Affinity column purification was carried out using Invitrogen's ProBond resin.

which contains bound divalent nickel ions. Resin was washed with 10 volumes of H_2O and equilibrated with denaturing binding buffer, pH 7.8. Resuspended inclusion bodies were mixed with pre-equilibrated resin and incubated for 2 h with shaking at 4 °C. The solution that flows through the column was collected and the protein-resin complex was washed with 5 volumes of denaturing washing buffers (pH 6.0, 5.3). The protein was eluted from the column using 5 mL of denaturing elution buffer (pH 4). In order to use the alternative elution method using imidazole, the resin was re-equilibrated to pH 7.8 and an increasing gradient of imidazole solution was added to the nickel bound protein.

HPLC purification was carried out using a Waters C4 column with a gradient-based solvent system changing the concentration of the organic solvent from 10 % to 100% at a rate of 5 mL/min. HPLC purification of the TrpΔLE-M5I3M6 was carried out using 2-propanol:1-butanol (2:1) as the organic solvent system, acetonitrile was the organic solvent for TrpΔLE-M6 or M6. The HPLC analytical columns used were Vydac C4 5 μm, 300 Å, 5 mm X 250 mm, and Waters Delta Pak 5 μm, 100 Å, 3.9 mm X 250 mm. For large-scale purifications the semipreparative column used was (Vydac C4 5 μm, 300 Å, 10 mm X 250 mm) and preparative columns Waters Delta Pak C4 15 μm, 100 Å, 19 mm X 300 mm, and Waters Delta Pak C18 15 μm, 100 Å, 19 mm X 300 mm.

Protein cleavage

Cleavage of the transmembrane peptides from their N-terminal leader peptide

was carried using CNBr (Walker et al., 1988). The lyophilized pure fusion protein was dissolved in 98% HCOOH or TFA at a concentration of 1-5 mg/mL. After dissolution of the fusion peptide, the concentration of the organic acid was diluted to 70% by addition of the required amount of water. Approximately 500 fold molar excess of CNBr was added to the protein solution. The reaction mixture was stirred in the dark and intermittently sonicated for 5 s during a period of 24 h. At the end of the cleavage reaction the mixture was diluted 5 times with 5% acetic acid and lyophilized. To ensure complete sublimation of CNBr, the resulting solid was resuspended in 5% acetic acid and re-lyophilized.

Protein solubilization

Solvents used for solubilization of the fusion proteins were TFE, HFIP, glacial acetic acid, or aqueous SDS. Lipid vesicles were prepared by sonication (Ong et al., 1981). Briefly, lipids were dissolved in a small amount of chloroform ($CHCl_3$), the solvent was removed under vacuum, and phosphate buffer or ammonium buffer was added to disperse the film to a final concentration of lipid of 1-2 mg/mL. Sonication was carried out in a temperature-controlled bath at 45 °C using a water bath sonicator at a frequency of 40 kHz. Vesicles containing protein were prepared in the same way; the protein solubilized in $CHCl_3$:CH_3OH:H_2O (4:4:1) was added to the mixture of the $CHCl_3$ and phospholipids with a lipid/protein ratio of 50:1. Large unilamellar vesicles (LUVETs) were prepared by the extrusion method (Demmers et al., 2000).

Determination of protein concentration and sample preparation

Protein concentration was determined by quantitative amino acid analysis and the Bradford method (Hammond et al., 1988). The primary sequence and molecular weight of the recombinant proteins was confirmed by amino acid analysis and by mass spectrometry (ESI-MS).

Preparation of samples in TFE/H_2O mixtures was accomplished by dissolving the membrane proteins in the organic solvent and then adding water. When using the $CHCl_3$:CH_3OH:H_2O (4:4:1) solvent system the proteins were directly dissolved into this solvent mixture. Protein-lipid vesicle complexes were prepared as outlined previously in the protein solubilization section of Materials and Methods. The final concentration of the proteins in lipid vesicles was determined by quantitative analysis of their HPLC chromatograms using as a reference the chromatogram of a stock solution of each protein of known concentration.

Instruments

High-performance liquid chromatography (HPLC) was carried out in a Hewlett Packard Series II 1090 instrument for qualitative analysis and a 1050 series for large-scale purification.

Circular dichroism (CD) experiments were performed using an AVIV instrument model 62A DS, Lakewood, New Jersey. The spectrometer was equipped with a water bath for temperature control. Cuvettes with 0.1 cm path length were employed. Each curve is the average of 5 scans recorded between 260 and 180 nm. Protein concentrations obtained by quantitative amino acid analysis ranged from 0.1 to 0.5 mg/mL. To assess protein stability and thermally induced helix to coil transitions CD spectra were collected at 2 °C increments from 5 °C to 82-85 °C and returned to 5 °C. Blanks were collected in the same manner except that no protein was dissolved in the solvent. Blanks were subtracted from the spectra containing the protein. Mean residue ellipticity values were calculated according to the formula:

$$[\theta] = \Delta\varepsilon/(10 \times C \times l \times AA),$$

where $\Delta\varepsilon$, is the measured ellipticity; C is the concentration of the sample in moles/L; l is the path length in centimeters and AA represents the number of amino acids of the sample. Mathematical subtraction of the CD spectra of TrpΔLE from the fusion proteins (TrpΔLE-M5I3M6 and TrpΔLE-M6) spectra was carried out using the calculated **molar ellipticity** values of each spectrum. Finally, the mean residue ellipticity ([θ]) of M5I3M6 was obtained by dividing the molar ellipticity by the appropriate number of residues.

Infrared spectra were measured in a Nicolet Magna 550 Infrared spectrometer, Madison, Wisconsin. Fluorescence experiments were carried out in a photon-counting Spex Fluoromax-3, Edison, New Jersey.

All solution nuclear magnetic resonance (NMR) experiments were performed using a Varian Unity Inova-600 MHz NMR spectrometer. The final concentration of TrpΔLE-M5I3M6 and TrpΔLE in 50% TFE were between 0.7 and 0.8 mM, respectively. The sample concentrations in $CDCl_3$:CD_3OH:H_2O (4:4:1) were 0.5 mM for TrpΔLE-M5I3M6, TrpΔLE and TrpΔLE-M6. Chemically synthesized peptides of the sixth transmembrane domain (Xie et al., 2000), M635P and M635L were dissolved in the $CDCl_3$:CD_3OH:H_2O (4:4:1) system to a final concentration of 0.71 and 0.87 mM, respectively. The concentration of the ^{15}N-M6 sample in the 4:4:1 solvent system was 0.70 mM. NMR spectra were processed using NMRPipe software (Delaglio et al., 1995). The analyses of the NMR spectra were carried out using NMRVIEW software (Johnson et al., 1994).

Molecular modeling was performed using the Crystallography & NMR System (CNS) software (Brunger et al., 1998). The NOE data were normalized according to the r^6 function correlating the median of the NOEs intensity to a distance of 2.7 Å.

Mass spectrometry analyses were carried out on Thermo Finnigan LCQ DECA with an electrospray source, San Jose California. Samples were dissolved in a mixture of acetonitrile:water:acetic acid (50:50:1).

Chapter III: Results

Characterization of mutated Ste2p

Generation of Ste2p (Table 4) containing mutations in methionine 218 and 250 was accomplished following the site directed mutagenesis method described elsewhere (Kunkel et al., 1987). The nucleotide sequence of the mutated receptors was confirmed by DNA sequencing. Yeast cells were transformed with the plasmids harboring the corresponding mutations and the generated strains are listed in Table 4. In order to determine the rate of growth in SD-URA media for each strain, the growth of the cells was monitored using the same conditions and the optical density at 600 nm was plotted against time (Figure 3).

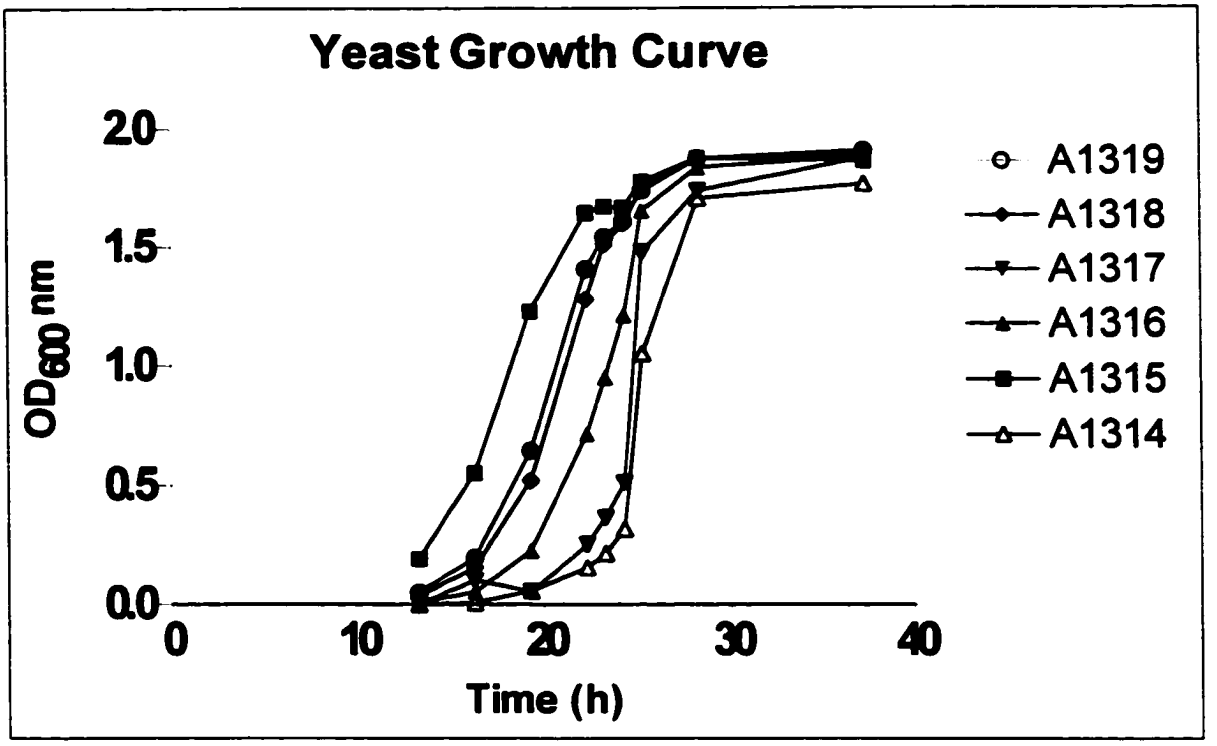

Figure 3. Growth curves for wild type and mutant Ste2p. The plasmids (see Table 4) carrying mutated Ste2p were transformed into *S. cerevisiae* and the growth of these cells was measured in SD-URA medium.

It can be observed that all strains reach $OD_{600} \approx 1$ within a 7-h range. The fact that some strains showed early exponential growth, like A1315 in the graph, arises from variability of the size of the single colony taken for the particular experiment. It was observed in the other biological assays that the rate of growth of the cells containing mutated Ste2p was, in general, similar to the wild type. The transformed yeast strains were tested in the growth arrest assay (also called the halo assay). The halo assay measures the response and adaptation of the cells to the exposure of the pheromone over a period of 2 days. It provides an assessment of the efficacy of signal transduction by

the α-factor receptor. Figure 4 shows the average of three independent measurements for each strain. The linear regression analysis of the semi-logarithmic plot of the data points for strains harboring the different mutations is shown in Table 6.

Table 6

Regression Analysis of the Growth Arrest Data

Strain	B*	$R^{2\,\#}$
A1314	13.18	0.9838
A1315	16.61	0.9243
A1316	12.18	0.9610
A1317	9.30	0.9648
A1318	12.40	0.9685
A1319	13.29	0.9771

* The data in Figure 8 were fit with the following equation: $Y = A + B \times \log(X)$

R^2 is defined as the coefficient of determination, calculated from the Pearson correlation coefficient.

All the strains were able to maintain the pheromone response for two days and show a halo corresponding to the growth inhibition due to the presence of the α-factor. The plot of the halo diameters against the logarithm of the amount of ligand added fit fairly well with a linear equation as described by the "R squared" (R^2) value in Table 6. All mutant receptors except for single mutation Met250Leu (A1315) and double mutation Met218Leu, Met250Leu (A1317) show similar slopes to that of the line for the wild type receptor. It is generally accepted that strains with different EC_{50} values (concentration of α-factor at which 50% of the maximum activation occurs) for

pheromone yield lines with similar slopes but different y-intercepts (Sommers et al., 1999).

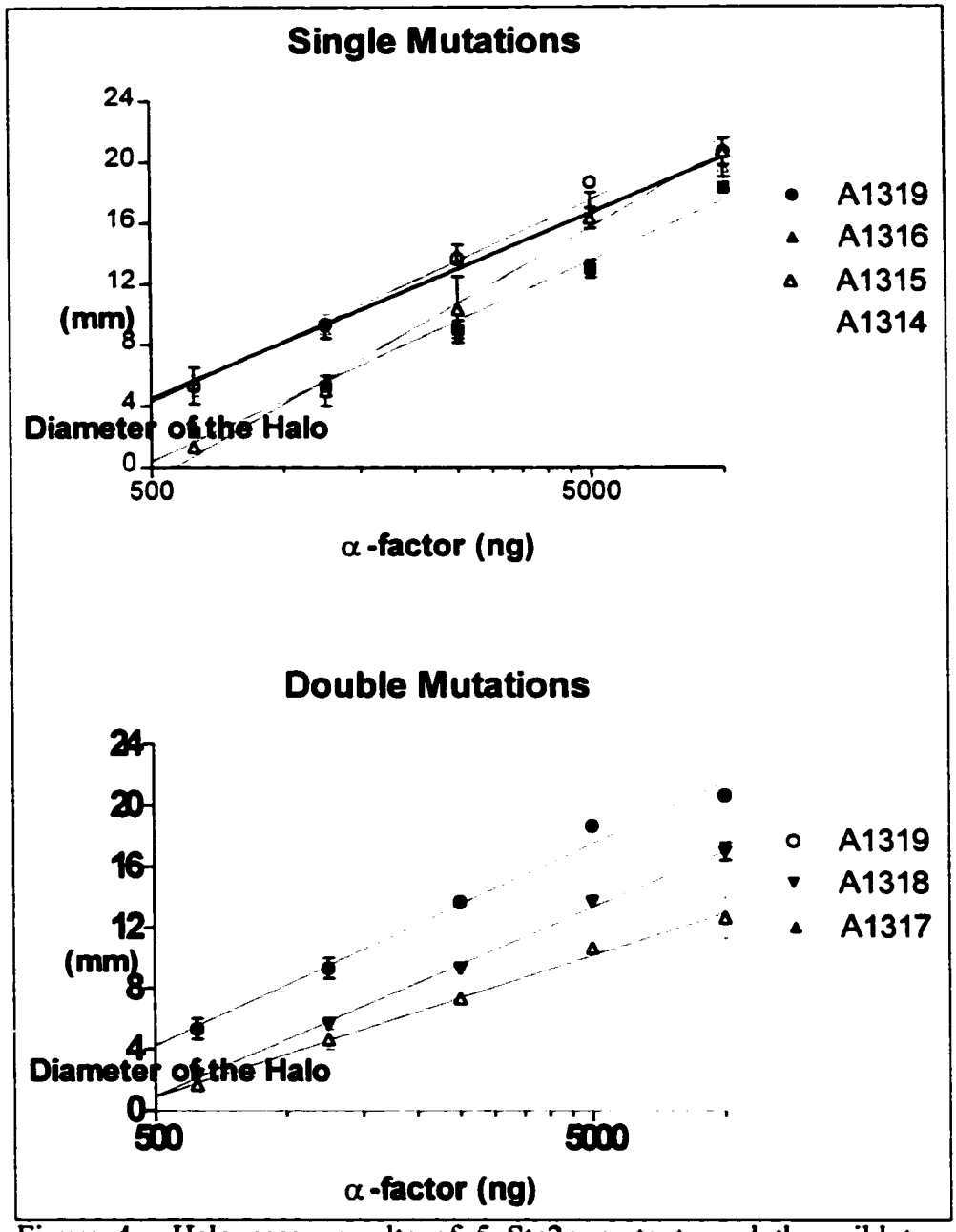

Figure 4. Halo assay results of 5 Ste2p mutants and the wild type receptor. The top panel shows the results of the strains containing the single mutation receptor compared to the strain harboring the wild type receptor. The bottom panel shows the strains containing the double mutation receptors compared to the wild type strain.

Figure 5 shows the percentage of response of each of the mutant receptors relative to the wild type at different amounts of pheromone. It can be noticed that the Met250Ala mutation in Ste2p resulted in the highest response of all the mutants and that the response of the other mutated receptors increased upon increase of the α-factor.

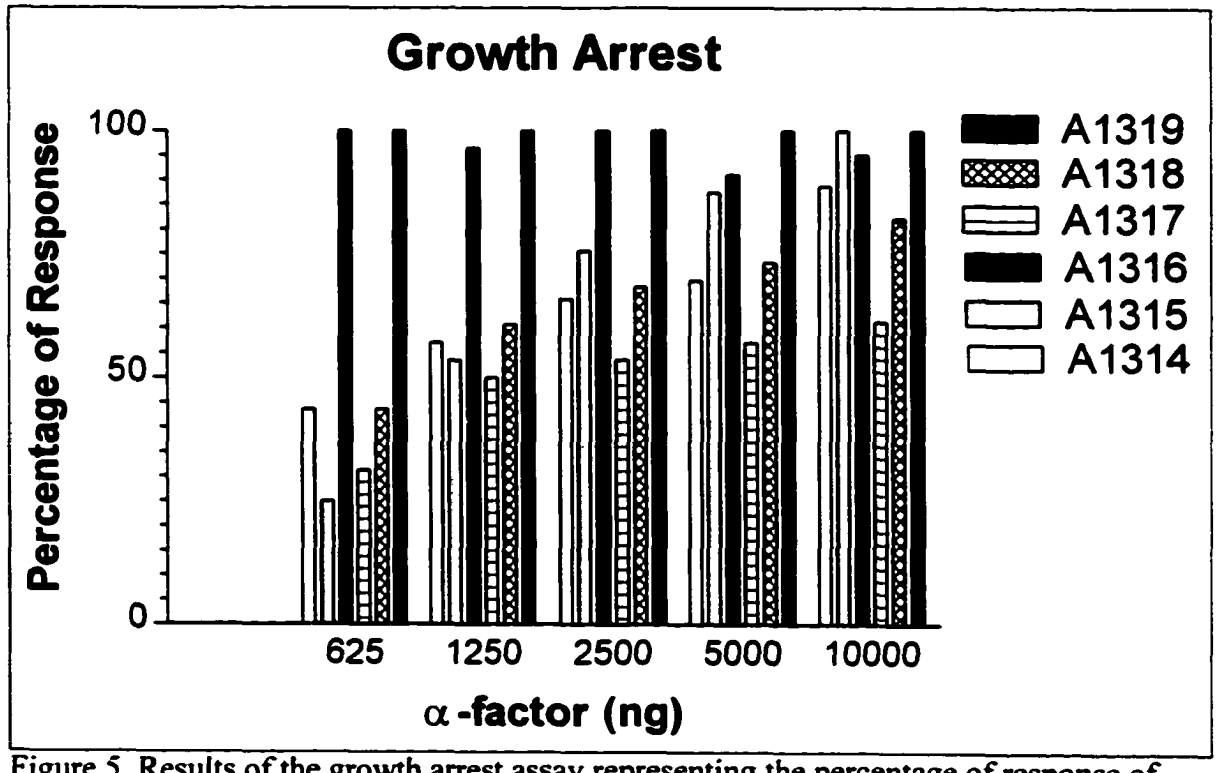

Figure 5. Results of the growth arrest assay representing the percentage of response of different strains upon incubation with different amounts of α-factor. The maximum response is based on the values of the wild type strain.

Strain A1314 was found to contain certain additional mutations and was not considered in the final analysis of all receptor characteristics. A "clean" strain with the Met218Leu mutation was received from Dr. Mark Dumont of the University of Rochester. Data on this strain are reported in Figures 8-10 and Table 6.

The *FUS1-LacZ* assay, also known as the β-galactosidase assay is a direct measurement of the expression levels of the *FUS*1 gene upon exposure to the pheromone. This assay provides reliable quantitation of the gene induction triggered by the ligand (Sommers, et al. 1999). In this assay, the response of Ste2p harboring single and double mutations correlates fairly well to the results obtained in the halo assay. A comparison of the gene induction by mutant receptors in response to different amounts of α-factor is shown in Figures 6A and 6B.

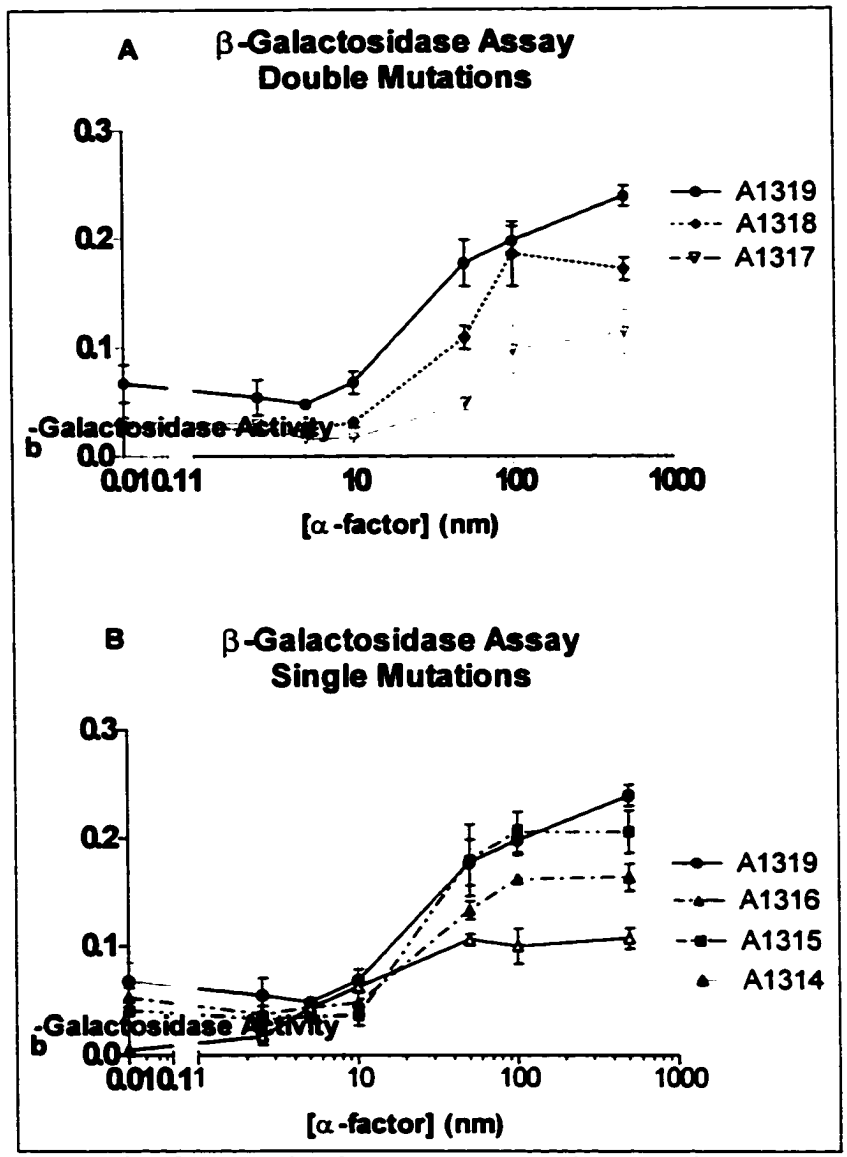

Figure 6. Gene activation assay on strains carrying mutant receptors. Graph A, depicts the β-galactosidase activity measured from the receptors containing the double mutations in the presence of 0, 2.5, 5, 10, 50, 100 and 500 nM α-factor. Graph B shows the activity measured from the receptors harboring the single mutations under the same conditions described above.

In all cases the wild-type receptor resulted in more efficient signaling than any of the mutants. However, it can be observed that the EC_{50} is located between 10 and 50

nM for all of the mutated receptors. This range is indicative of receptors that maintain significant ability to transduce signal. Despite the fact that the mutant receptors show less activation upon induction with the ligand, the percentage of response is still significant as illustrated in Figure 7.

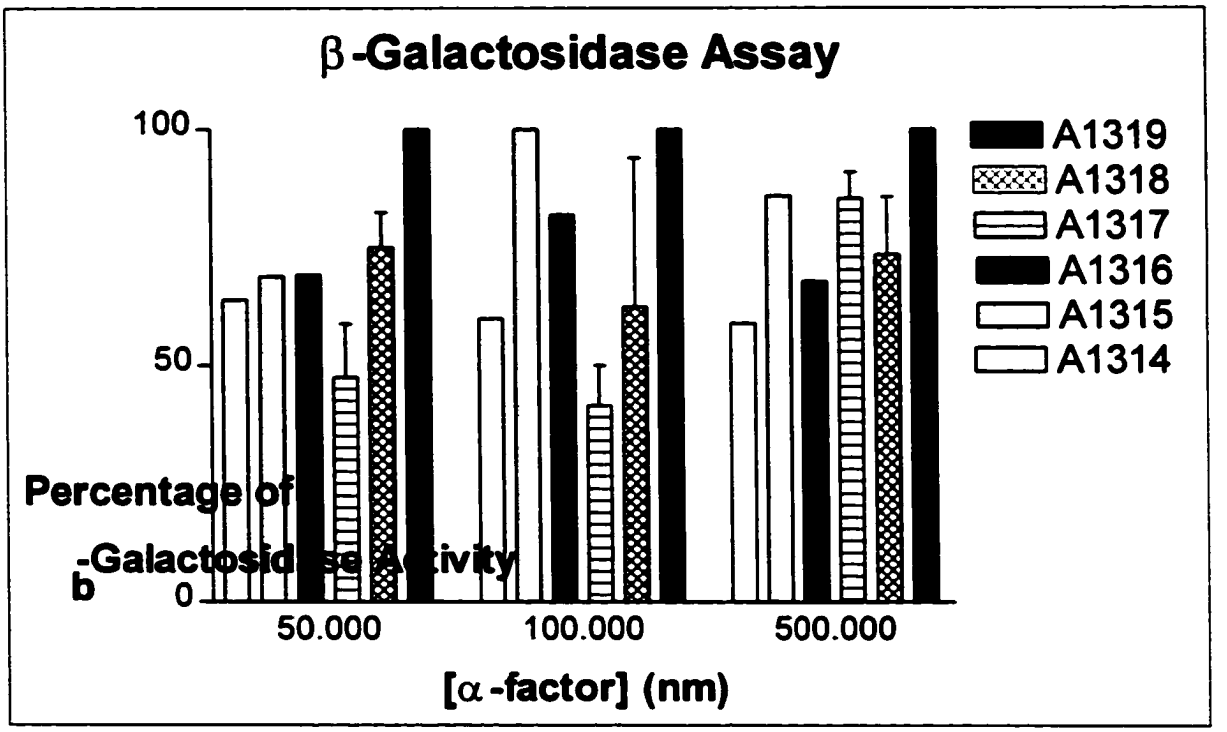

Figure 7. Percentage of gene response induction of strains carrying mutated α-factor receptors. Response upon induction with 50, 100 and 500 nM α-factor. The maximum response is based on the values of the wild type strain.

The conclusion is that the mutation of methionine 218 and/or methionine 250 does not significantly interfere with the biological function of Ste2p.

Binding assay experiments were carried out to determine the number of binding sites per cell in each strain as well as the dissociation constant of the radiolabeled peptide (K_D) (Linden, 1982). In order to find out the range of concentration of the

ligand to work with in the successive experiments, saturation binding assays were performed (Figure 8).

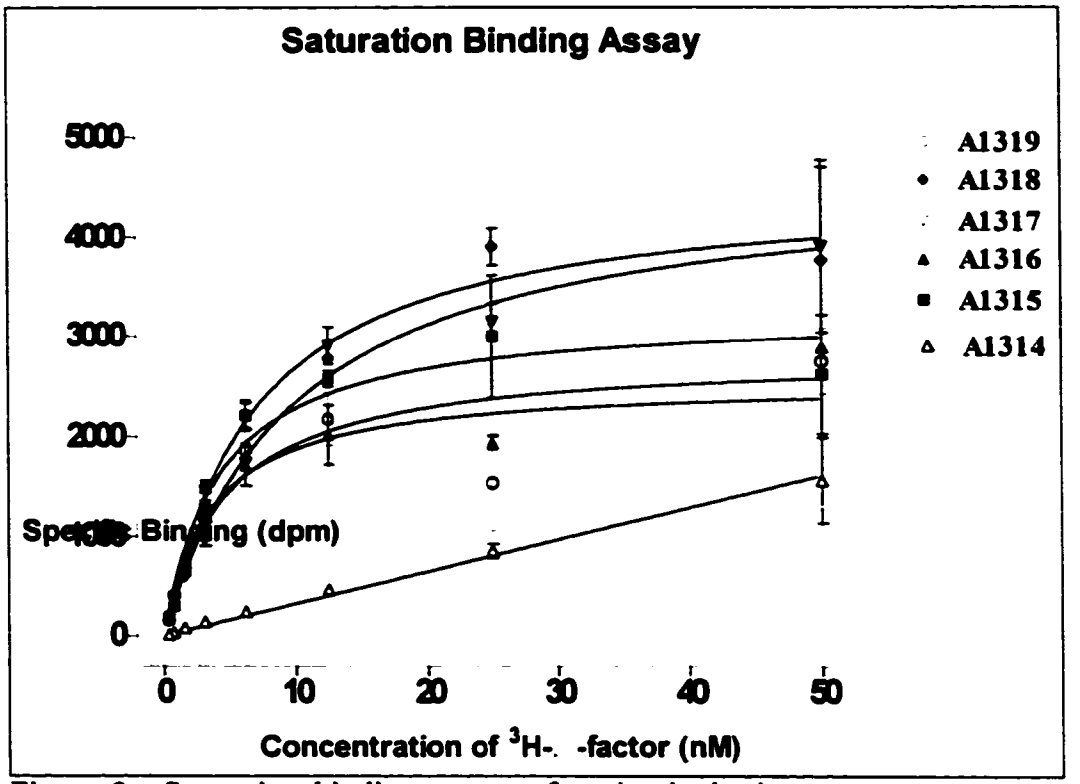

Figure 8. Saturation binding assays of strains harboring mutant receptors. Cells were incubated with increasing amounts of ^{3}H-α-factor and the specific binding (dpm) associated with the cells containing the different mutant receptors is plotted against α-factor concentration.

Subsequently, competition binding assays were carried out (Figure 9) and the results were analyzed to determine the dissociation constant and the number of binding sites per cells listed in Table 7.

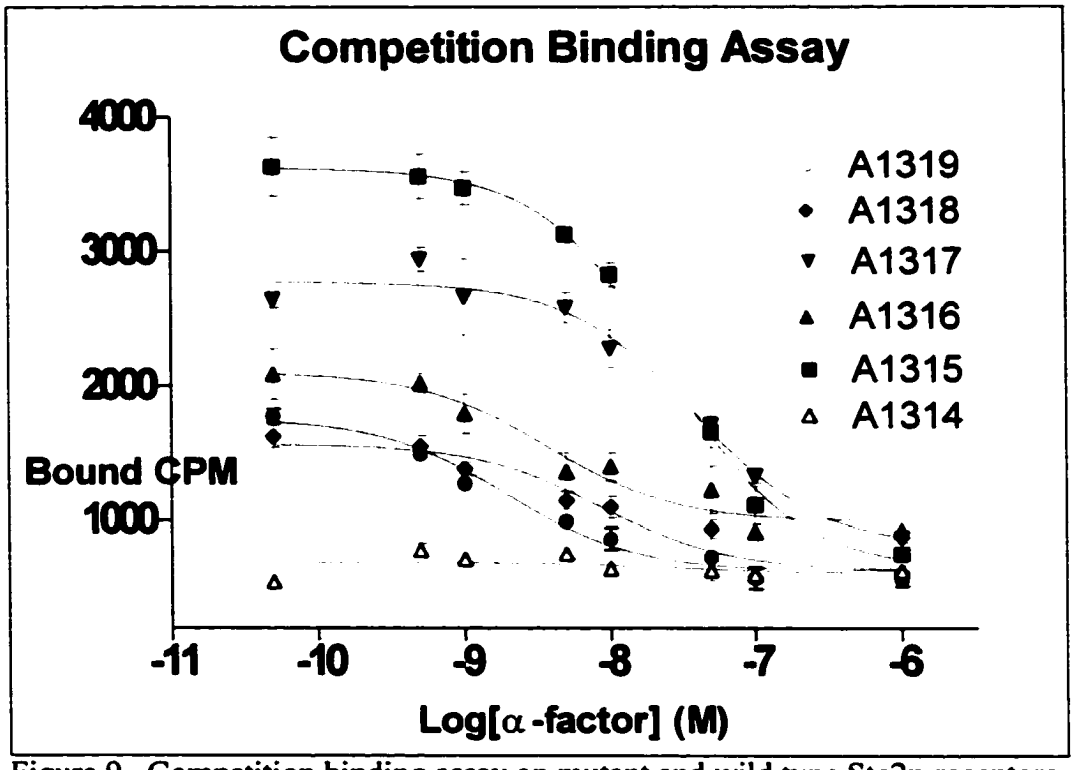

Figure 9. Competition binding assay on mutant and wild type Ste2p receptors. Binding of cold α-factor was performed in competition with ^{3}H-α-factor. The remaining radioactivity of labeled α-factor was plotted against the concentration of cold α-factor.

Table 7

Binding Constants Determined from the Competition Binding Assay

Receptors	K_D (nM)	*Bmax* (dpm)	Binding Sites/Cell
A1319 (WT)	3.59±1.60	2547±319	6785±850
A1314 (M218L)	NA	NA	NA
A1315 (M250L)	4.24±1.12	3243±251	8639±669
A1316 (M250A)	4.68±1.35	2818±242	7507±645
A1317 (M218L,M250L)	9.97±1.52	4659±258	12411±687
A1318 (M218L,M250A)	6.98±1.20	4549±258	11991±680

NA: Specific binding was not determined under the experimental conditions

The binding assays indicate that results obtained in the growth arrest and β-galactosidase experiments were not due to the lack of expression of the receptors on the surface of the cells but mostly because of a subtle alteration in the binding affinity of the mutant receptors to the α-factor pheromone. Strain A1316 carrying the receptor mutant Met218Ala and strain A1315 carrying mutant Ste2p with Met250Leu have similar α-factor dissociation constants to that of the wild receptor. Of the two double mutations, A1318 (Met218Leu and Met250Ala) exhibits less than a two-fold decrease in binding affinity and has about twice the number of receptors per cell. Interestingly binding could not be detected in strain A1314, and A1317 had the lowest affinity for the pheromone.

The binding results correlate fairly well with the halo assays and the β-

galactosidase assays. The receptors with only one mutation functioned more closely to the wild type than the receptors with the double mutations. In the β-galactosidase assay, points corresponding to 500 and 1000 nM did not correlate with the growth arrest or binding results. However, the effect of α-factor probably saturates near 100 nM. Points taken above this concentration probably involve non-specific effects of the peptide and may not have physiological relevance. Indeed at 100 nM, and below, the response of the receptors with single mutations correlates with the growth arrest results.

These biological and biochemical assays provided information supporting the conclusion that the substitutions of Met218Leu and Met250Ala do not dramatically change the structural conformation of the receptor. These mutations do not change the ability of the cells to express and localize the mutated receptors in the cell membrane, and the mutated receptor functions similarly to the wild-type receptor. Thus the Met218Leu, Met250Ala Ste2p mutant was selected for further structural analyses.

Cloning, expression and purification of the M5I3M6 and M6 regions of Ste2p in *Escherichia coli*

pNED1 is a bacteria-yeast shuttle vector that harbors the *STE2* wild type gene and the β-lactamase gene, which hydrolyzes the peptide bond of the lactam ring of ampicillin (Figure 10) (David et al., 1997).

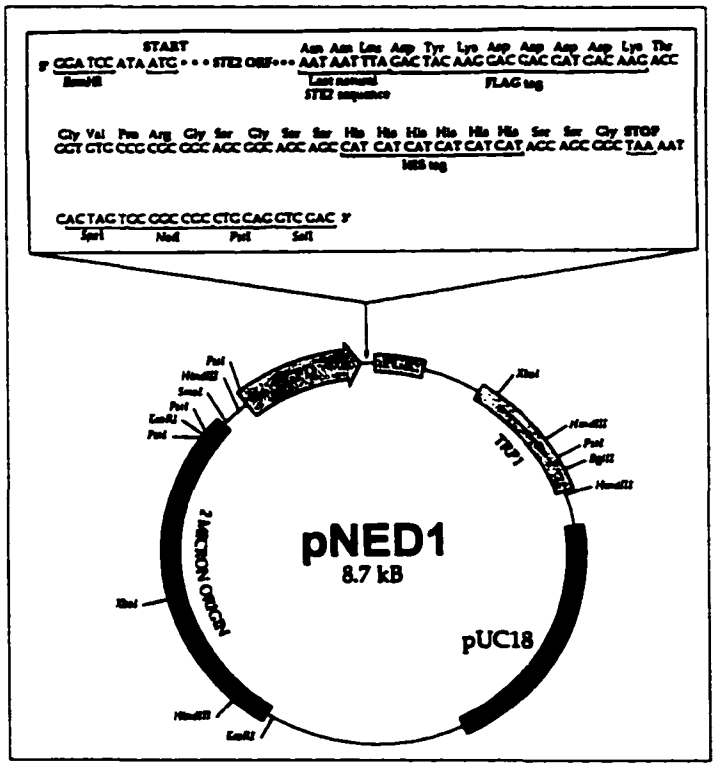

Figure 10: pNED1. Vector derived from the progenitor plasmid pG3 (Schena et al., 1991). It contains the yeast genes, *STE2*, *GPD* promoter (glyceraldehyde-3-phosphate dehydrogenase), *PGK* transcription terminator and polyadenylation signal (phosphoglycerate kinase), *TRP1* gene, 2μm origin of replication and the ampicillin resistance gene from pUC18 (David et al., 1997).

The wild type M5I3M6 region of the *STE2* gene was initially amplified from

pNED1 by means of the polymerase chain reaction (PCR). The amplified region was inserted into a high expression vector called pRSETB between the *Hin*d III and *Bam*H I restriction sites. This commercial plasmid contained an N-terminal fusion peptide which included an ATG translation initiation codon, a histidine tag that functions as a metal binding domain, a stabilizing sequence from gene 10 of phage T7, and an antigenic sequence (-Asp-Leu-Asp-Asp-Asp-Asp-Lys-) for protein detection and the enterokinase cleavage recognition sequence (Asp-Asp-Asp-Asp-Lys-) designed to release the target the recombinant peptide (Figure 11).

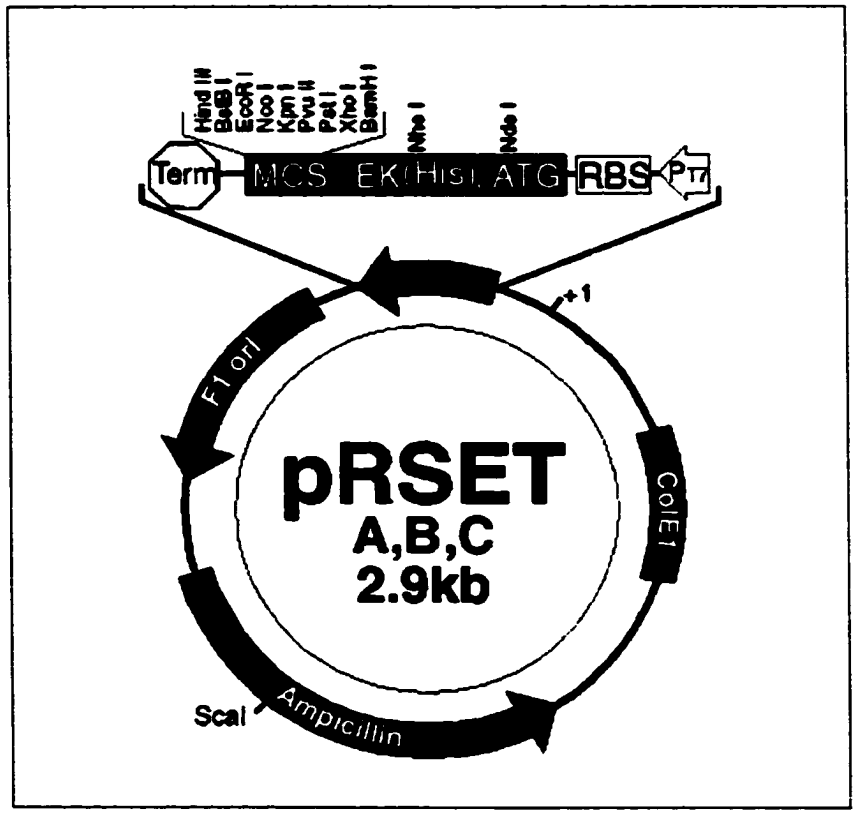

Figure 11. High expression vector pRSETB. It contains the ampicillin resistance gene, the F1 ori, the colE1 and a multiple cloning site (MCS). The M5I3M6 DNA sequence was inserted between the *Hin*d III and *Bam*H I restriction sites.

The cloned DNA (pJenMa01) was propagated in a RecA$^-$ strain of *E. coli* (DH5α cells) that lacks the gene for T7 RNA polymerase. The incorporation of the DNA sequence of M5I3M6 was confirmed initially by restriction enzyme digestion at the unique site of *Cla* I located in the M6 domain and the corresponding analysis of the size of the digested plasmid (Figure 12).

Figure 12: Verification through agarose electrophoresis of the presence of pJenMa01 in 3 independent transformed cells. The plasmid was isolated from each independent cell and digested with *Cla* I for 2 h. Lane 1 is a 1 Kb DNA ladder (10, 8, 6, 5, 4, 3, 2.5, 2, 1.5, 1 Kilobases). Lanes 2, 3 and 4 are digested pJenMa01 with Cla I. The digested plasmids were in the range of 3-4 Kb.

Ultimately DNA sequencing confirmed the incorporation of the correct DNA sequence and the correct reading frame of the gene in line with the leader sequence and the T7 RNA polymerase initiation codon.

Transformation of *E. coli* cells BL21(DE3)pLysS with pJenMa01 yielded strain EB002 (Table 5). Expression of the fusion protein was carried out following the plasmid manufacturer's protocol. A single colony was grown overnight in 5 mL culture

of SOB containing ampicillin and chloramphenicol. The overnight culture was used to inoculate a larger volume of the medium. Induction was carried out by addition of IPTG to 1 mM final concentration. The time course of the expression was followed by examination of 1 mL aliquots of the induced culture at different incubation times. The cell aliquots were lysed in lysis buffer and the supernatant and pellet separated by ultracentrifugation at 14,000 RPM. The supernatant and the pellet were mixed with SDS-PAGE sample buffer, vortexed, and incubated at 60 °C for 3 min, loaded on gels and subjected to electrophoresis. Analysis of the Coomassie stained gel did not show any significant change in protein expression in the supernatant fractions (Figure 13).

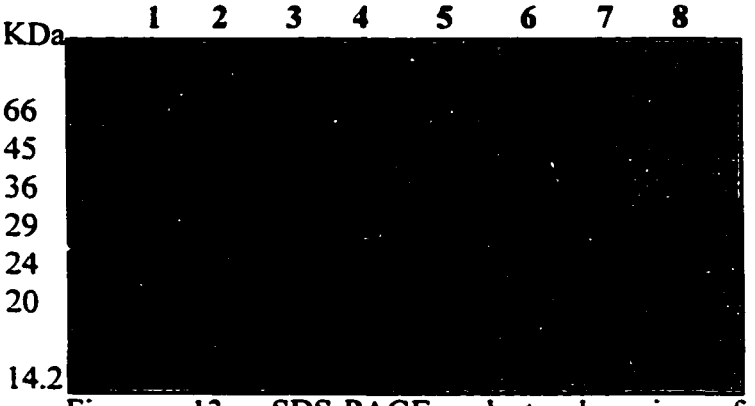

Figure 13: SDS-PAGE electrophoresis of supernatant fractions from the induction time course experiments of cells transformed with pJenMa01. The gel does not show significant change in protein expression.
Lane 1: Molecular marker: 66, 45, 36, 29, 24, 20, 14.2, 6.5 KDa
Lane 2: No induction with IPTG
Lane 3: 1 h after induction with IPTG
Lane 4: 2 h after induction with IPTG
Lane 5: 3 h after induction with IPTG
Lane 6: 4 h after induction with IPTG
Lane 7: 5 h after induction with IPTG
Lane 8: 6 h after induction with IPTG

The pellet fraction did show a band at 18 KDa and another band at 22 KDa. Considering the fact that the expected molecular weight of the fusion protein was 12.8 KDa, the band at 22 KDa was presumed to be a dimerization product of the expressed protein (Figure 14).

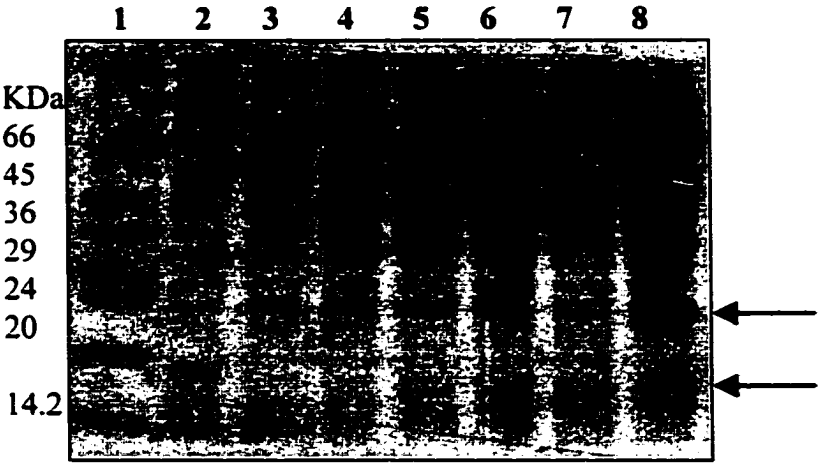

Figure 14: SDS-PAGE electrophoresis of pellet fractions from the induction time course experiments of cells transformed with pJenMa01. The arrows indicate the bands at ~18 and ~22 KDa suspected to be the expression protein.
Lane 1: Molecular marker: 66, 45, 36, 29, 24, 20, 14.2, 6.5 KDa
Lane 2: No induction with IPTG
Lane 3: 1 h after induction with IPTG
Lane 4: 2 h after induction with IPTG
Lane 5: 3 h after induction with IPTG
Lane 6: 4 h after induction with IPTG
Lane 7: 5 h after induction with IPTG
Lane 8: 6 h after induction with IPTG

Western blot analysis of this sample showed a cross reactivity of the protein at 18 KDa but this same cross reactivity was found with non-transformed cells (Figure 15).

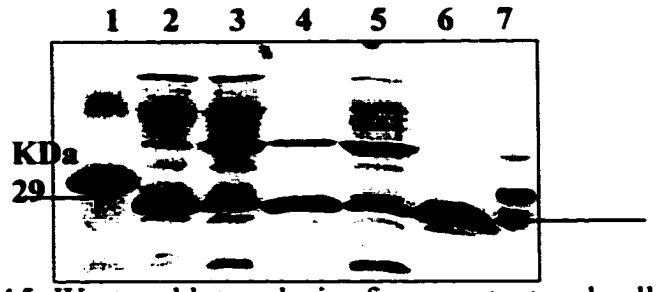

Figure 15: Western blot analysis of supernatant and pellet fractions of cells transformed with pJenMa01. The arrows indicate the bands that were also observed in the SDS-PAGE that would correspond to the protein expression. It can be noticed that the same band is present in both transformed and non-transformed cells.
Lane 1: Molecular marker band of 29 KDa
Lane 2: Pellet fraction of transformed cells after 6 h of induction
Lane 3: Pellet fraction of transformed cells after 5 h of induction
Lane 4: Supernatant fraction of transformed cells after 5 h of induction
Lane 5: Pellet fraction of transformed cells before induction
Lane 6: Pellet fraction of non-transformed cells after 5 h induction
Lane 7: Molecular marker bands of 29 and 20 KDa

Tests were performed to investigate the possibility that the cross reactivity was caused by a non-specific reaction of the secondary antibody but no reactivity was observed. Therefore it was reasonable to conclude that the protein visualized below 20 KDa in the Coomassie stained gels and the western blot was an endogenous protein and not the expected expression protein. The concentration of ampicillin in the media was also increased from 100 μg/mL to 200 μg /mL in order to maintain selective pressure so that the plasmid was not lost.

BL21(DE3)pLysS is a B strain that is deficient in the *lon* protease and the *ompT*

outer membrane protease that can degrade proteins during purification. It contains the bacteriophage DE3 derivative that carries the DNA fragment of the *lacI* gene, the *lac*UV5 promoter, the beginning of the *lacZ* gene and the gene for the T7 RNA polymerase. The DE3 lysogen integrated into the chromosomal DNA is only induced to express the T7 RNA polymerase gene in the presence of IPTG, which acts on the *lac*UV5 promoter (Studier et al., 1990). Addition of at least 0.4 mM IPTG to the growing culture induces the expression of T7 RNA polymerase, which in turn transcribes the target DNA in the plasmid. Nonetheless, there are always some basal levels of expression of T7 RNA polymerase that will promote the expression of the expected protein in uninduced cells (Dubendorff et al., 1991). The presence of the additional plasmid pLysS decreases the levels of free T7 RNA polymerase, because it expresses T7 lysozyme, a natural inhibitor of T7 RNA polymerase, thereby inhibiting undesired transcription of the target gene (Moffat et al., 1987; Studier et al., 1990).

Plasmid instability experiments carried out in the EB002 strain showed that the cells would lose their plasmid during the logarithmic phase of growth and even increasing the concentration of ampicillin in the medium could not maintain the selection pressure. This may result from the catalytic activity of β-lactamase secreted into the medium, which rapidly depletes the amounts of ampicillin. Thus even using a tight expression system like BL21(DE3)pLysS could not assure the expression of the transmembrane regions of Ste2p that apparently are significantly toxic to the bacterial cells.

Based on these results it was decided to try the expression of Ste2p transmembrane domains using a leader sequence that would trigger the formation of protein aggregates within the cell in order to minimize the lethal effect of these hydrophobic regions to the bacterial cells. It was found that a protein called TrpΔLE had been successfully used for expression of a soluble protein like bovine pancreatic trypsin inhibitor (Staley et al., 1994) and an integral membrane protein "Vpu" from HIV-1 (Marassi et al. 1999).

The plasmid pMMHa, received as a gift from Dr. Peter Kim at the Massachusetts Institute of Technology (Staley et al., 1994) is a pET-derived plasmid that also contains the gene for β-lactamase. This plasmid contains the T7 RNA polymerase binding site and the N-terminus leader peptide TrpΔLE derived from the tryptophan operon (Mozzari et al., 1978; Staley et al., 1994). The expression vector also contains restriction enzyme cloning sites; *Bam*H I and *Hind* III. PCR reaction was carried out using the plasmid pMD604 (Table 4) as a template to copy the M5I3M6 or the M6 region of the Ste2p receptor containing mutations Met218Leu and Met250Ala. The amplified regions were inserted in pMMHa in between the *Bam*H I and Hind III restriction sites. The resulting clone contained the sequence of TrpΔLE followed by the M5I3M6 or the M6 DNA sequence of Ste2p (Figures 16 and 17).

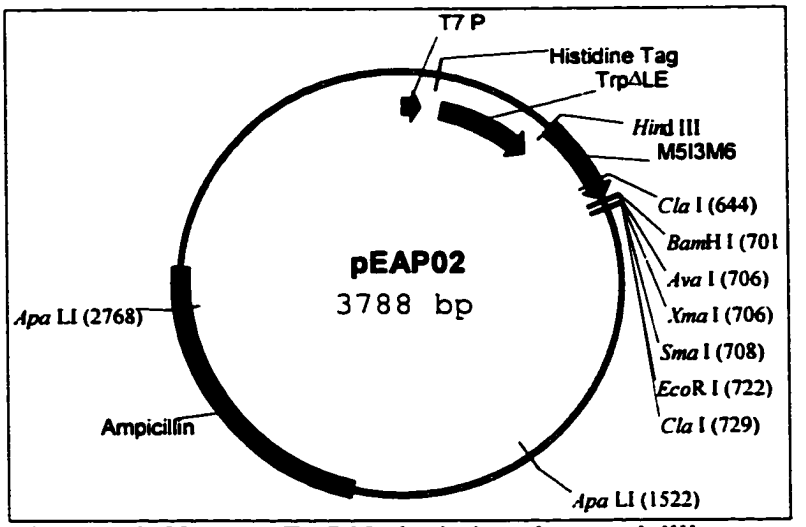

Figure 16. Vector pEAP02 depicting the ampicillin gene, T7 RNA polymerase binding site, His-tag-TrpΔLE leader sequence and M5I3M6 sequence inserted between *Hin*d III and *Bam*H I restriction enzyme sites.

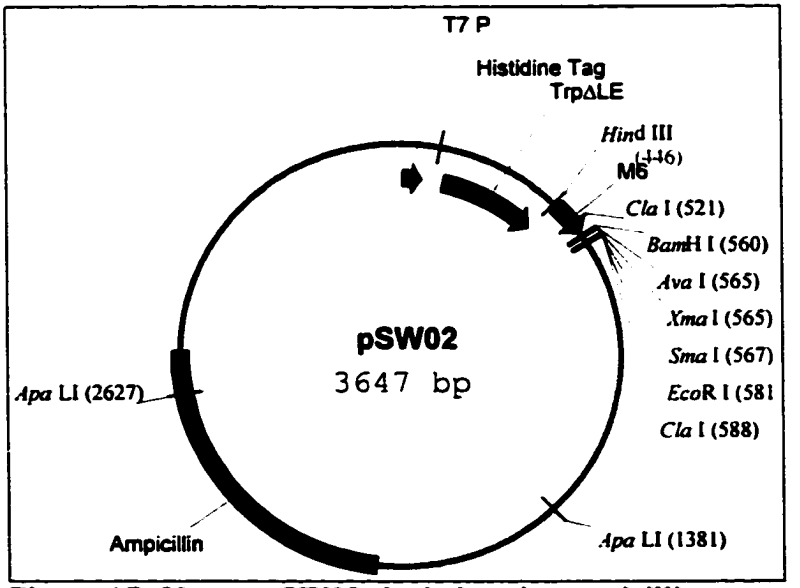

Figure 17. Vector pSW02 depicting the ampicillin gene, T7 RNA polymerase binding site, His-tag-TrpΔLE leader sequence and M6 sequence between *Hin*d III and *Bam*H I restriction enzyme sites.

E. coli strain BL21(DE3)pLysS was transformed with the generated plasmid. In order to construct a plasmid as a positive control for expression carrying just the TrpΔLE, the DNA sequence of bovine pancreatic trypsin inhibitor (BPTI) was removed from pMMHa by enzymatic digestion with *Bam*H I and *Hin*d III. The stop codon was introduced using primers listed in Table 3 and then annealed and ligated to TrpΔLE preserving the *Bam*H I and *Hin*d III restriction sequences.

The M5I3M6 region fused to the TrpΔLE leader sequence was designed so that the hydrophobic region could be released from the leader peptide by chemical cleavage using CNBr. CNBr cleaves at methionine residues (Walker et al., 1988). In order to eliminate internal cleavage of the transmembrane domains during CNBr reaction the template used during the PCR reaction was the gene of the receptor containing the mutations Met218Leu and Met250Ala. The special characteristics of the pMMHa plasmid are the presence of the TrpΔLE leader sequence that contains a ribosome-binding site for promotion of the initiation of translation (Miozzari et al., 1978) and that the expression of this fusion polypeptide triggers the formation of inclusion bodies (protein aggregates) within the cells. This minimizes both the exposure of the expressed protein to proteases and the toxicity of deleterious proteins (Staley et al., 1994).

The fusion of the DNA coding for M5I3M6 to the TrpΔLE leader sequence was confirmed by DNA sequencing. The generated plasmid pEAP02 (Figure 16) was used to transform the BL21(DE3)pLysS strain into the EB008 strain (Table 5). The time course of expression was again investigated in the EB008 strain using a 1 mM final

concentration of IPTG. The SDS-PAGE results did not clearly show the classical band of an over expressed protein (data not shown) as was found when the TrpΔLE fused to bovine pancreatic trypsin inhibitor (BPTI) was used as a positive control of expression (Figure 18).

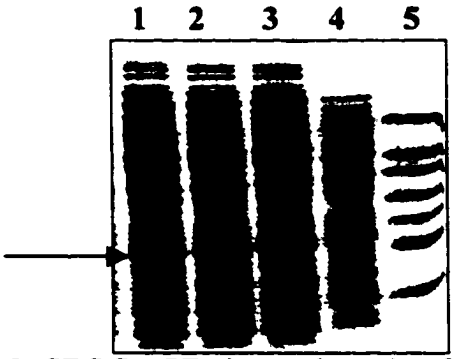

Figure 18: SDS-PAGE electrophoresis of the expression of BPTI by cells transformed with pMMHa. The arrow indicates the band corresponding to BPTI. The strong band of the target protein is already observed at 2 h after induction.
Lane 1: 5 h after induction with IPTG
Lane 2: 3 h after induction with IPTG
Lane 3: 2 h after induction with IPTG
Lane 4: No induction with IPTG
Lane 5: Molecular marker: 66, 45, 36, 29, 24, 20, 14.2 KDa

After the isolation of inclusion bodies, solubilization of the cell pellet with 8 M urea did not prove to be efficient (Figure 19), and it was noticed that a significant amount of pellet was still present after the attempt to solubilize it in 8 M urea.

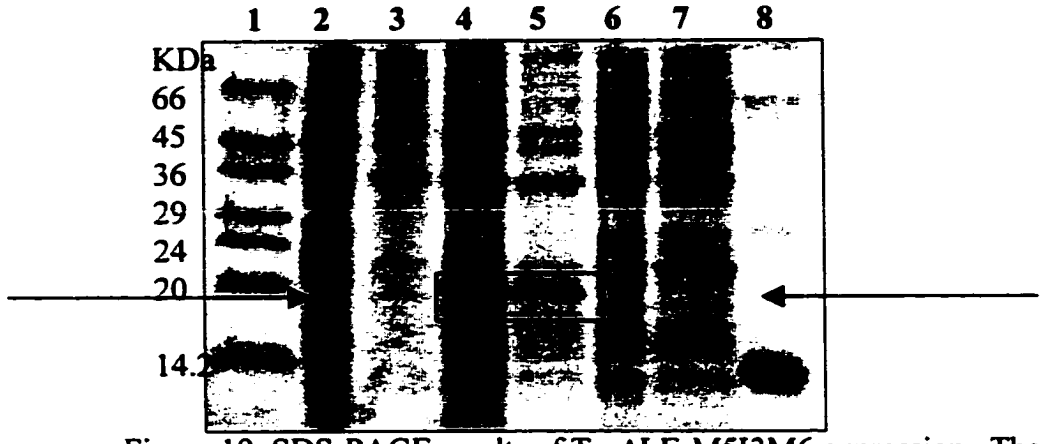

Figure 19: SDS-PAGE results of TrpΔLE-M5I3M6 expression. The arrows indicate the bands corresponding to TrpΔLE-M5I3M6. The rectangle encloses the faint band corresponding to TrpΔLE-M5I3M6. The different lanes correspond to aliquots taken during the isolation of inclusion bodies.
Lane 1: Molecular marker: 66, 45, 36, 29, 24, 20, 14.2 KDa
Lane 2: First supernatant extract from the isolation of inclusion bodies
Lane 3: Second supernatant extract from the isolation of inclusion bodies
Lane 4: First solubilization attempt in 8 M urea
Lane 5: Second solubilization attempt in 8 M urea
Lane 6: Aliquot of non induced cells with IPTG. The cells were directly lysed with loading buffer
Lane 7: Aliquot of cells induced four hours with IPTG. The cells were directly lysed with loading buffer
Lane 8: Molecular marker band of 14.2 KDa

Two treatments of the isolated inclusion bodies using PBS containing 0.5% SDS solubilized the entire observable pellet. A significant amount of TrpΔLE-M5I3M6-was detected in both of these washes (Figure 20, lanes 3 and 4).

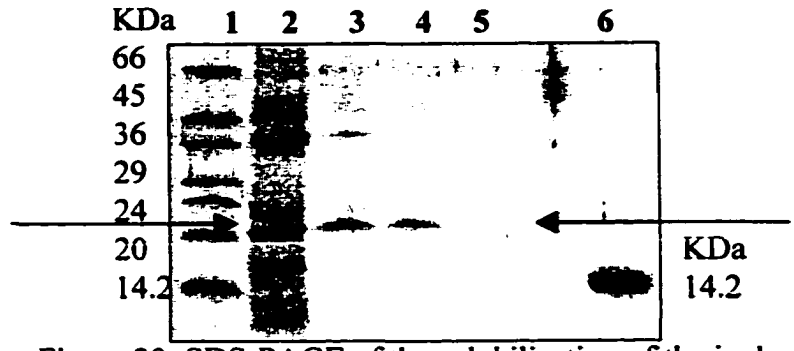

Figure 20: SDS-PAGE of the solubilization of the inclusion bodies in 8M urea and PBS buffer containing 0.5% SDS. The arrows indicate the bands corresponding to TrpΔLE-M5I3M6. Significant amount of pellet was still present after attempts of solubilization with 8 M urea.
Lane 1: Molecular marker: 66, 45, 36, 29, 24, 20, 14.2 KDa
Lane 2: Solubilization of inclusion bodies pellet with 8 M urea
Lane 3: First solubilization trial with PBS 0.5% SDS
Lane 4: Second solubilization trial with PBS 0.5% SDS
Lane 5: Washing of the remaining pellet with H_2O
Lane 6: Molecular marker band of 14.2 KDa

Although SDS proved to be an efficient ionic detergent to solubilize the inclusion bodies, it could interfere during the purification step using the affinity column. The affinity column uses immobilized Ni^{+2} that sequesters histidine tagged proteins because the nickel cation is able to form a stable complex with four non-protonated histidine rings. Since SDS is negatively charged it would interact with Ni^{+2} ions and interfere with the binding of the histidine-tagged proteins. For this reason non-ionic detergents like n-octyl-β-D-glucoside and n-dodecyl-β-D-glucoside were also tried for solubilization of the inclusion bodies pellet. n-Octyl-β-D-glucoside or n-dodecyl-β-D-glucoside (2%) was dissolved in PBS and used in an attempt to resuspend the protein pellet. Although the lanes corresponding to n-octyl-β-D-glucoside-containing sample show the band of the M5I3M6 fusion protein (data not shown), inspection of the pellet

revealed that significant amounts of protein were still not resuspended. The same results were observed with n-dodecyl-β-D-glucoside (data not shown).

A pellet obtained from the isolation of the inclusion bodies was treated with CNBr in order to cleave the hydrophobic membrane protein from the TrpΔLE leader peptide and to confirm that the band within the range of 20-24 KDa was the targeted fusion protein. Figure 21 shows the results of the chemical cleavage of the M5I3M6- fusion protein and the BPTI fusion protein used as the positive control of the chemical reaction. The negative controls of the reactions were samples treated in the same way with the exception that they did not contain CNBr during the incubation of the reaction.

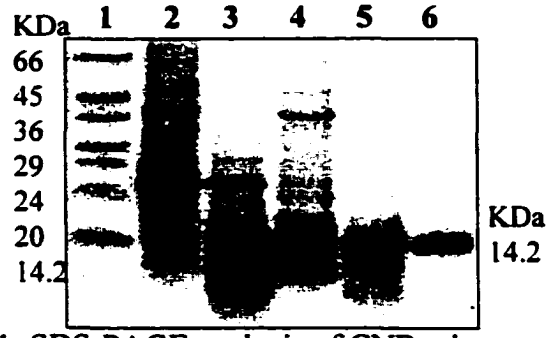

Figure 21: SDS-PAGE analysis of CNBr cleavage of TrpΔLE-M5I3M6 and TrpΔLE-BPTI.
Lane 1: Molecular marker: 66, 45, 36, 29, 24, 20, 14.2 KDa
Lane 2: TrpΔLE-BPTI from inclusion bodies not treated with CNBr
Lane 3: TrpΔLE-BPTI from inclusion bodies treated with CNBr
Lane 4: TrpΔLE-M5I3M6FP from inclusion bodies not treated with CNBr
Lane 5: TrpΔLE-M5I3M6FP from inclusion bodies treated with CNBr
Lane 6: Molecular marker band of 14.2 KDa

It can be observed in Figure 21 that the BPTI samples reacted as expected as the main band corresponding to the fusion peptide almost disappeared after the reaction

with CNBr and the reaction products showed clearly the bands of TrpΔLE and BPTI at 12 and 8 KDa, respectively. The unreacted sample corresponding to the M5I3M6-fusion protein showed a very faint band near 22 KDa (the other more prominent bands are additional proteins in the inclusion bodies), because of the low solubility of the fusion protein caused by aggregation that occurs after the lyophilization of the reaction solvent. The reaction products of the M5I3M6-fusion protein showed a clear band corresponding to the leader peptide (~12 KDa) and a smeared band in the region below. The smeared band may represent M5I3M6 (8,875.5 Da). This peculiar way of running of the M5I3M6 under the conditions of the SDS-PAGE seems to be a characteristic of membrane proteins (Kleinschmidt et al., 1999; Jones et al., 2000; Tatulian et al., 2000) and especially of the Ste2p transmembrane domains as reported by Xie et al. (2000). These results supported the conclusion that the expression of TrpΔLE-M5I3M6 and its isolation from inclusion bodies was feasible using the methodology outlined above.

Purification of M5I3M6-fusion protein was first attempted using the Ni^{+2} affinity column. Since TrpΔLE contains 9 histidines at the beginning of its amino acid sequence, purification using a nickel column would allow specific purification of the fusion protein from the inclusion bodies. Figure 22 shows the result of the purification using the immobilized Ni^{+2} column.

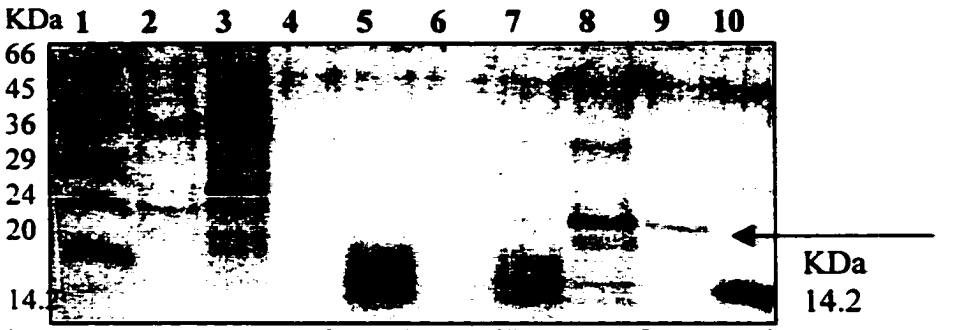

Figure 22: SDS-PAGE of the nickel affinity purification of TrpΔLE-M5I3M6.
Lane 1: Molecular marker: 66, 45, 36, 29, 24, 20, 14.2 KDa
Lane 2: Flow trough the column
Lane 3: Washing step with pH 6
Lane 4: Washing step with pH 5.3
Lane 5: Elution with pH 4
Lane 6: Elution with 50 mM imidazole
Lane 7: Elution with 200 mM imidazole
Lane 8: Elution with 350 mM imidazole
Lane 9: Elution with 500 mM imidazole
Lane 10: Molecular marker band of 14.2 KDa

Almost no protein in the range of 20-24 KDa is observed in the washing steps and in the elution with urea at pH 4. The only bands within the expected range are observed in lanes corresponding to elution with 200, 350, and 500 mM imidazole. Most of the fusion protein was eluted by 350 mM imidazole. The low molecular weight bands observed in the pH 4 elution and 200 mM imidazole lanes seem to correspond to degradation products of the fusion peptide that occurred during the isolation of inclusion bodies or during the purification process in the affinity column. The results of the affinity column further indicated that the isolated protein from inclusion bodies was indeed the target fusion protein of TrpΔLE-M5I3M6.

With the intention to have an internal control for expression, purification and biophysical studies, generation of a plasmid that expressed only the TrpΔLE was also accomplished following standard molecular biology procedures (Table 5). The resulting plasmid pEAP04 was used to generate strain EB012 for subsequent expression of the leader protein. Expression of the protein and its isolation from the inclusion bodies was achieved in the same fashion as for the TrpΔLE-M5I3M6. Purification of the TrpΔLE using the Ni^{+2} column yielded the purified protein in the 200, 350, and 500 mM imidazole fractions. Figure 23 shows the histidine tag detection results of the TrpΔLE and TrpΔLE-M5I3M6 purified through the affinity column.

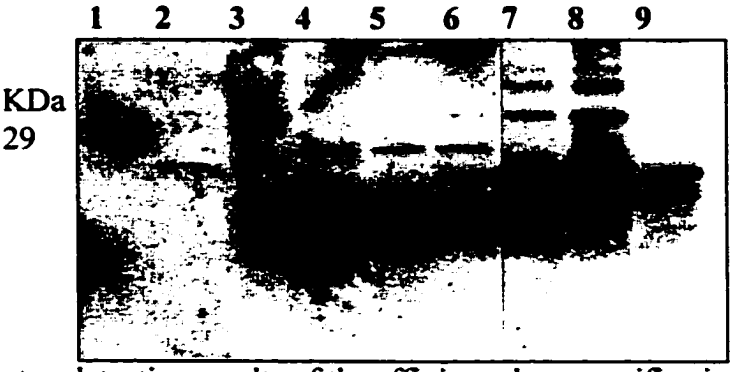

Figure 23: His-tag detection results of the affinity column purification of the TrpΔLE and TrpΔLE-M5I3M6 fusion proteins.
Lane 1: Molecular marker: 29 KDa
Lane 2: Aliquot of EB012 cells at OD_{600}=0.6 (no induction)
Lane 3: Aliquot of EB012 cells at OD_{600}=1.2 (no induction)
Lane 4: Elution of TrpΔLE with 200 mM imidazole
Lane 5: Elution of TrpΔLE with 350 mM imidazole
Lane 6: Elution of TrpΔLE with 500 mM imidazole
Lane 7: Elution of TrpΔLE-M5I3M6 with 350 mM imidazole
Lane 8: Elution of TrpΔLE-M5I3M6 with 500 mM imidazole
Lane 9: Elution of TrpΔLE-M5I3M6 with 3 M imidazole

The India HisProbe reacts exclusively with poly-histidine-tagged fusion proteins and reveals them via a horseradish peroxidase catalyzed reaction that results in chemiluminescence. The His-tag detection shows the difference in molecular size of the TrpΔLE and the fusion of the TrpΔLE with the M5I3M6 polypeptide. In this gel, no colored molecular markers were loaded. For this reason an accurate estimate of the molecular weights of the respective bands cannot be made. The regular molecular markers used in the gel showed only one band corresponding to 29 KDa. Apparently this particular protein marker also contained a region that cross-reacted with the India probe. Nevertheless, the difference in molecular size of the TrpΔLE and TrpΔLE-M5I3M6 is clearly noticeable in the gel. This result confirmed the expression, isolation and purification of the TrpΔLE polypeptide and the target fusion protein TrpΔLE-M5I3M6.

The fractions from the affinity chromatography were further purified by high performance liquid chromatography with an 2-propanol:1-butanol/H_2O system as specified in the Materials and Methods section. A chromatogram of the purification process is shown in Figure 24.

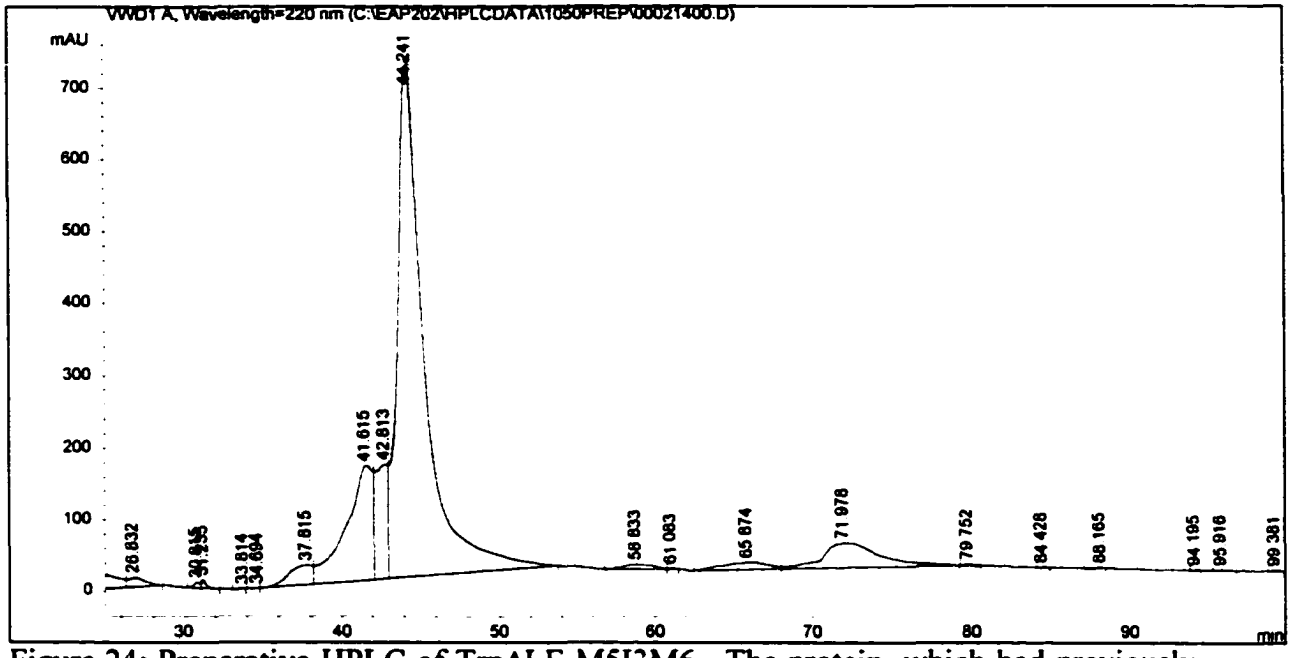

Figure 24: Preparative HPLC of TrpΔLE-M5I3M6. The protein, which had previously been purified through a Ni^{+2} column; was eluted from a C4 reversed phase column using water (0.1% TFA)/2-propanol:1-butanol (0.1% TFA) linear gradient. The gradient used for this sample was 10% 2-propanol:1-butanol (0.1% TFA) the first 10 min and a linear increase of 10% to 70% of the organic mixture the last 80 min.

It can be observed that the fraction corresponding to the 500 mM imidazole elution still contains some contaminants with retention times of 42, 43, and 72 min. The sample after purification by HPLC is depicted in Figure 25.

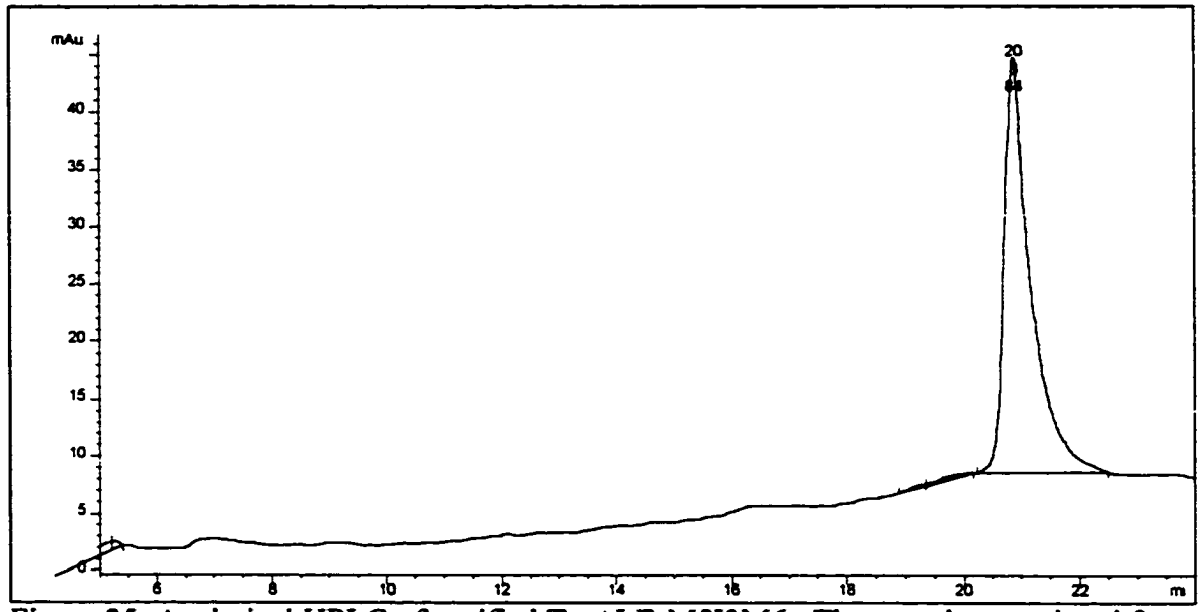

Figure 25: Analytical HPLC of purified TrpΔLE-M5I3M6. The sample was eluted from a C4 reversed phase column using water (0.1% TFA)/2-propanol:1-butanol (0.1% TFA) linear gradient. The gradient used for this sample was 10% 2-propanol:1-butanol (0.1% TFA) the first 5 min and a linear increase of 10% to 50% of the organic mixture the last 20 min.

After this HPLC fraction was lyophilized and loaded into SDS-PAGE, the fusion protein still showed 3 bands and a smeared region in between the upper bands (Figure 26, lane 2). However, when the same sample was loaded at different concentrations into the polyacrylamide gel, two of the bands disappear rapidly (Figure 26, lanes 4-6), indicating that the higher and the lower molecular weight bands were an effect of intermolecular and intramolecular aggregation.

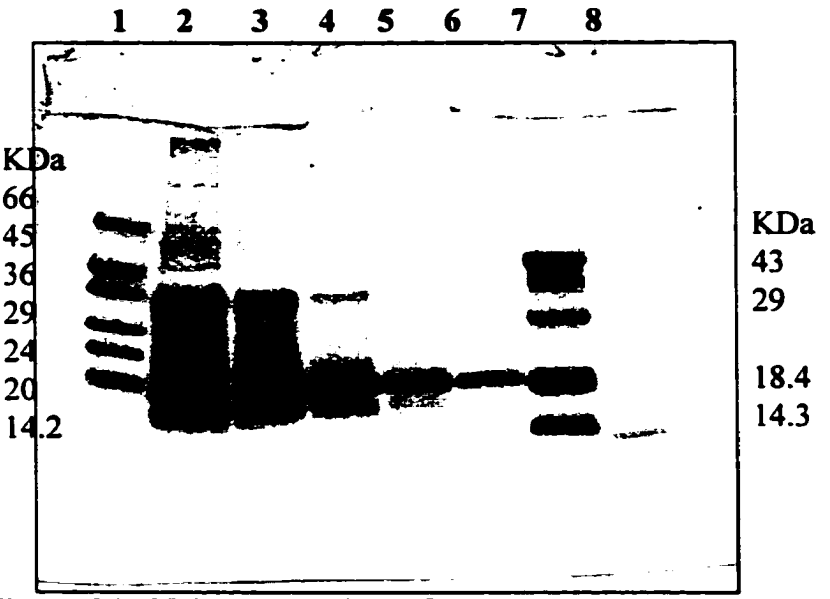

Figure 26: SDS-PAGE of purified TrpΔLE-M5I3M6. The gels were loaded with 20 μL of the fusion protein at different concentrations in order to find out the effect of the concentration in the behavior of the fusion protein during the electrophoresis.
Lane 1: Molecular marker: 66, 45, 36, 29, 24, 20, 14.2 KDa
Lane 2: Concentration of TrpΔLE-M5I3M6: 36 mg/mL
Lane 3: Concentration of TrpΔLE-M5I3M6: 18 mg/mL
Lane 4: Concentration of TrpΔLE-M5I3M6: 9 mg/mL
Lane 5: Concentration of TrpΔLE-M5I3M6: 4.5 mg/mL
Lane 6: Concentration of TrpΔLE-M5I3M6: 1.8 mg/mL
Lane 7: Molecular marker: 43, 29, 18.4, 14.3 KDa
Lane 8: Molecular marker: 14.2 KDa

Densitometric analysis of the gel supports this conclusion as the absorption values of the band at ~40 KDa decays rapidly with sample concentration. In contrast, the 22 KDa band decreases in a nearly linear fashion with change of the concentration of the sample (Figure 27).

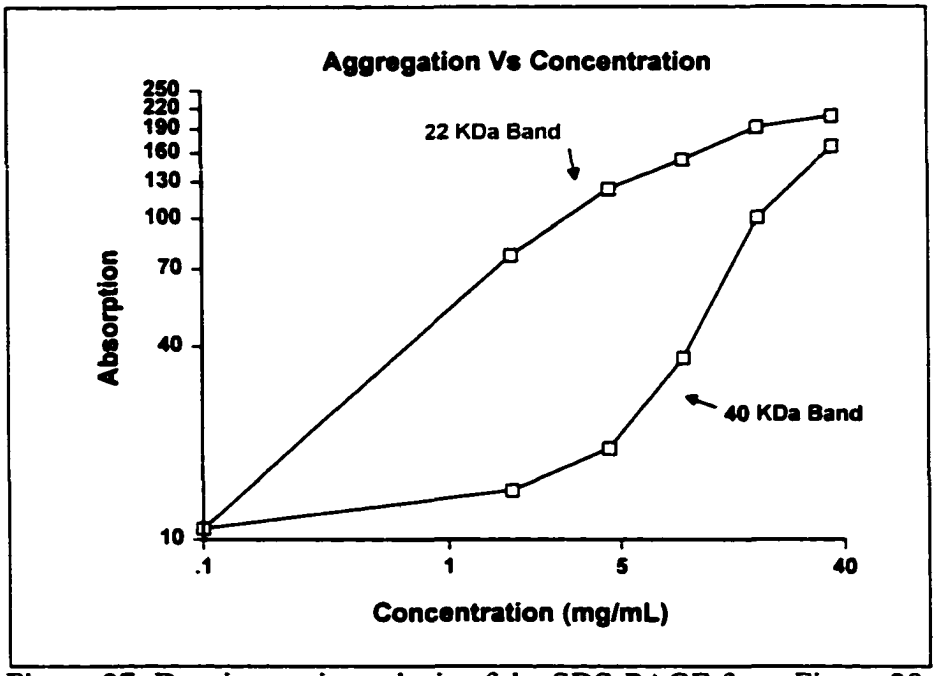

Figure 27: Densitometric analysis of the SDS-PAGE from Figure 28. Plot of the densitometric values (absorption) versus concentration of the sample loaded in each lane.

This result clearly indicates that the unexpected bands of the purified sample were artifacts of the sample as a consequence of its tendency to aggregate. The tendency of TrpΔLE-M5I3M6 to aggregate with other molecules and with itself is determined by the M5I3M6 region, since the TrpΔLE itself does not show this particular behavior (Figure 23). All of these results demonstrate the plausibility of expressing two transmembrane domains of Ste2p in bacterial system like *E. coli* and of isolating and purifying the target proteins using affinity chromatography and reversed phase HPLC.

Similar to the cloning of M5I3M6 into the pMMHa plasmid, the single transmembrane domain of Ste2p (M6) was cloned into the parent plasmid and the newly

generated plasmid, pSW02 (Table 5) was sequenced to confirm the presence and the correct insertion of the DNA sequence corresponding to the M6 domain. Transformation of BL21(DE3)pLysS with pSW02 generated strain EB011 (Table 5), which was used for expression of the TrpΔLE-M6 fusion protein as described above. Isolation and purification of the M6 fusion protein was also carried out by separation of the inclusion bodies and by HPLC. Figure 28 shows the Coomassie staining of the SDS-PAGE loaded with purified TrpΔLE, TrpΔLE-M6 and TrpΔLE-M5I3M6.

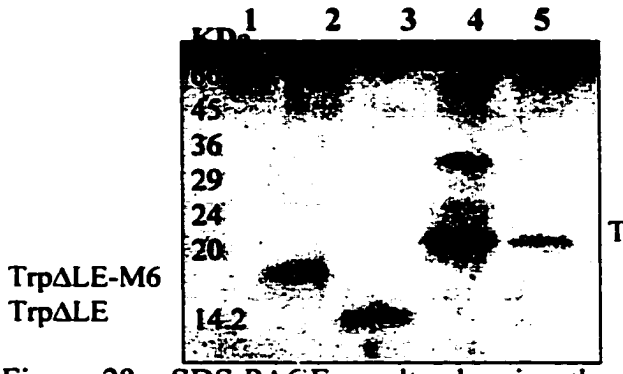

Figure 28: SDS-PAGE results showing the fusion proteins and TrpΔLE. The gel depicts the different molecular weights of TrpΔLE (lane3), TrpΔLE-M6 (lane 2) and TrpΔLE-M5I3M6 (lanes 4 and 5).
Lane 1: Molecular weight markers: 66, 45, 36, 29, 24, 20, 14.2 KDa
Lane 2: TrpΔLE-M6
Lane 3: TrpΔLE
Lane 4: TrpΔLE-M5I3M6, 1x dilution with sample buffer
Lane 5: TrpΔLE-M5I3M6, 10x dilution with sample buffer

The difference in molecular weight of the three samples is evident, as is the tendency of TrpΔLE-M5I3M6 to aggregate at higher concentrations (compare lanes 4 and 5). Both the M6 and M5I3M6 fusion proteins show bands at the interface between

the stacking gel and the separating gel further indicating the tendency of these samples to aggregate even in the presence of SDS. Figure 29 represents the His-tag detection of the same samples and confirms the identity of the bands seen in the acrylamide gel stained with Coomassie blue.

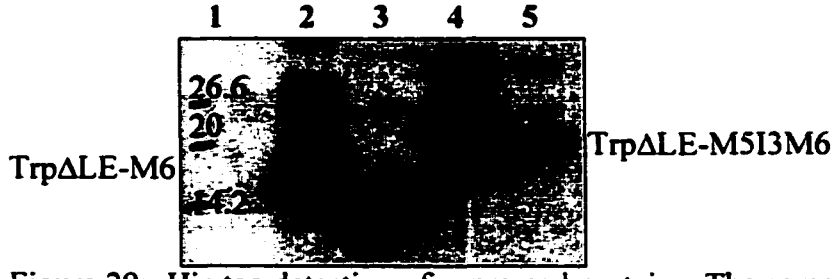

Figure 29: His-tag detection of expressed proteins. The same samples were loaded in Figure 28.
Lane 1: Molecular weight markers: 26.6, 20, 14.2 KDa
Lane 2: TrpΔLE-M6
Lane 3: TrpΔLE
Lane 4: TrpΔLE-M5I3M6, 1x dilution with sample buffer
Lane 5: TrpΔLE-M5I3M6, 10x dilution with sample buffer

The higher molecular weight species of the M6 and M5I3M6 fusion proteins are clearly revealed by the India probe.

Purification of TrpΔLE-M6 fusion protein was also carried out using affinity chromatography and it was again observed that the imidazole fractions were at most 80-85% pure. This level of purity indicated the necessity for further purification by HPLC. For this reason direct injection of the resuspended inclusion bodies into the HPLC instrument was attempted in order to increase the efficiency of purification. Table 8 summarizes the amount of material obtained in each step of the purification process per

liter of culture.

Table 8

Summary of Recoveries of Materials from Expression and Isolation of TrpΔLE-Fusion Proteins

	Weight of Material Isolated During TrpΔLE-M5I3M6 expression (rich media)*	Weight of Material isolated during TrpΔLE-M6 expression (minimal media)*
Wet cells	3.646 g	3.696 g
Wet inclusion bodies	0.163 g	0.196 g
HPLC purification	2.6 mg	3.7 mg

* Based on one liter fermentation

The HPLC fractions containing the M6 and M5I3M6 fusion proteins were frequently higher than 85% homogeneous, indicating that direct purification of inclusion bodies by HPLC is an efficient approach to isolate the desired proteins. The values in Table 8 represent the average of all expressions carried in rich media (LB) and minimal media (M9) for M5I3M6 and M6 fusion proteins, respectively. The HPLC chromatograms of pure TrpΔLE-M6 and pure TrpΔLE-M5I3M6 fusion proteins are represented in Figures 30 and 31, respectively. The chromatograms indicate that fusion proteins with >93% homogeneity can be obtained.

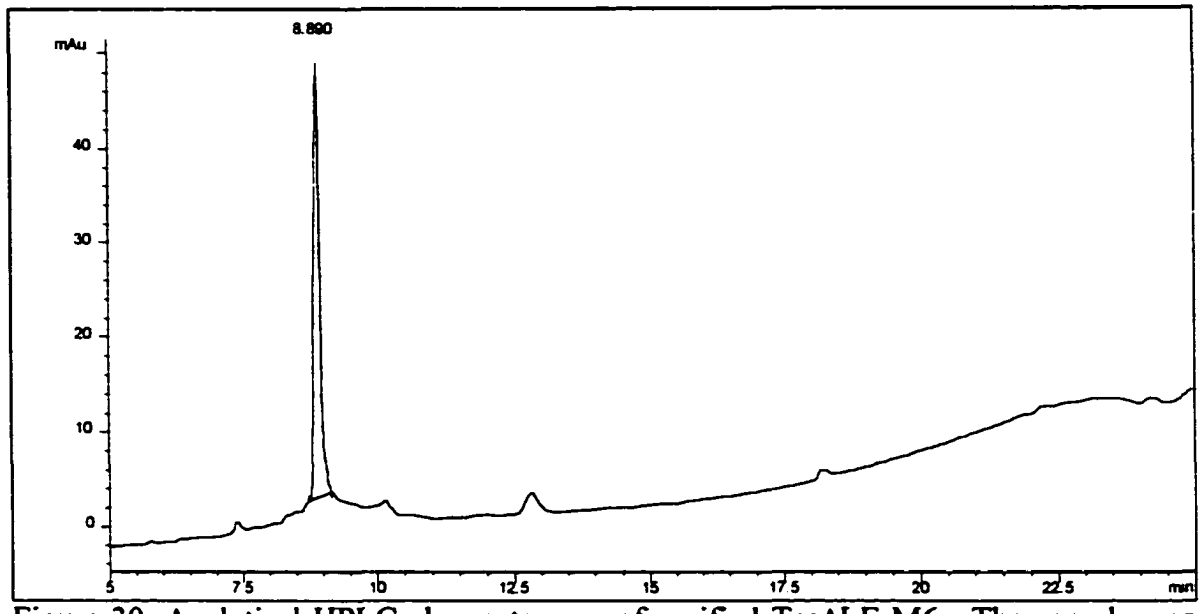

Figure 30: Analytical HPLC chromatogram of purified TrpΔLE-M6. The sample was eluted from a C4 reversed phase column using water (0.1% TFA)/ acetonitrile (0.1% TFA) linear gradient. The gradient used for elution of this sample was from 30% to 100% acetonitrile (0.1% TFA) the first 20 min and 100% acetonitrile (0.1% TFA) the last 5 min.

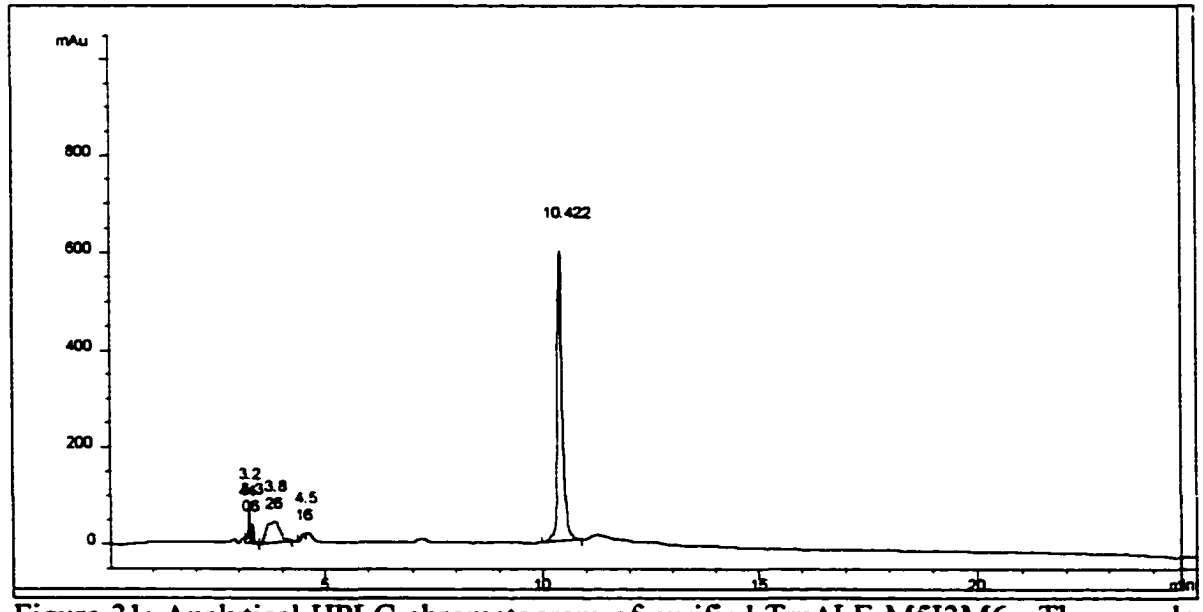

Figure 31: Analytical HPLC chromatogram of purified TrpΔLE-M5I3M6. The sample was eluted from a C4 reversed phase column using water (0.1% TFA)/ acetonitrile (0.1% TFA) linear gradient. The gradient used for elution of this sample was from 30% to 100% acetonitrile (0.1% TFA) the first 20 min and 100% acetonitrile (0.1% TFA) the last 5 min.

CNBr cleavage of ^{15}N-TrpΔLE-M6 and purification of the ^{15}N-M6 peptide

The ultimate goal of this line of research is to produce domains of Ste2p that can be subjected to high resolution structural analysis using NMR spectroscopy. Our experience with Ste2p single transmembrane domains indicates that such studies will require isotopically labeled peptides. In order to develop procedures for this goal the expression of TrpΔLE-M6 polypeptide in M9 minimal media containing $^{15}NH_4Cl$ as a sole source of nitrogen was evaluated. This procedure generated a fusion-peptide containing ^{15}N in all of the amino acids of the primary structure of the fusion protein. The labeled peptide was isolated in the inclusion bodies after cell lysis and purified by HPLC as detailed in the previous section of the results. Chemical cleavage with CNBr was carried out following the procedure outlined in the Materials and Methods section. Figure 32 shows the HPLC chromatogram of the purified ^{15}N-M6. From 20.85 mg of ^{15}N-TrpΔLE-M6 (80% pure) cleaved with CNBr approximately 1 mg of ^{15}N-M6 peptide was obtained. The amount of isotopically labeled peptide represents 28% recovery of the expected amount of peptide.

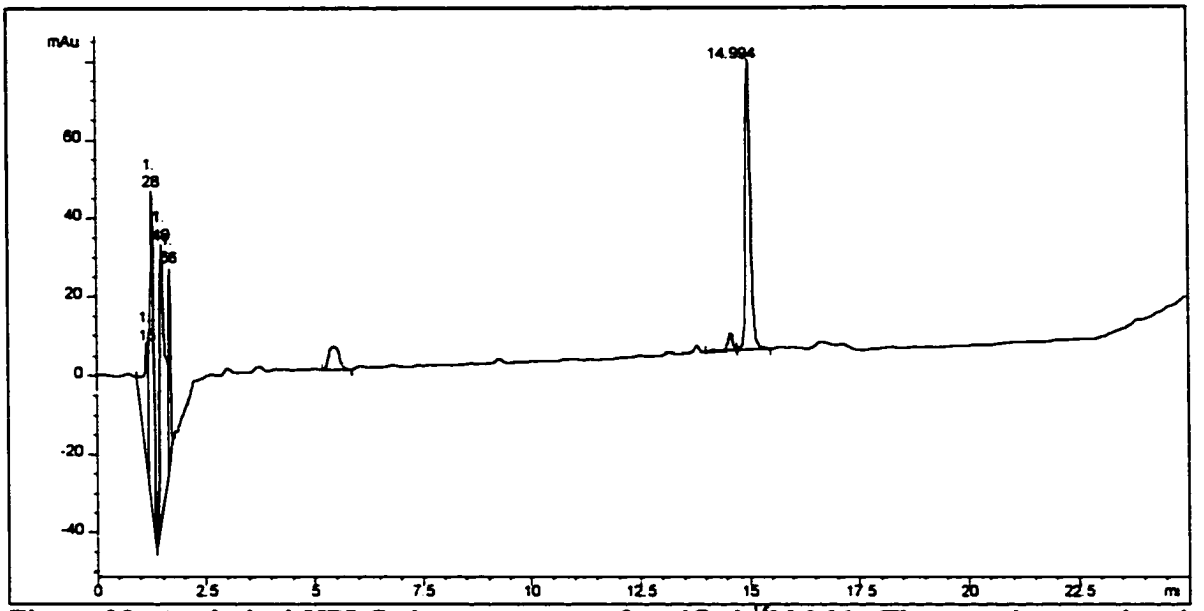

Figure 32: Analytical HPLC chromatogram of purified ^{15}N-M6. The sample was eluted from a C4 reverse phase column using water (0.1% TFA)/acetonitrile (0.1% TFA) linear gradient. The gradient used for elution of this sample was from 30% to 60% acetonitrile (0.1% TFA) for the first 20 min and from 60% to 100% acetonitrile (0.1% TFA) for the last 5 min.

Characterization of the M6 and M5I3M6 fusion proteins

Mass spectrometry results of the TrpΔLE-M6, ^{15}N-M6, and the TrpΔLE-M5I3M6 agreed within 2 Da with the theoretical mass of the fusion peptides. Figure 33 shows the mass spectrum of the unlabeled TrpΔLE-M6 fusion protein. The molecular mass, 17,313 Da, corresponds to the theoretical mass.

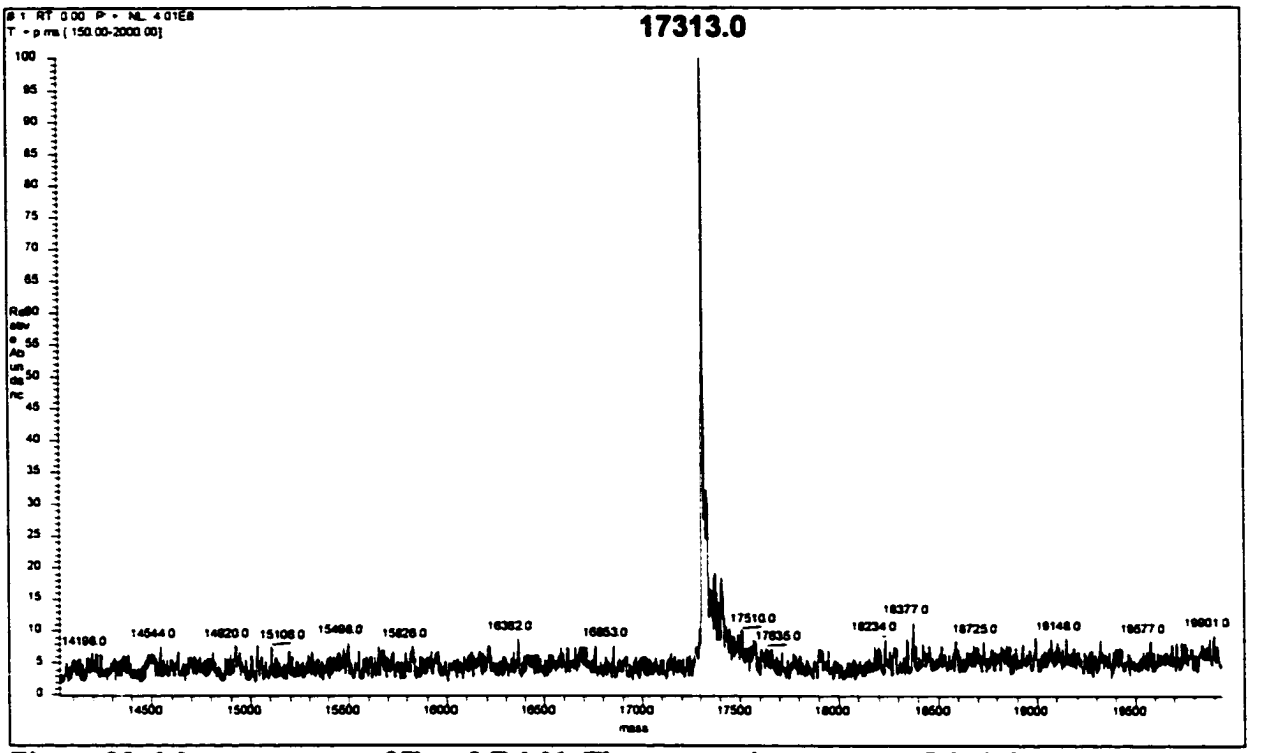

Figure 33: Mass spectrum of TrpΔLE-M6. The expected mass was 17,313 Da.

Figure 34 shows the non-deconvoluted mass spectrum of the ^{15}N-M6 peptide. Four peaks of molecules with charges (z) [+6, +5, +4, +3] correspond to the m/z values of 632.6, 758.9, 948.4, and 1264.3, respectively. Calculation of the mass from the formula {[(m/z) x z] – z} yielded an average value of 3,789.65 Da. Therefore, the mass of the peptide enriched with ^{15}N also agreed with the theoretical value.

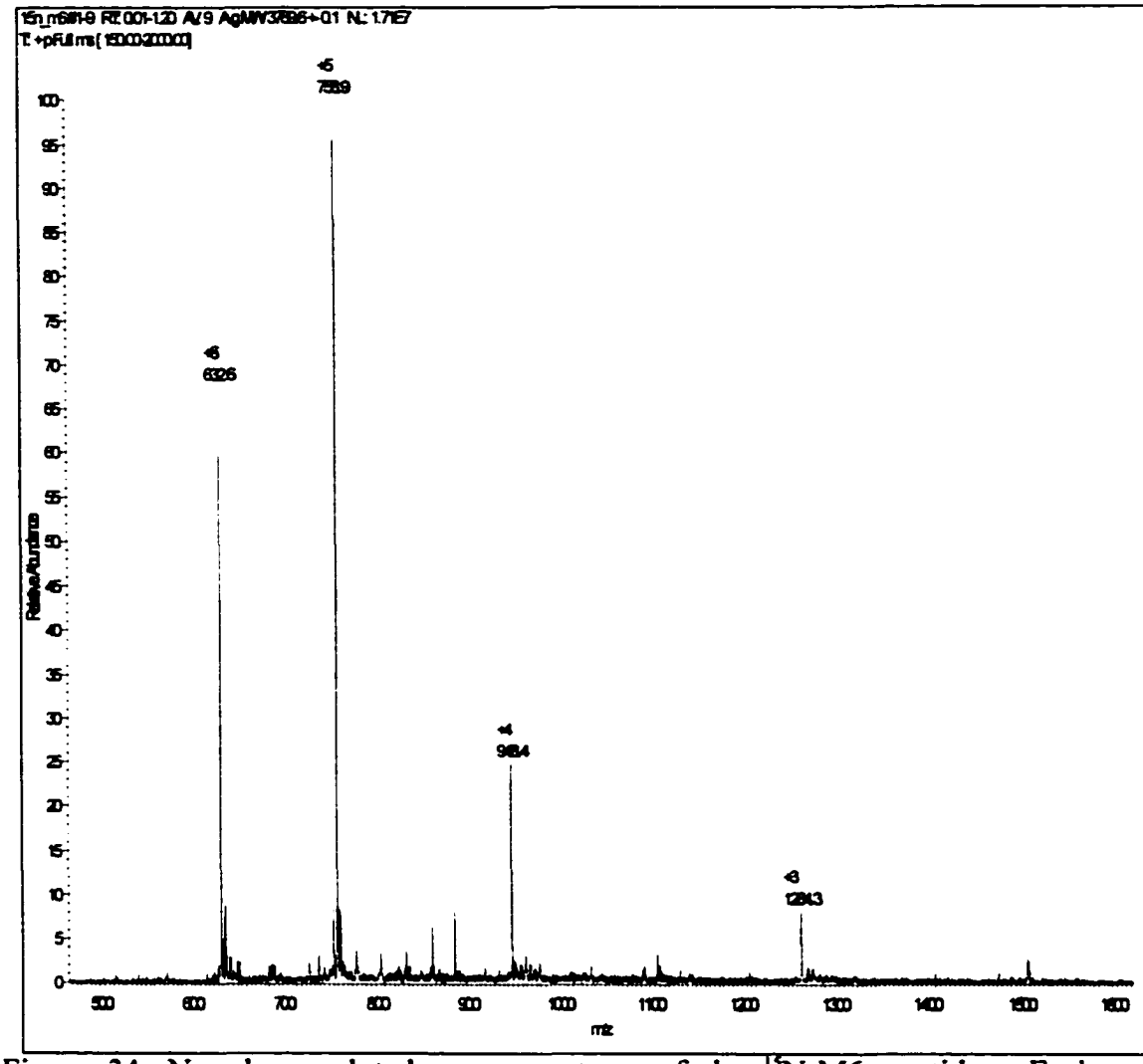

Figure 34: Non-deconvoluted mass spectrum of the ^{15}N-M6 peptide. Each peak represents a specific mass-to-charge (*m/z*) ratio from the electrospraying ionization. The calculation of the mass of ^{15}N-M6 yielded 3,789.6 Da.

Figure 35 depicts the results of the mass spectrometry of the TrpΔLE-M5I3M6. The obtained mass for this fusion protein lies within 2 Da of the expected mass (22,438 Da).

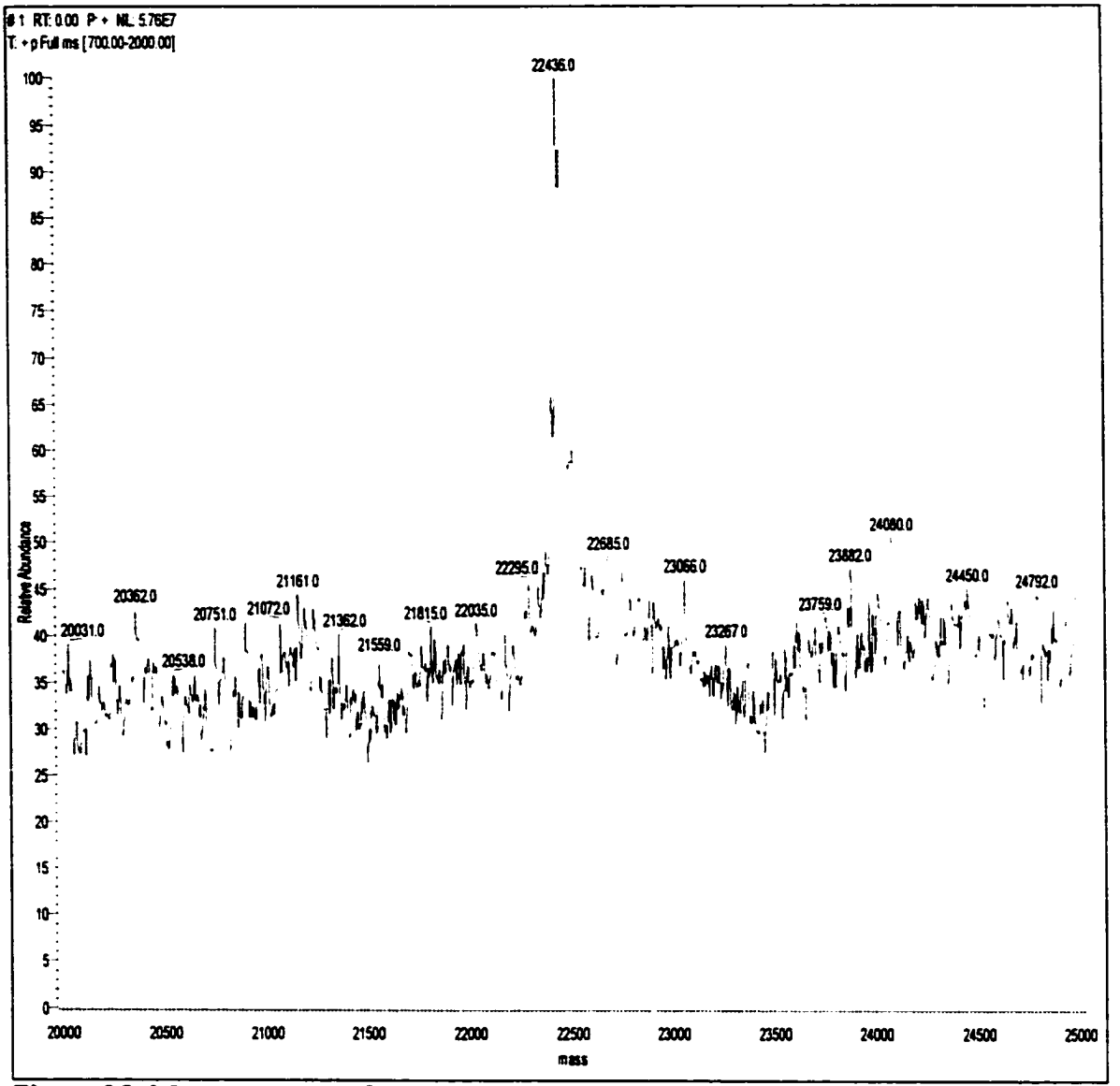

Figure 35: Mass spectrum of TrpΔLE-M5I3M6. The expected mass was 22,438 Da.

Table 9 summarizes the mass spectrometry analysis of the peptides and proteins.

Table 9

Mass Spectrometry of Membrane Proteins and Peptides

Protein	Observed mass (Da)	Expected mass (Da)	% Error
TrpΔLE-M6	17,313	17,313	-
^{15}N-TrpΔLE-M6	17,539	17,540	0.005
TrpΔLE-M5I3M6	22,436	22,438	0.008
^{15}N-M6	3,789.6	3790.0	0.01

Amino acid analysis was carried on the TrpΔLE-M6 and TrpΔLE-M5I3M6 to confirm the identity of the isolated fusion proteins by inspection of their amino acid composition. A bar chart depicts the observed number of residues of each amino acid and the respective expected value in TrpΔLE-M6 (Figure 36).

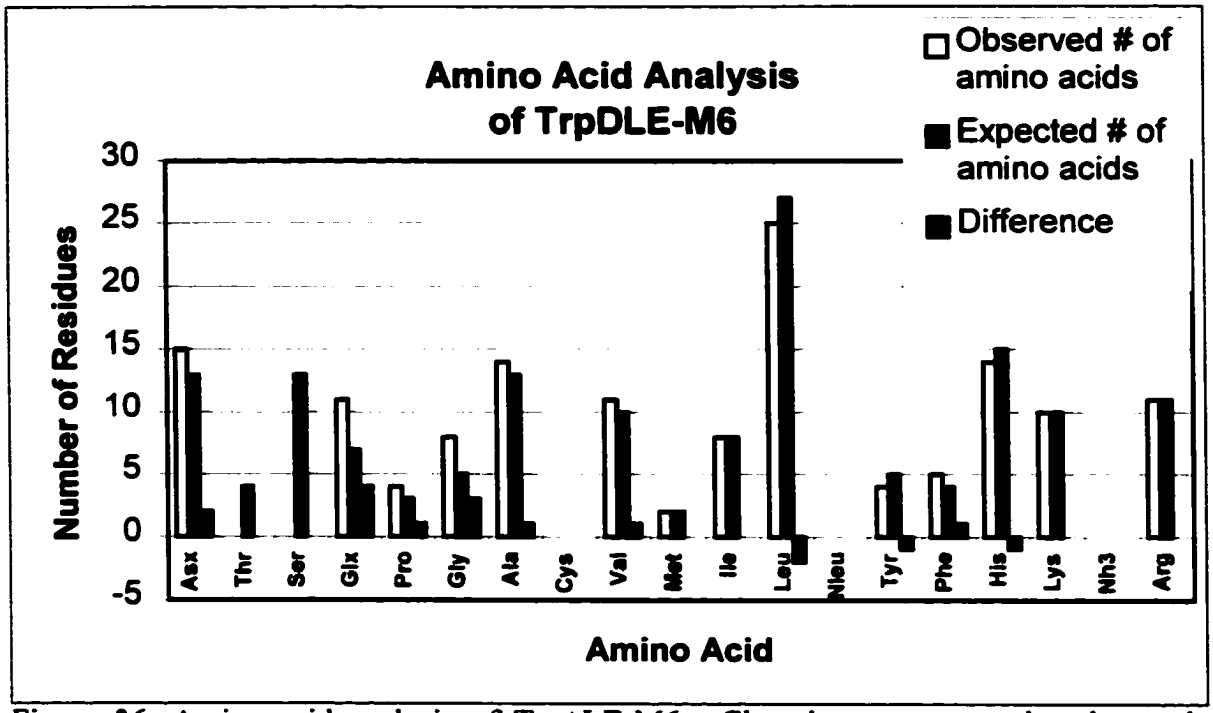

Figure 36: Amino acid analysis of TrpΔLE-M6. Clear bars represent the observed number of amino acids after 72 h of acid hydrolysis. Dark bars represent the number of amino acids expected from the primary structure. Gray bars depict the difference between the observed and the expected number of amino acids.

Except for serines and threonines, most of the values found for the other amino acid residues were close to their theoretical values, given the expected experimental error of this procedure (10% to 20%). Similar results were found in TrpΔLE-M5I3M6 (Figure 37). We observed that Ser and Thr residues are often degraded during the 3-day hydrolysis used for membrane peptides (Arshava et al., 1998). Therefore, together with the mass spectrometry analysis, the amino acid ratios confirm the structure of the fusion proteins.

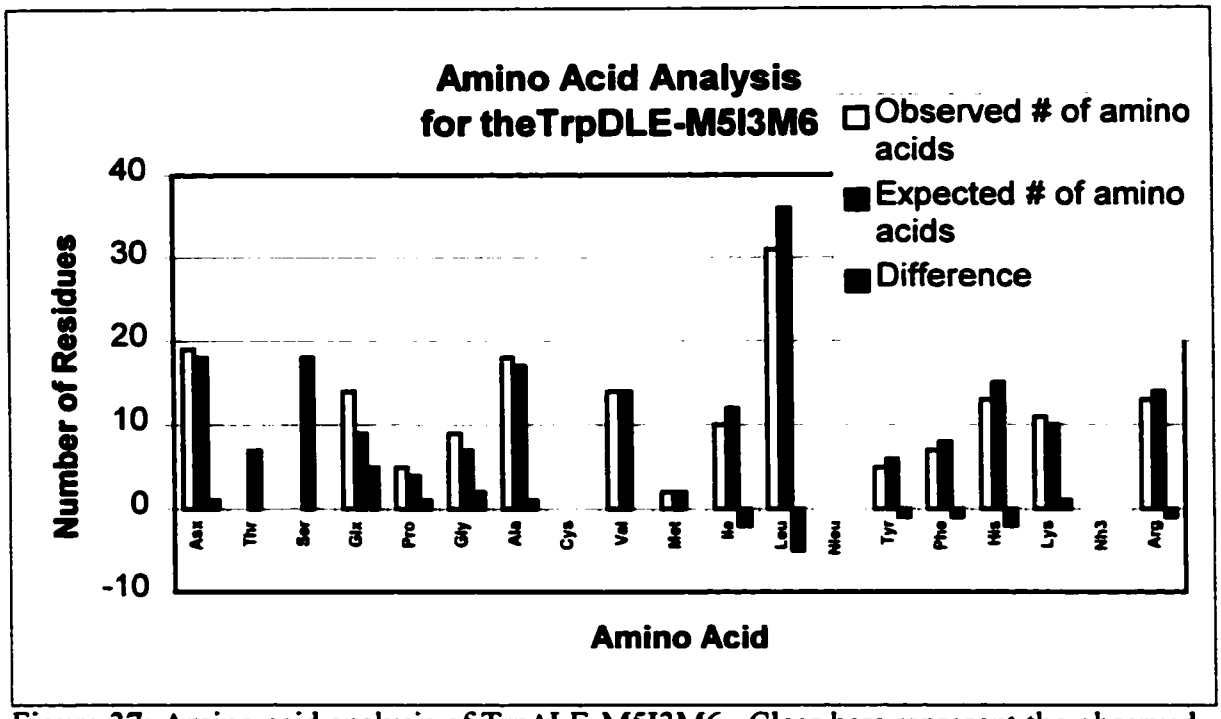

Figure 37: Amino acid analysis of TrpΔLE-M5I3M6. Clear bars represent the observed number of amino acids after 72 h of acid hydrolysis. Dark bars represent the number of amino acids expected from the primary structure. Gray bars depict the difference between the observed and the expected number of amino acids.

Biophysical studies of the fusion peptides: circular dichroism spectroscopy

Figures 38 and 39 show CD profiles of TrpΔLE-M5I3M6 and TrpΔLE in trifluoroethanol, a putative membrane mimetic organic solvent.

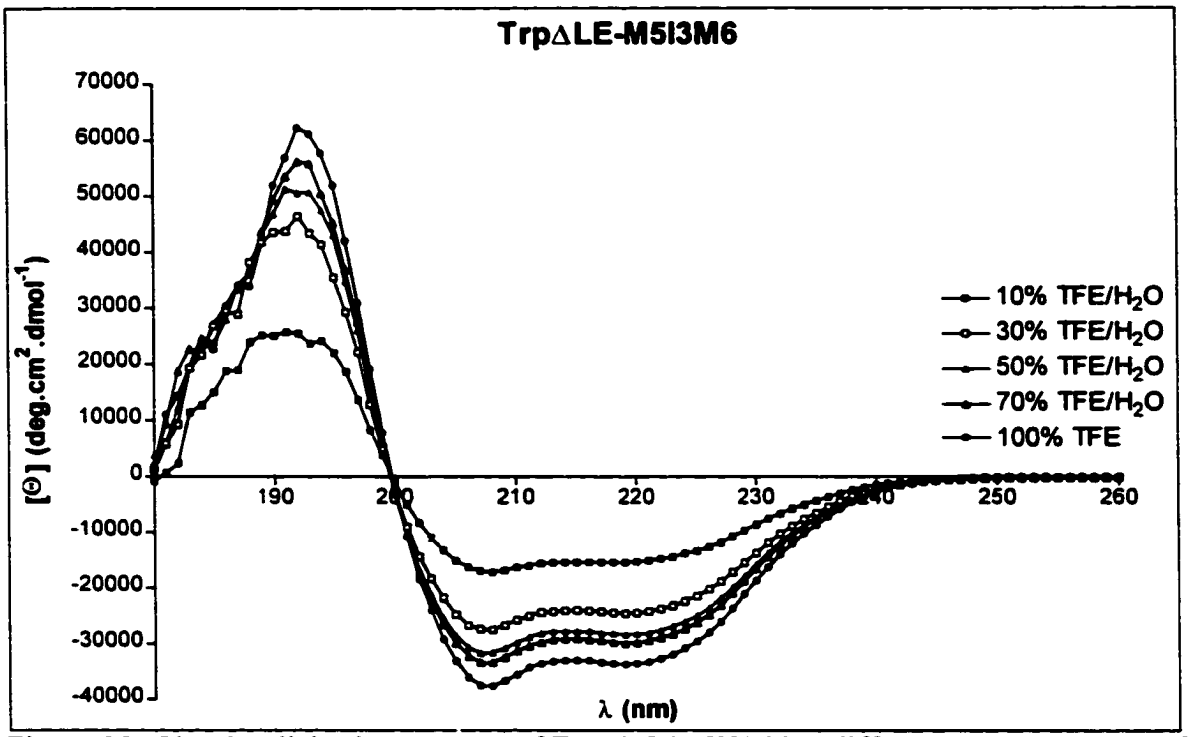

Figure 38: Circular dichroism spectra of TrpΔLE-M5I3M6 at different concentrations of TFE in H_2O. The stock solution was prepared by dissolving the protein sample in TFE to a final concentration of 3 mg/mL. The resulting stock was used to prepare samples with final protein concentrations of 0.3 mg/mL at different TFE concentrations in H_2O.

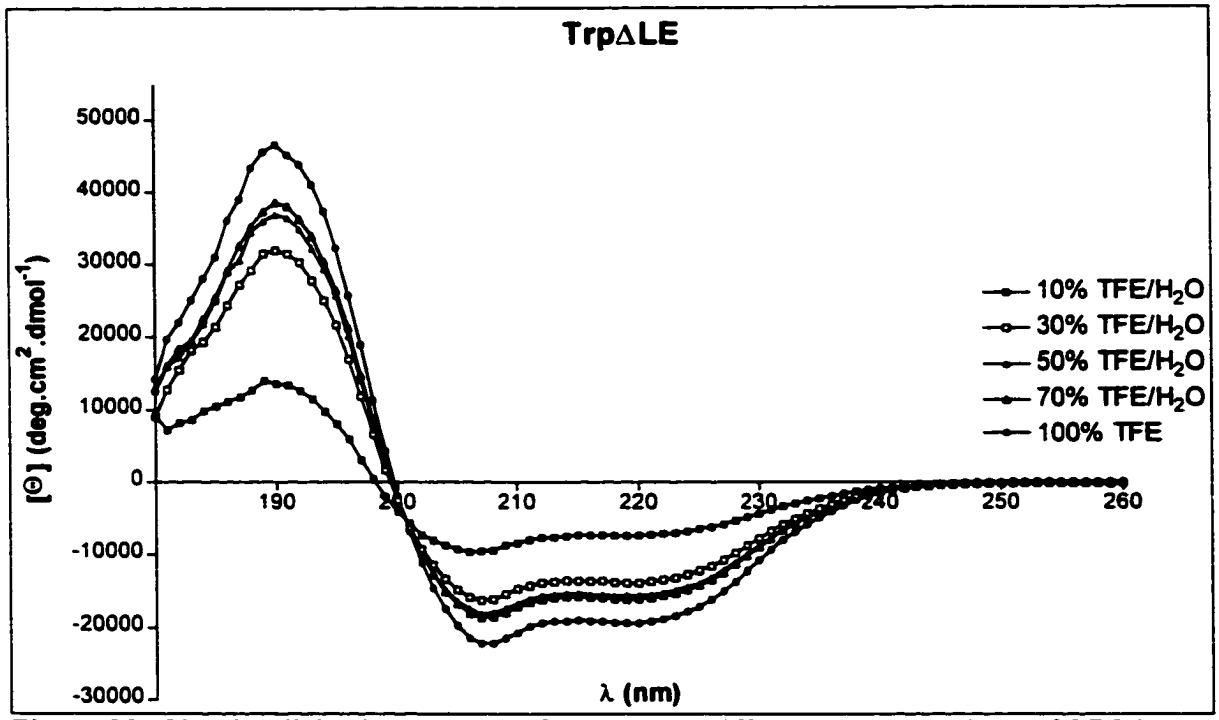

Figure 39: Circular dichroism spectra of TrpΔLE at different concentrations of TFE in H_2O. The samples were prepared as indicated in Figure 38.

Direct comparison of the mean residue ellipticity values shows a higher helical content for the TrpΔLE-M5I3M6 with respect to the TrpΔLE polypeptide. This indicates that the M5I3M6 peptide significantly contributes to the increase in helical structure in the fusion protein. At different concentrations of TFE in water both polypeptides adopt a mostly helical conformation, but the helical content decreases with the decrease of the percentage of TFE. With the assumption that the TrpΔLE and the hydrophobic domains behave independently from each other, the dichroic content of TrpΔLE can be subtracted from that of TrpΔLE-M5I3M6 or of TrpΔLE-M6 to obtain information related to the M5I3M6 or the M6 peptides. This assumption was considered and discussed previously by Venyaminov and Yang (1996). Figure 40 shows the results

of the calculation of the new spectra for TrpΔLE-M5I3M6.

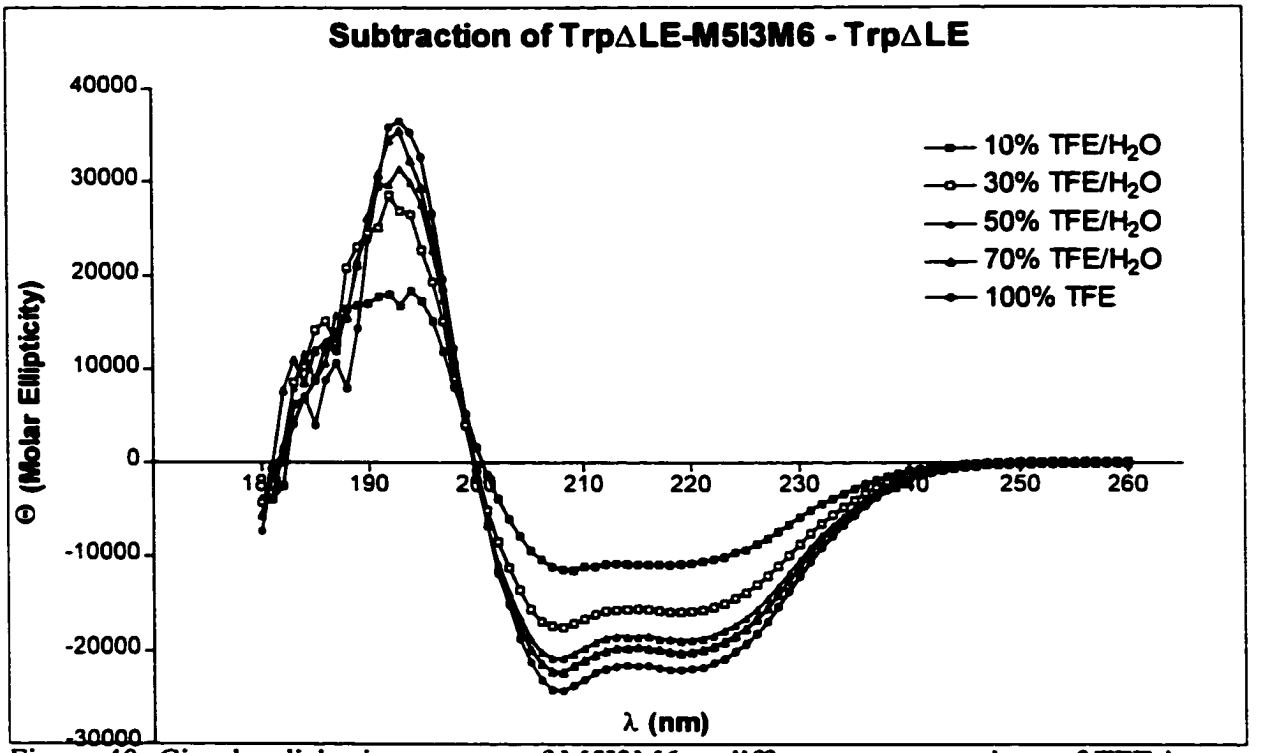

Figure 40: Circular dichroism spectra of M5I3M6 at different concentrations of TFE in H_2O. The spectra were obtained after subtraction of TrpΔLE spectra from that of TrpΔLE-M5I3M6 as indicated in the Materials and Methods section.

The TrpΔLE-M6 fusion protein was also studied in different ratios of TFE:H_2O (Figure 41).

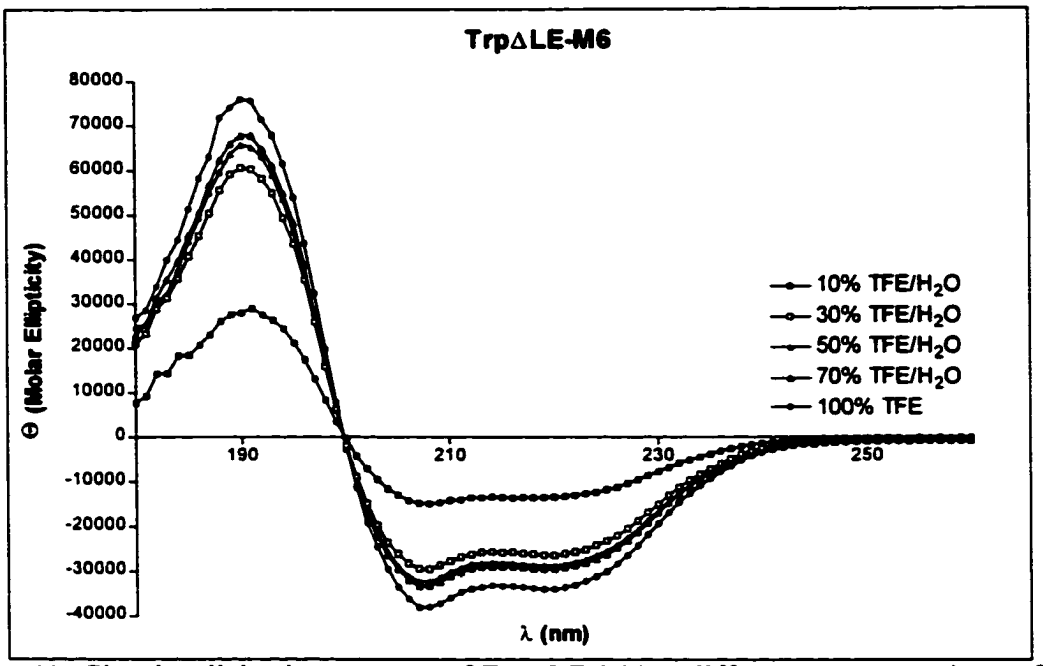

Figure 41: Circular dichroism spectra of TrpΔLE-M6 at different concentrations of TFE in H_2O. The samples were prepared as indicated in Figure 38.

Similar to the results obtained with TrpΔLE and TrpΔLE-M5I3M6, the fusion protein of M6 shows a typical helix profile in the circular dichroism spectrum. Subtraction of the TrpΔLE spectrum from the TrpΔLE-M6 spectrum also resulted in a helical profile that would correspond to the M6 peptide (Figure 42).

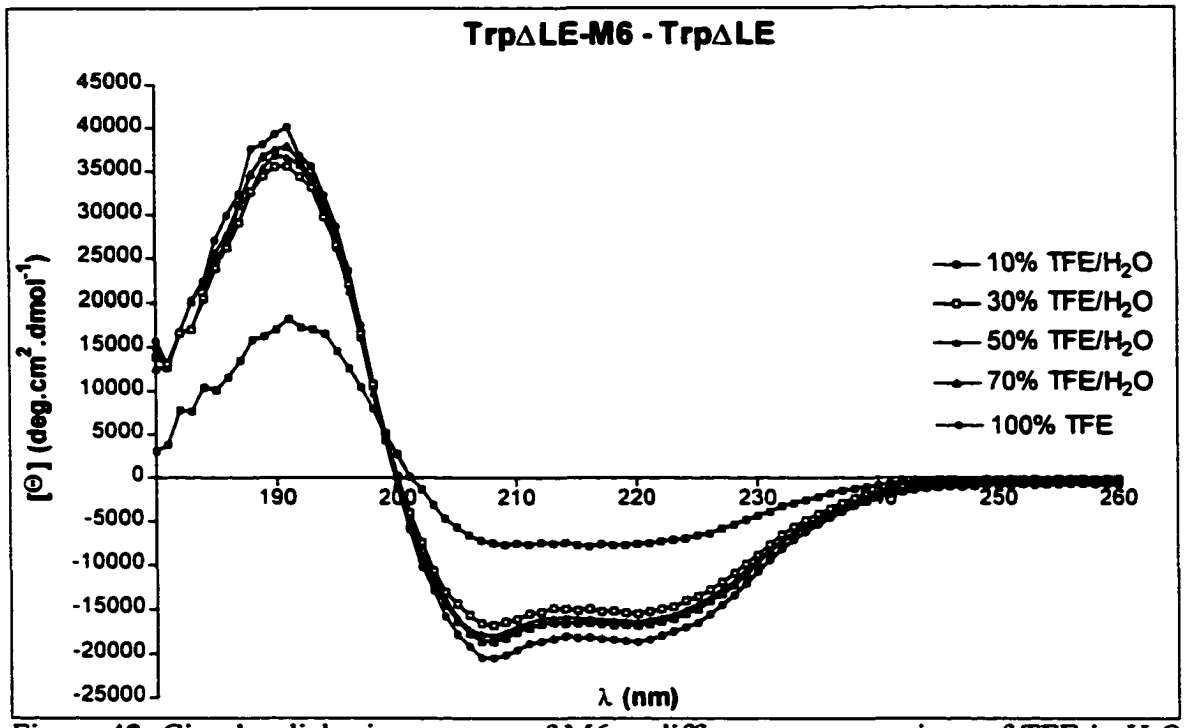

Figure 42: Circular dichroism spectra of M6 at different concentrations of TFE in H_2O. The spectra were obtained after subtraction of TrpΔLE spectra from that of TrpΔLE-M6 as indicated in the Materials and Methods section.

The maximum positive rotational strength of TrpΔLE was found to be around 190 nm (Figure 39), but for both fusion proteins it was red-shifted toward 195 nm (Figures 38 and 41). This effect is in accord with the observations made by Fasman (1996) about the circular dichroism characteristics of membrane proteins.

Quantitative analysis of the amount of helical structure of all the fusion proteins and the leader peptide was carried out using methods based on model peptides and the application of the backpropagation neural network approach (Bohm et al., 1992). Greenfield and Fasman (1969) proposed a quantitative method to calculate the amount of secondary structure from the CD spectra of poly-L-lysine under various conditions.

The CD spectrum (S(λ)) can be analyzed as a linear combination of basis spectra according to the following equation:

$$S(\lambda)=f_H B_H(\lambda) + f_\beta B_\beta(\lambda) + f_R B_R(\lambda) \qquad (1)$$ (Greenfield et al., 1969)

where B_k is the basis spectrum, H = α helix, β = β-sheet, R = random coil, and f_k is the fraction of kth secondary structure. Since

$$\Sigma f_k = 1 \qquad (2)$$

the estimation of the protein secondary structure can be carried out by a least-squares fit of the experimental CD spectrum of the protein according to equation (1).

Assuming the existence of only helical and random coil structures (f_β = 0), the percentage of helical structure can be calculated as follows:

$$\% \text{ Helical Structure} = ([\theta_\lambda] - [\theta_\lambda]_0)/([\theta_\lambda]_H - [\theta_\lambda]_0) \times 100 \qquad (3),$$

where $[\theta_\lambda]$ is the mean residue ellipticity of the polypeptide at wavelength (λ). $[\theta_\lambda]_0$ is the mean residue ellipticity of an unstructured polypeptide (-2000 deg cm^2 mol^{-1} at 222 nm and 4000 deg cm^2 mol^{-1} at 208 nm). $[\theta_\lambda]_H$ is the mean residue ellipticity of an infinite helix (Chen et al., 1974). The computed dichroic strengths of an infinite helix for λ= 208 and 222 nm were calculated to be –33,800 and –39,500 deg cm^2 mol^{-1}, respectively. Wu et al. (1981) revised the mean residue ellipticity value of the infinite helix at 222 nm to –37,400 deg cm^2 mol^{-1}.

Tables 10 and 11 summarize the amount of the helical content calculated using equation (3) (Greenfield et al., 1969; Chen et al., 1974; Wu et al., 1981).

Table 10

Percentage of Helical Structure in Membrane Peptides Calculated Using the Mean Residue Ellipticity at 222 nm of Model Peptides. (Greenfield et al., 1969; Chen et al., 1974; Wu et al., 1981)

TFE%	TrpΔLE-M5I3M6	TrpΔLE	M5I3M6*	TrpΔLE-M6	M6*
100	87	48	54	87	45
70	76	38	49	75	40
50	72	37	46	73	39
30	61	32	38	67	37
10	36	15	23	31	15

* The percentages of M5I3M6 and M6 are based on the spectra obtained mathematically. The subtraction method is described in the Materials and Methods section.

Table 11

Percentage of Helical Structure in Membrane Peptides Calculated Using the Mean Residue Ellipticity at 208 nm of Model Peptides. (Greenfield et al., 1969; Chen et al., 1974; Wu et al., 1981)

TFE%	TrpΔLE-M5I3M6	TrpΔLE	M5I3M6*	TrpΔLE-M6	M6*
100	116	63	70	117	57
70	101	50	63	100	50
50	95	48	58	97	48
30	81	42	47	88	44
10	45	19	26	38	12

* The percentages of M5I3M6 and M6 are based on the spectra obtained mathematically. The subtraction method is described in the Materials and Methods section.

The standard deviation in the helical structure estimations using 222 nm and 208 nm ellipticities vary by 2-21%. The M5I3M6 and M6 fusion proteins show helical structures higher than 100% in pure TFE. This indicates that this method overestimates the amount of secondary structure in these membrane proteins. This is not surprising since the theoretical mean residue ellipticity values used in this empirical method were based on model peptides. Park and others (1992) found that the ellipticities for membrane proteins at 208 and 222 nm were –50,000 and –60,000 deg cm^2 $dmol^{-1}$, respectively. The difference in the rotational strength observed in membrane proteins may result from the immersion of the membrane proteins in a lower dielectric medium compared to the interior of soluble proteins. This affects the $\pi\pi^*$ and $n\pi^*$ transitions of the amide chromophore (Fasman, 1996). The percentages of helical structure using the mean residue ellipticity values determined for membrane peptides (Park et al., 1992) are summarized in Tables 12 and 13.

Table 12

Percentage of Helical Structure in Membrane Peptides Calculated Using the Mean Residue Ellipticity at 222 nm of Membrane Proteins. (Park et al., 1992)

TFE%	TrpΔLE-M5I3M6	TrpΔLE	M5I3M6*	TrpΔLE-M6	M6*
100	53	29	33	53	28
70	46	23	30	46	25
50	44	23	28	45	24
30	37	20	23	41	22
10	22	9	14	19	9

* The percentages of M5I3M6 and M6 are based on the spectra obtained mathematically. The subtraction method is described in the Materials and Methods section.

Table 13

Percentage of Helical Structure in Membrane Peptides Calculated Using the Mean Residue Ellipticity at 208 nm of Membrane Proteins. (Park et al., 1992)

TFE%	TrpΔLE-M5I3M6	TrpΔLE	M5I3M6*	TrpΔLE-M6	M6*
100	73	40	44	74	36
70	64	31	40	63	32
50	60	30	36	61	31
30	51	26	29	55	28
10	29	12	16	24	8

* The percentages of M5I3M6 and M6 are based on the spectra obtained mathematically. The subtraction method is described in the Materials and Methods section.

The calculation of the helical content using the maximum rotational strength

reported by Park and others (1992) yielded lower percentages of secondary structure in each of the polypeptides compared to the percentages obtained using the values of model peptides. The percentages of helical structure calculated using the values at 208 nm were higher than those at 222 nm. However, the standard deviation of the calculated helical structure based on the values at 222 and 208 nm are less than or equal to 14%.

The sixth transmembrane domain (M6) fused to TrpΔLE is composed of 33 amino acids, and the results from the CD experiments on this fusion protein can be compared to those previously reported on the synthetic M6-31 and M6-35 peptides that were thoroughly studied in the laboratory (Xie et al., 2000). These synthetic peptides were 57% and 68% helical in 100% TFE, respectively. Using the data obtained by the mathematical subtraction of the TrpΔLE spectrum from the TrpΔLE-M6 spectrum, by the Wu et al., (1981) method, the percentage of α-helical structure of the biosynthetic M6-33 was determined to be 38% helical. The lower number can be attributed, in part, to the influence of TrpΔLE on the secondary structure formation of M6. Part of the difference in helical content between the two synthetic peptides can be attributed to the additional four residues in M6-35, which appear to increase the helical tendency of the entire sequence. Some differences may also arise from the presence of a sulfur-containing side chain of Met250 in M6-35 and its absence (Nle250) in M6-31 (Krittanai, et al., 1997). Although the quantitative analysis of the biosynthetic M6-33 covalently joined to the TrpΔLE leader peptide gave values for the helicity of the M6 domain that differed by approximately 20% from those of synthetic model peptides with

similar sequences, the subtraction method suggested that this domain is approximately 40% helical in TFE when bound to the leader sequence. Furthermore, using the same calculation method employed by Xie et al. (2000), the M5I3M6 peptide in the fusion protein was found to be 42% helical.

In M6, 24 residues are expected to be located within the membrane according to the hydropathy analysis, whereas in M5I3M6, 48 residues are expected to be in the membrane. These predictions suggest that the single transmembrane domain peptide sequence should be approximately 75% helical, whereas the double domain peptide should be 55% helical. Thus, the quantitative analysis of the CD results gives better agreement for the double domain peptide. It is possible that the presence of two transmembrane helical domains help to induce and stabilize the formation of each other in a synergistic fashion.

The approach of directly studying membrane peptides fused to leader peptides that increase solubility of the target protein is at least partially supported by the M6 results and comparison of the results of the fused peptide to those published on similar synthetic peptides. Given the similar findings with the M5I3M6 fusion polypeptide, study of this protein can lead to biophysical information relevant to the two transmembrane helices and the connecting loop of Ste2p. As mentioned before, as long as the leader peptide does not significantly interact with the target peptide the subtraction of the CD spectra is considered to be valid. Similar approaches have been considered with different leader peptides such as glutathione S-transferase (Cetin et al.,

1998) and the B1 domain of protein G (Zhou et al., 2001).

Calculation of the secondary structure content was also carried out using the backpropagation neural network method. The difference between this method and the classical statistical methods is that it uses a non-linear analysis like a neural network algorithm (Böhm et al., 1992) to calculate the secondary structure content of a protein. Table 14 summarizes the quantitative analysis of each polypeptide in different concentrations of TFE in water.

Table 14

Helix Content of Membrane Peptides at Different Percentages of TFE in Water Calculated Using the Backpropagation Neural Network Method

TFE%	TrpΔLE-M5I3M6	TrpΔLE	M5I3M6*	TrpΔLE-M6	M6*
100	91	64	65	93	59
70	86	53	60	88	55
50	83	52	56	87	53
30	74	45	48	83	51
10	46	27	35	43	29

* The percentages of M5I3M6 and M6 are based on the spectra obtained mathematically. The subtraction method is described in the Materials and Methods section.

This method calculates the secondary structure content based on different regions of the spectrum unlike the aforementioned model peptide methods that calculate the content of helical structure based on the absorption value of two points of the

spectrum.

Figure 43 depicts the CD results for TrpΔLE and TrpΔLE-M5I3M6 in phosphate buffer containing 0.5% SDS.

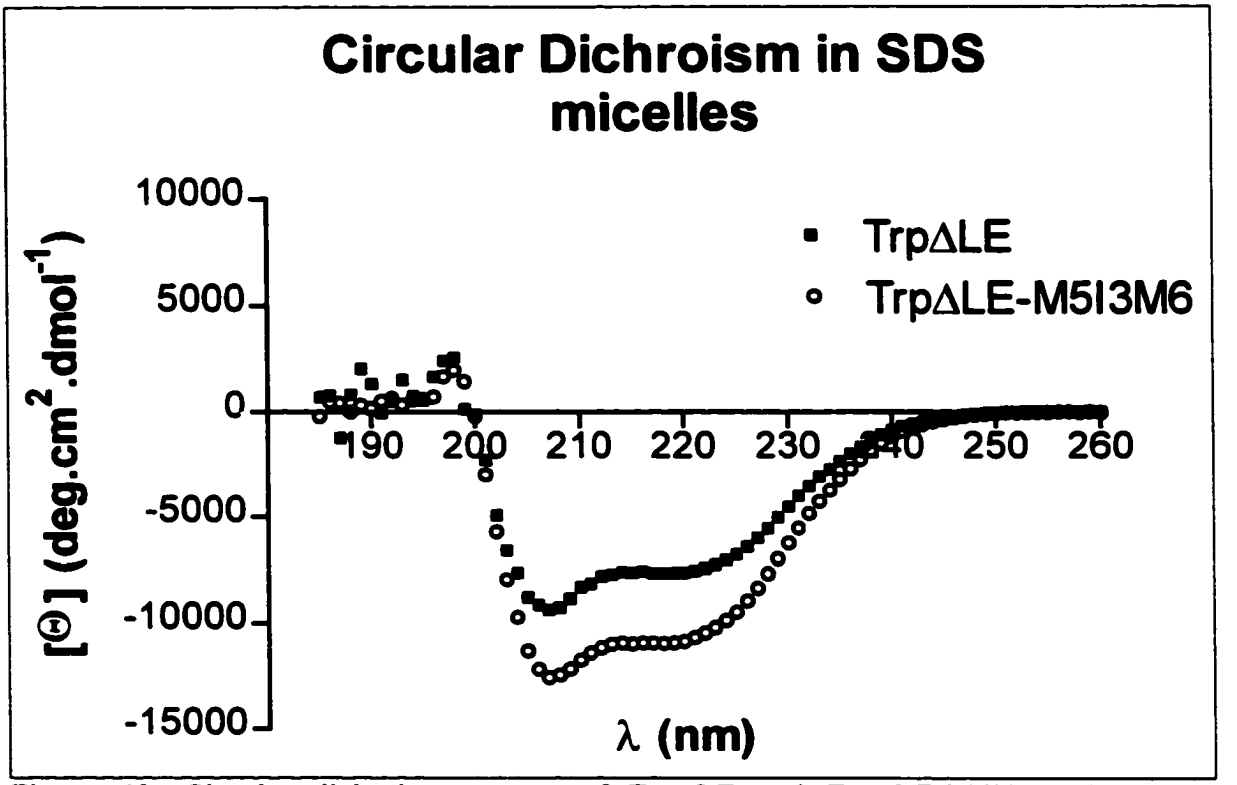

Figure 43: Circular dichroism spectra of TrpΔLE and TrpΔLE-M5I3M6 in SDS micelles. Samples were dissolved in PBS buffer containing 0.5% SDS.

The CD profiles of both proteins match those expected for a helical protein. The mean residue value of the TrpΔLE-M5I3M6 is indicative of a higher helicity content than the leader protein, indicating that the Ste2p domains are helical under the conditions of the experiment. The percentage of α-helical structure was calculated using data at 222 nm and equations based on pure secondary structures from model peptides

and the backpropagation neural network procedure (Table 15). The results yielded 44% and 33% helicity for TrpΔLE-M5I3M6 and TrpΔLE, respectively, from the statistical approach and 34% and 26% from the neural network calculation. These values are based on a concentration found by the weight of the samples. Therefore, a significant amount of error can be expected in the calculations of the helical content of each protein.

Table 15

Percentage of Helical Content of TrpΔLE-M5I3M6 in SDS Micelles

	TrpΔLE-M5I3M6	TrpΔLE
Model Peptides Method [a]	44	33
Neural Network Method [b]	34	26

a. Greenfield et al., 1969; Chen et al., 1974; Wu et al., 1981
b. Böhm et al., 1992

In general, the agreement between the two calculation methods is reasonable. The lower helicity values obtained in SDS micelles compared to those in 100% TFE may arise from the imprecise concentration determined for the samples in detergent.

Table 16 summarizes the helical contents of various polypeptides in 100% TFE calculated with the different methods. The values vary by as much as 50% for a given polypeptide. The neural network estimations are usually closer to values calculated using model polypeptides than those using model membrane proteins. Therefore, the use of the model membrane proteins probably underestimates the helicities of the transmembrane domains of Ste2p.

Table 16

Comparison of the Different Methods Used to Calculate the Helical Content of Membrane Peptides [a]

	TrpΔLE-M5I3M6	TrpΔLE	M5I3M6 [b]	TrpΔLE-M6	M6 [b]
222 nm [c]	87	48	54	87	45
208 nm [c]	116	63	70	117	57
222 nm [d]	53	29	33	53	28
208 nm [d]	73	40	44	74	36
Neural network [e]	91	64	65	93	59

a. Data were measured in 100% TFE.
b. The percentages of M5I3M6 and M6 are based on the spectra obtained mathematically. The subtraction method is described in the Materials and Methods section.
c. Model peptide method using fixed infinite helical content (peptides Greenfield et al., 1969; Chen et al., 1974; Wu et al., 1981)
d. Model membrane protein method (Park et al., 1992)
e. Neural network method (Böhm et al., 1992)

In summary, the CD analyses provide qualitative evidence that the single (M6) and double (M5I3M6) membrane domains of Ste2p retain their helical structures in TFE and TFE/water environments when bound to the TrpΔLE leader peptide. The quantitative analysis, while not totally satisfactory, provides insights into the number of residues in these peptides that are helical. It appears that TrpΔLE may influence the helicity of M6 because unstructured residues in the leader are directly connected to a relatively short transmembrane domain peptide and disturb its secondary structure.

Apparently this influence is less pronounced with the double transmembrane domain which manifests a helicity closer to that expected on the basis of hydropathy predictions.

To determine the stability of the helical structures formed by the domains a temperature profile of the CD spectra was obtained for TrpΔLE-M6 and TrpΔLE-M5I3M6 in TFE/water (30/70) (Figure 44).

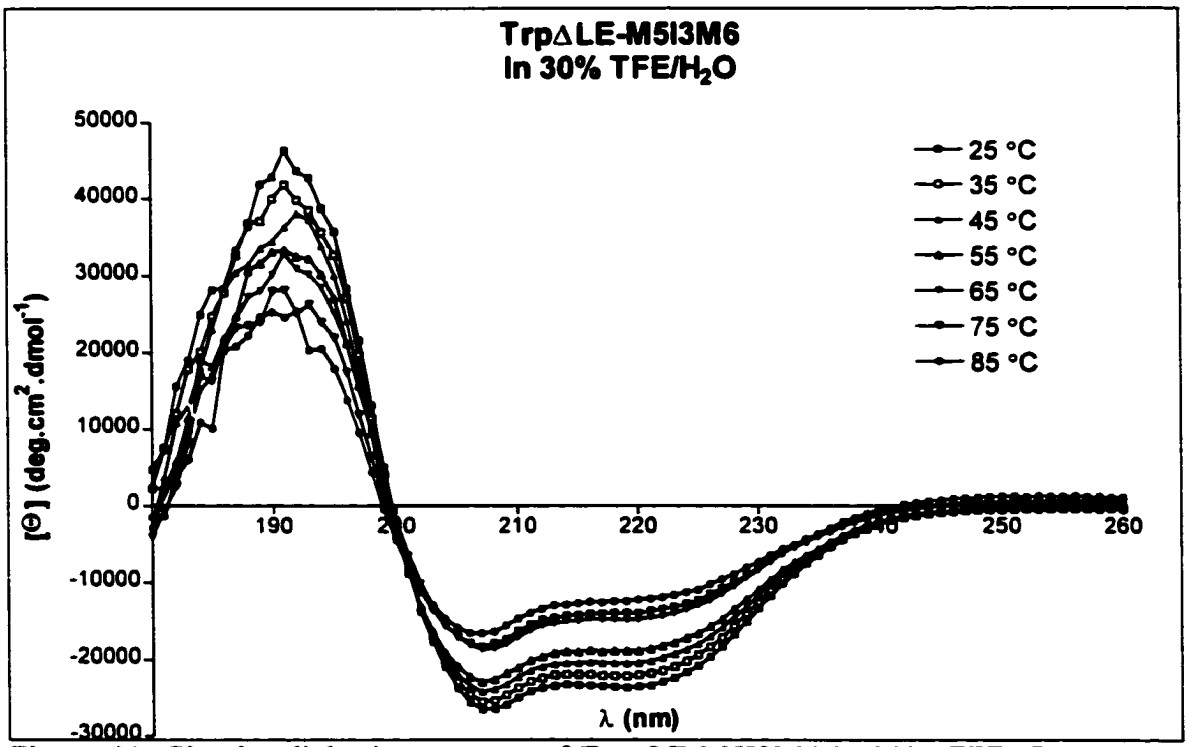

Figure 44: Circular dichroism spectra of TrpΔLE-M5I3M6 in 30% TFE/ 70% H_2O at different temperatures.

All profiles are qualitatively α-helical in shape with the proportion of helix diminishing as the temperature is raised from 25 to 85 °C. A region around 200 nm,

where all the spectra converge, resembles an isodichroic point but at this point there is no evidence for the co-existence of distinct interchangeable secondary structures. A ratio $[\theta_{190}]:[\theta_{220}]$ of approximately 2:1 was also observed in experiments with different concentrations of TFE. Figure 45 depicts the CD spectra of the TrpΔLE leader peptide and also indicates that the helical regions of this polypeptide decrease with increasing temperature.

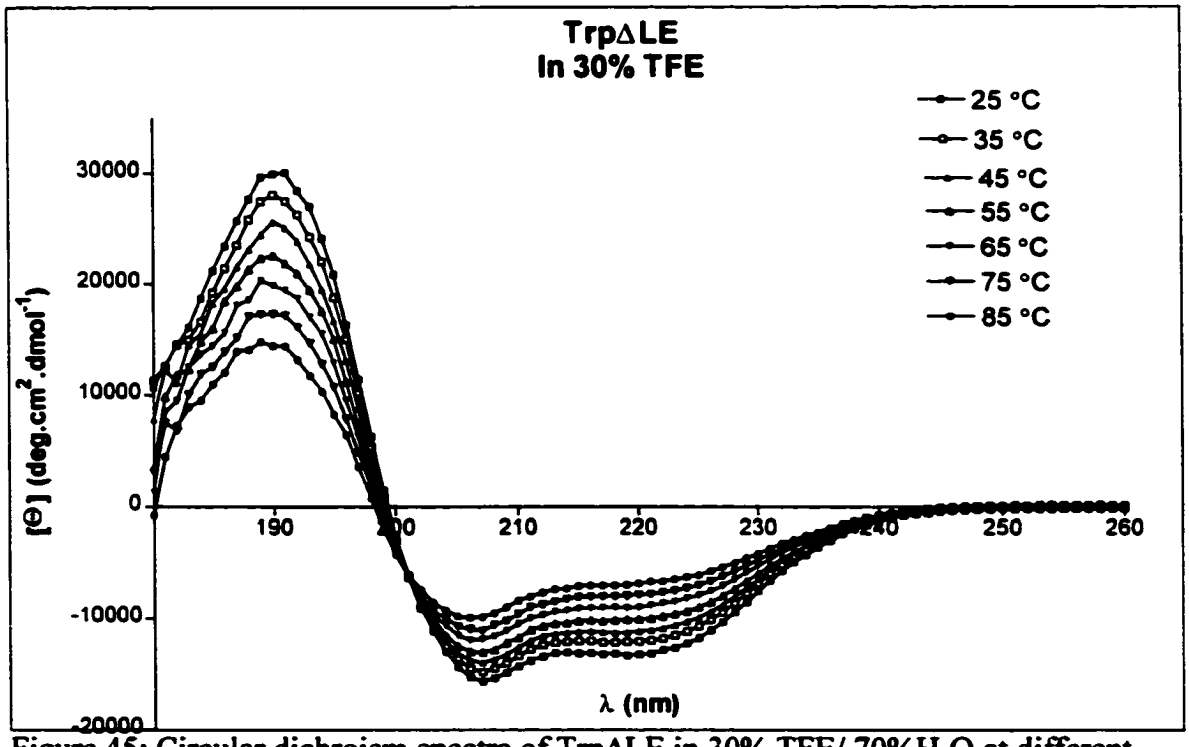

Figure 45: Circular dichroism spectra of TrpΔLE in 30% TFE/ 70%H_2O at different temperatures.

The determination of the CD profile for M5I3M6 obtained by subtraction is represented in Figure 46, and indicates that M5I3M6 is helical in the presence of 30% TFE and does not unfold in a two-state manner when the temperature is increased.

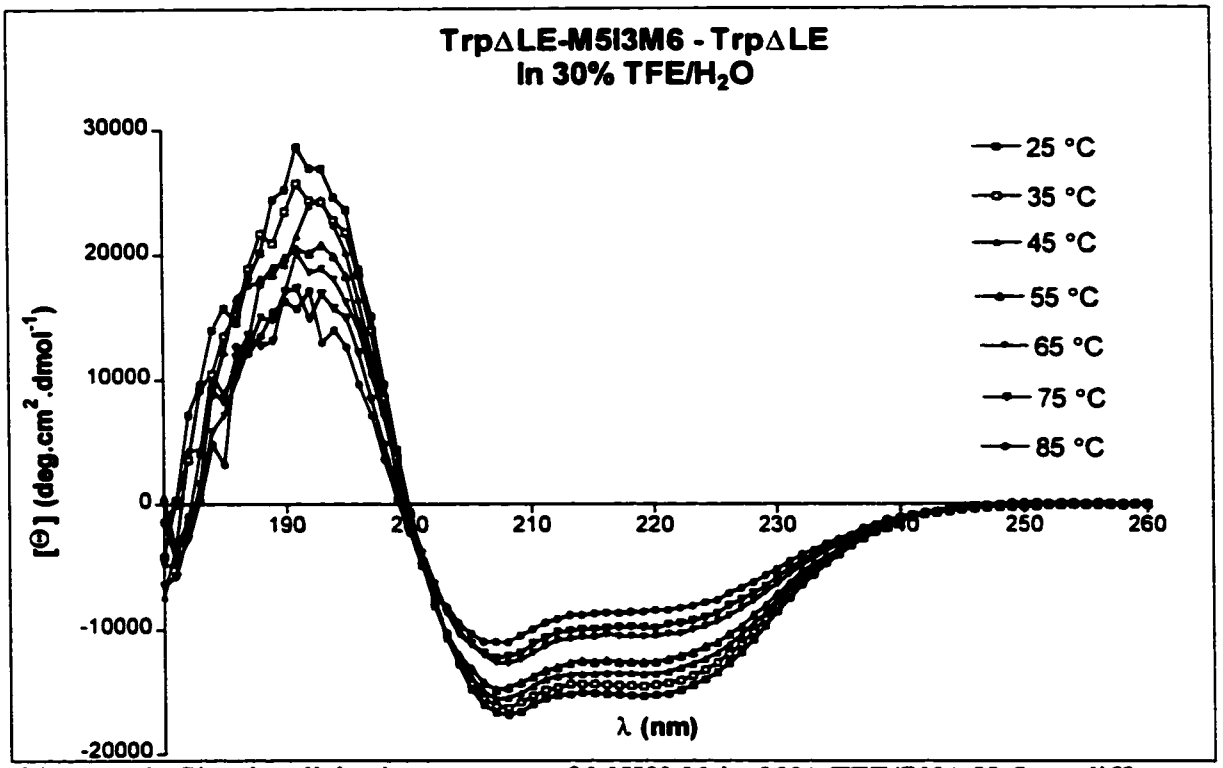

Figure 46: Circular dichroism spectra of M5I3M6 in 30% TFE/70% H_2O at different temperatures. The spectra were obtained after subtraction of TrpΔLE spectra from that of TrpΔLE-M5I3M6 as indicated in the Materials and Methods section.

It can be observed that even at 85 °C there is a measurable helical content in this polypeptide. When the same analysis was applied to TrpΔLE-M6 (Figure 47), it was seen that this fusion protein also contains a significant amount of helical structure that decreases with an increase in temperature.

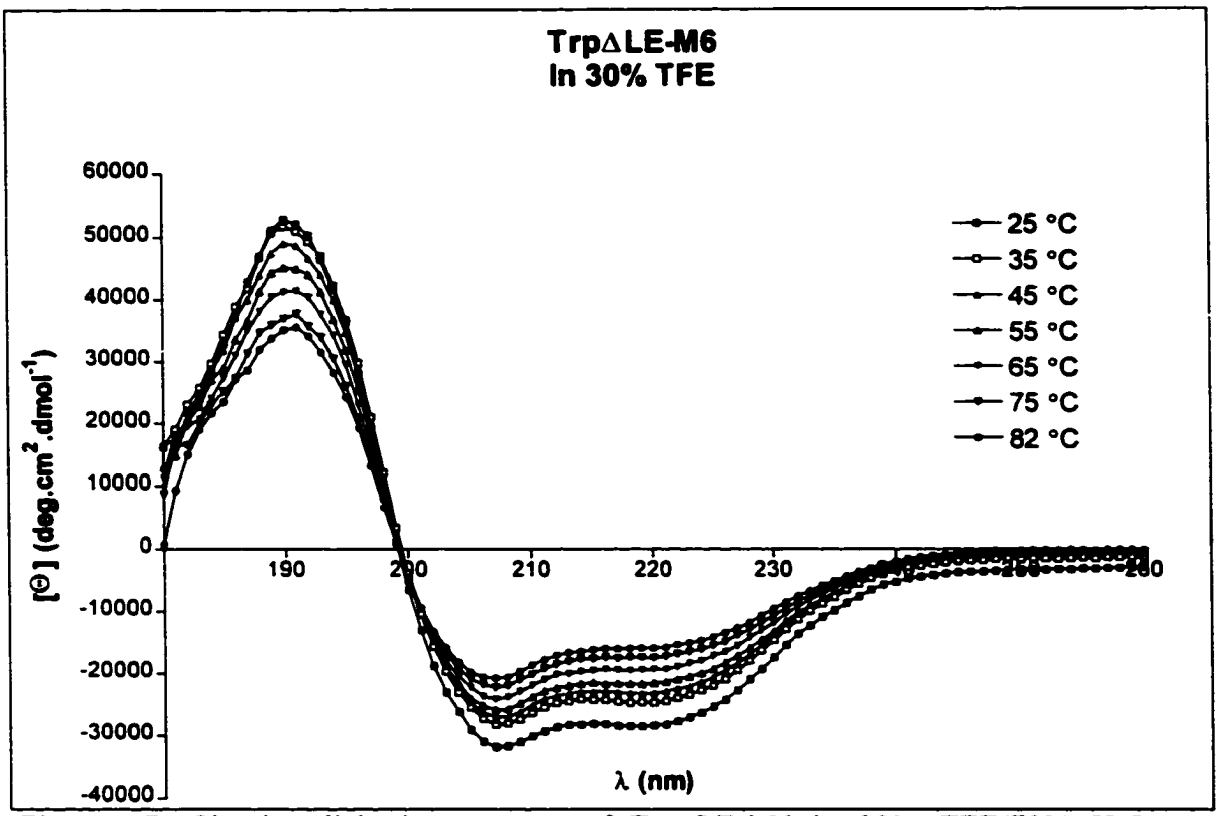

Figure 47: Circular dichroism spectra of TrpΔLE-M6 in 30% TFE/70% H_2O at different temperatures.

Plots of $[\Theta]_{222}$ versus temperature for TrpΔLE, TrpΔLE-M6, and TrpΔLE-M5I3M6 polypeptides dissolved in 30% TFE are shown in Figure 48.

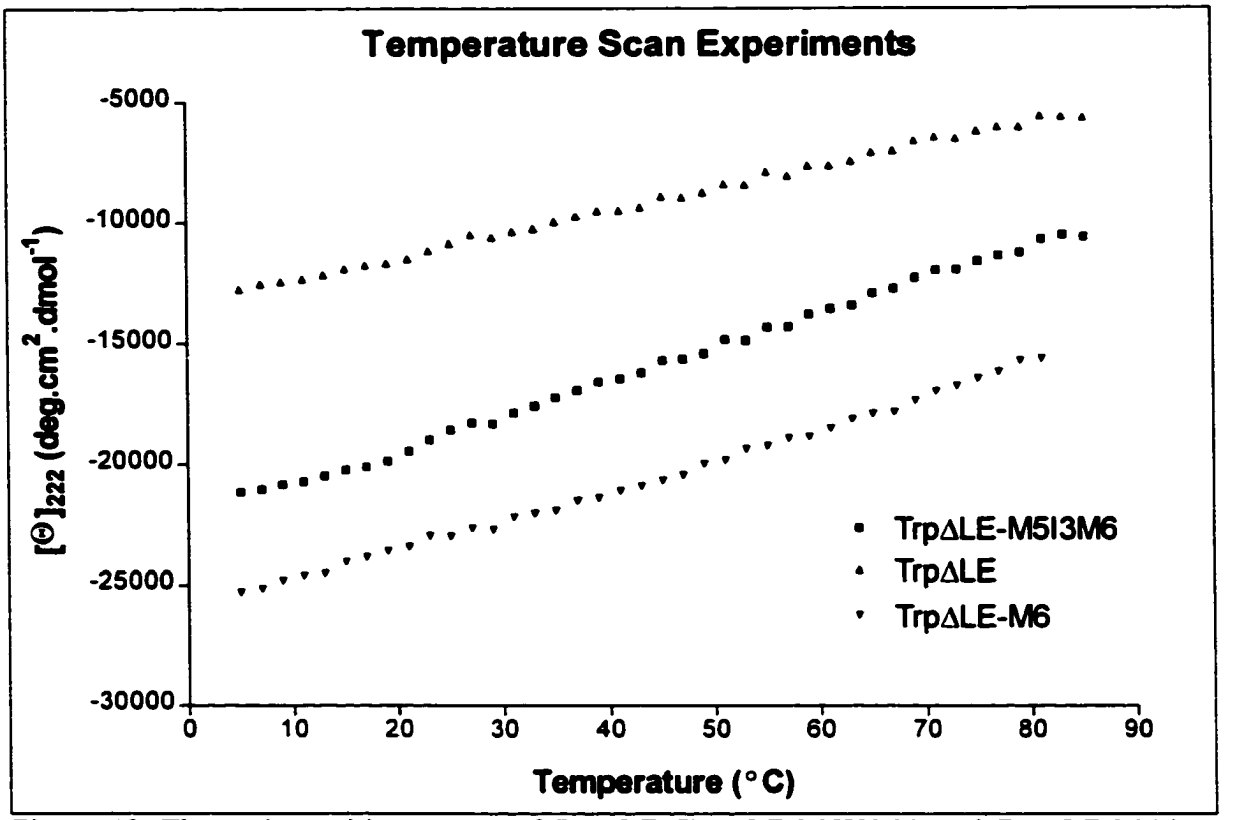

Figure 48: Thermal transition curves of TrpΔLE, TrpΔLE-M5I3M6, and TrpΔLE-M6 in 30% TFE/70% H_2O. The mean residue ellipticity of each protein at 222 nm was plotted against temperature.

The linear trend of the ellipticity at 222 nm at different temperatures indicates that there is no abrupt conformational transition in these molecules over the temperature range examined. It is known that tertiary interactions can contribute significantly to the CD spectra of proteins (Kallenbach et al., 1996; Luo et al., 1997). Given the absence of any cooperative transition in the CD spectra with temperature, it is concluded that the various domains of the fusion proteins are not interacting.

Circular dichroism experiments were also conducted in DMPC and

DMPC:DMPG (1:1) vesicles. Preliminary studies indicated that the polypeptides in the mixed lipid vesicles gave very low ellipticities – presumably because of poor insertion into the bilayers. Therefore, detailed CD studies were only carried out in DMPC vesicles. The CD spectra of the samples in DMPC vesicles are depicted in Figure 49.

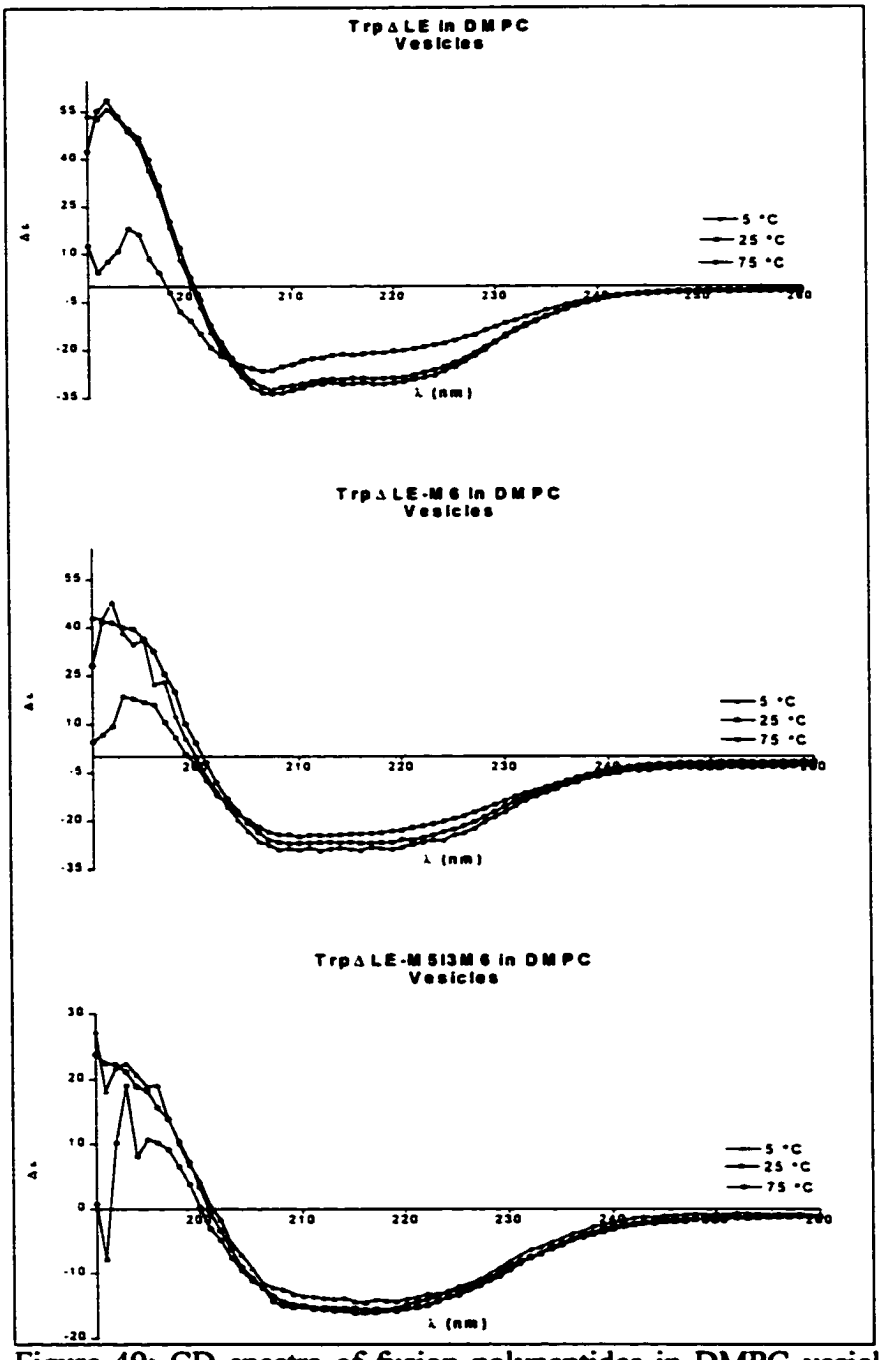

Figure 49: CD spectra of fusion polypeptides in DMPC vesicles. Spectra of TrpΔLE (top panel), TrpΔLE-M6 (middle panel), and TrpΔLE-M5I3M6 (bottom panel) at 5 °C (▲), 25 °C (O) and 75 °C (■).

Although the 208 nm and 222 nm minima were less pronounced than in the TFE spectra, the profiles indicated that all samples are mainly helical. Clearly the CD spectra of the different polypeptides in DMPC do not show patterns for a purely β-structure or random coil. The conclusion that the fusion peptides are helical in vesicles was confirmed by deconvolution performed with the neural network algorithm. The samples could not accurately be converted into mean residue ellipticity since the amount of peptide incorporated into the lipid vesicles could not be quantified. Determination of the protein concentration by amino acid analysis was not possible because of interference of the lipids during the hydrolysis and during the HPLC detection of the hydrolyzed amino acids. Spectroscopic determination of the protein concentration in lipids was also hampered by the presence of the phospholipids and the results did not provide a reliable concentration of the protein in the protein-lipid mixture.

It is interesting to note that the lipid vesicle samples containing the fusion proteins show very similar spectra at the three temperatures. But the TrpΔLE-containing vesicles show a distinct CD pattern at 75 °C. Visual inspection of the protein-lipid vesicles revealed a clear solution for the TrpΔLE-M6 and TrpΔLE-M5I3M6 lipid complex, but a less clear solution for the TrpΔLE vesicles. Vesicles containing no protein were unstable to the thermal denaturation experiments. The change in ellipticity at 222 nm upon increasing the temperature from 5 °C to 75 °C and then the reverse measurements from 75 °C to 5 °C were carried out for all peptides. The thermal transition experiments are summarized in Figure 50.

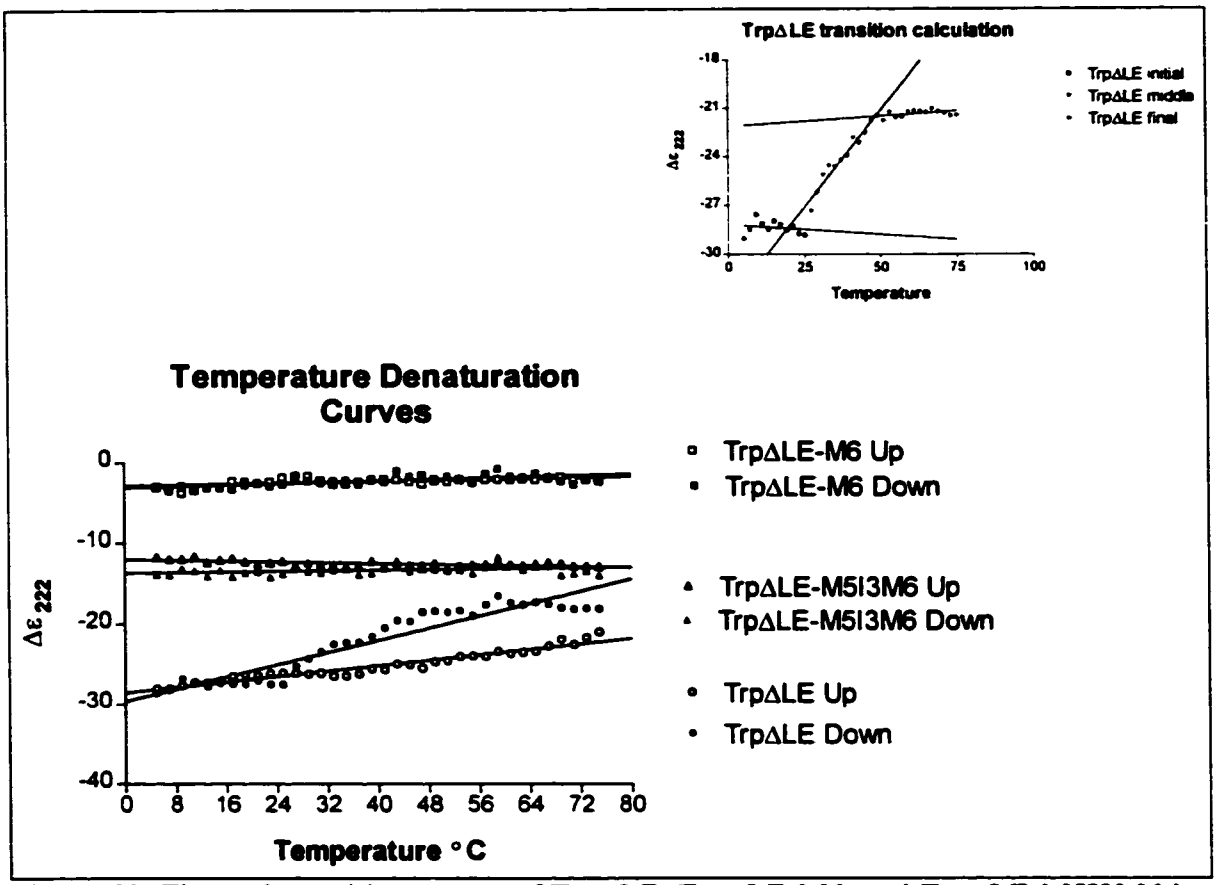

Figure 50: Thermal transition curves of TrpΔLE, TrpΔLE-M6, and TrpΔLE-M5I3M6 in DMPC vesicles. The increasing temperature curves are denoted by empty symbols and the decreasing temperature curves by filled symbols. The insert depicts the curves of the temperature descending profile of TrpΔLE showing the linear regression analysis in three different regions. The data for TrpΔLE-M6 were artificially offset for the purpose of presentation.

The TrpΔLE-M6 and TrpΔLE-M5I3M6 protein-lipid complexes show almost no change in $\Delta\varepsilon_{220}$. In contrast, TrpΔLE in lipid vesicles exhibited significant changes in $\Delta\varepsilon_{220}$ with increasing temperature. The fusion protein containing the sixth transmembrane showed a reversible thermal change in CD; the same rotational strength was observed at the beginning and at the end of the thermal denaturation cycle.

TrpΔLE-M5I3M6 in the presence of DMPC vesicles also exhibited nearly reversible and linear denaturation curves. This sample showed more negative ellipticity values at the end of the cycle, which would indicate that the heating of the protein-lipid vesicles induced an increase in the helical content of the protein. The TrpΔLE leader protein, on the other hand, showed different temperature profiles for the heating and cooling curves. The increase in temperature showed a quasi-linear thermal transition curve. However, the decrease in temperature followed a S-shaped curve where the final ellipticity value was similar to the value observed at the beginning of the denaturation experiments. Apparently the TrpΔLE leader protein assumes different structures at 75 °C and 5°C, and the unfolding and refolding pathways are distinct. The protein readopts its original conformation in the range of 22-30 °C.

The difference in the behavior of the TrpΔLE and the two fusion proteins is an indication that the helical structure observed in the leader polypeptide is adopted externally and independently from the structure of the membrane peptides. The higher stability of the vesicles containing the fusion polypeptides and the small change in the dichroic signal with temperature leads to the conclusion that the membrane domains were indeed inserted into the lipid vesicles and adopted a helical structure, as suggested by the CD patterns in Figure 49.

Biophysical studies of the fusion peptides: Fourier transform infrared spectroscopy

IR studies can help to evaluate the changes in conformation upon interaction of different domains of the receptor in lipid bilayers. Characteristic absorption bands of the amide group at 3300 cm^{-1} [υ (NH)], 1635 cm^{-1} [υ (C=O), amide I], 1540 cm^{-1} [δ (NH) + υ (C-N), amide II], 690 cm^{-1} (amide V), and 580 cm^{-1} (amide VI) can be observed in a typical IR spectrum (Siesler et al., 1980). The amide I band is sensitive to secondary structure because it adopts characteristic frequencies for α helices, β strands, β turns, and τ turns. The amide II mode is a more complex vibration because it includes NH in-plane vibration, C-N, C^{α}-C stretching, and CO in-plane bending vibrations. However, useful information can be obtained when a spectrum of the protein can be measured in H_2O and another in ^2H_2O. Deuteration of the nitrogen lowers its frequency of vibration by about 100 cm^{-1} (Siebert et al., 1995). By measuring the amide II intensity of the fusion protein after exposure to ^2H_2O, the extent of $^1H/^2H$ exchange can be calculated, and the inference of the region that is exposed to the solvent can be made.

Figure 51 shows the FT-IR results of the TrpΔLE-M6 and TrpΔLE alone in DMPC vesicles.

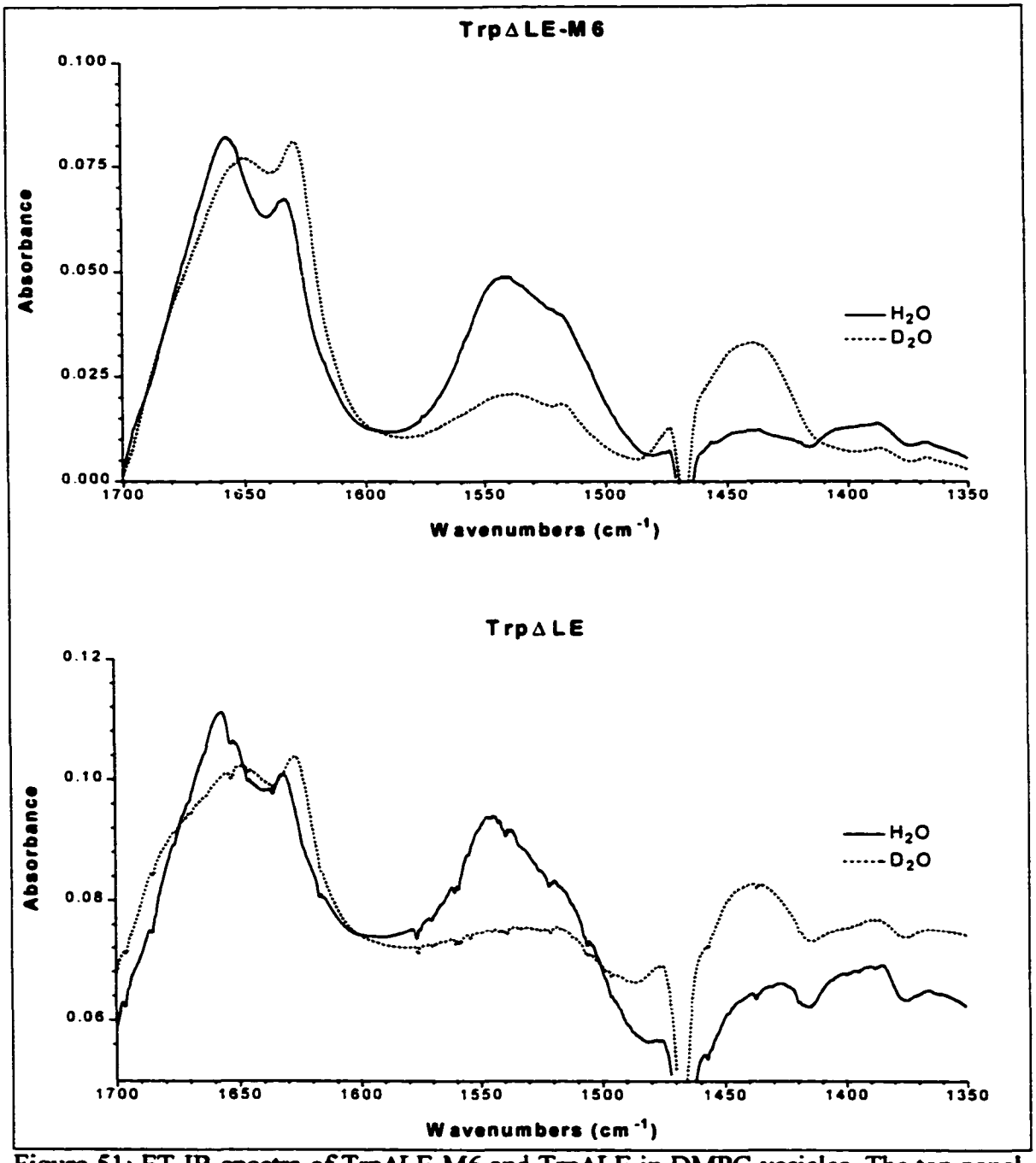

Figure 51: FT-IR spectra of TrpΔLE-M6 and TrpΔLE in DMPC vesicles. The top panel represents the results of the TrpΔLE-M6 and the bottom panel TrpΔLE. Solid lines denote the spectra of the protonated samples and the dashed lines the IR of the deuterated forms of the proteins.

The samples were exposed to H_2O and D_2O in order to determine the qualitative and quantitative effect of the exchange in the amides signal. It was expected that if the TrpΔLE is located on the surface of the lipid bilayer, H/D exchange in a period of 12 h could be observed by measuring the intensity of the amide II signal. In the fusion peptides, it was expected that higher protection would be observed for protons involved in hydrogen bonds or of the residues located inside the lipid bilayer. As illustrated in Figure 51, the amide II shows a broad absorption band that significantly shifts from 1540 cm^{-1} to 1440 cm^{-1} upon exposure to D_2O. A similar situation is observed in TrpΔLE. Problems involving the baseline made it difficult to accurately quantitate the FT-IR results. Therefore, from these experiments it was not possible to reproducibly determine the extent of deuteration in the leader peptide and the fusion proteins. Thus, at this time the location of TrpΔLE with respect to the lipid vesicles cannot be determined by IR spectroscopy.

The amide I band shows an intense absorption band at 1658 cm^{-1}, characteristic of α-helical membrane proteins (Parker, 1973; Haris et al., 1995; Ding et al., 2001b). The absorption band observed in all of the proteins around 1630 cm^{-1} could be attributed to the presence of a β-structure; however, similar bands were observed in proteins like myoglobin with very little or almost no β-structure (Haris et al., 1995). A band around 1640 cm^{-1} has also been assigned to a 3_{10}-helix in the antibiotic alamethicin (Haris et al., 1995). The shift of the amide I to lower frequencies and its reduction of the intensity observed in the protein:lipid mixtures upon exposure to ^2H_2O could be attributed to the

polypeptide segments in random coil conformations, as has been observed in the nicotinic acetylcholine receptor (Baenziger et al., 1995). In general, however, no additional insights into the structure of the M6 domain in lipid could be extracted from the IR experiments.

Nuclear magnetic resonance spectroscopy on fusion peptides:

Solution NMR experiments were carried out using 2D experiments including double quantum filtered correlation spectroscopy (DQF-COSY) (Rance et al., 1983), total correlation spectroscopy (TOCSY) (Braunschweiler et al., 1983), and nuclear Overhauser effect spectroscopy (NOESY) (Jeener et al., 1979). The NMR experiments on the fusion peptides and the TrpΔLE were run in 50% TFE. The 2D experiments show that there was significant overlap of the cross peaks and very little dispersion of the amide chemical shifts. This could be as a result of the slow tumbling of the proteins that causes peak broadening. Furthermore, residues located in helical structures do not show large chemical shift dispersions (Cavanagh et al., 1996). Although the CD results indicated that the TrpΔLE and both fusion proteins contain more than 35% and 60% helical structure, respectively, the quality of the NMR data obtained do not allow chemical shifts assignments to be made. These experiments lead to the conclusion that in 50% TFE the membrane peptides fused to the TrpΔLE leader peptide could not be studied by NMR.

Nuclear magnetic resonance spectroscopy on synthetic transmembrane peptides

In order to investigate other conditions for NMR analysis of fusion polypeptides we decided to study two synthetic peptides representing the sixth transmembrane domain in Ste2p (M6-35P and M6-35L) in $CDCl_3$:CD_3OH:H_2O (4:4:1). This solvent was previously used to investigate the structure of the EmrE transporter (Schwaiger et al., 1998) and subunit c of F_1F_0 ATP synthase (Rastogi et al., 1999). The peptides examined harbor the naturally occurring proline (M6-35P) and a substitution of Pro258Leu (M6-35L) that is known to cause the intact receptor to become constitutively activated (Konopka et al., 1996).

The peptides were dissolved in $CDCl_3$:CD_3OH:H_2O (4:4:1) to give about 0.8 mM concentrations, and the classical two-dimensional NMR experiments outlined above were conducted on each peptide. The strategy to assign the corresponding chemical shifts of the protons of each residue in the peptide was as follows: 1) The chemical shift frequencies of the backbone amide proton and the α-proton (C^{α}-H) were determined from their correlation peak in a DQF-COSY spectrum. 2) The spin system of each side chain was identified and correlated with their respective backbone resonance in the TOCSY spectrum. 3) After organizing the spin systems of all possible amino acid types, NOESY was used to connect these spin systems in sequential order. 4) NOESY was also used to verify the backbone assignment of the aromatic residues, by looking for the cross peaks between the β-protons (C^{β}-H) and the aromatic ring protons (Gesell et al., 1997; Wüthrich, 1986). After the sequential assignment, medium-

range NOEs were assigned based on the previously assigned backbone protons and the cross peaks resulting from the correlation through space to protons far removed in the primary sequence but close in space due to the presence of a specific conformation. The medium-range NOEs inspected were: NN (i, i+1), αN (i, i+1), βN (i, i+1), αN (i, i+3), αN (i, i+4), and αβ (i, i+3).

Tables 17 and 18 list the proton chemical shift assignments of the M6-35P and M6-35L synthetic peptides.

Table 17.

^{1}H Resonance Assignments* of M6-35P in CD_3Cl:CD_3OH:H_2O (4:4:1) at 25 °C

Residue	NH	αCH	β_1CH	β_2CH	γCH	δCH	εCH	Other
Lys								
Lys	8.75	4.28	1.69		1.39	1.44		
Lys	8.40	4.16	1.62	1.62	1.34	1.27		
Phe241	8.13		2.97	3.11				
Asp242	8.17	4.66	2.79	2.86				
Ser243	8.42	4.06	3.85	3.85				
Phe244	8.03	4.23	3.05	3.17				
His245	8.16	4.08	3.26	3.26				
Ile246	7.83	3.65	1.96		0.92			
Leu247	7.94	4.02	1.66	1.75				
Leu248	7.93	3.92						
Ile249	7.78	3.67	1.94		1.76			1.93, 0.91
Met250	8.56	4.15	2.14		2.72			
Ser251	8.39	4.17	3.89	4.10				
Ala252	8.18	4.08	1.55					
Gln253	8.30	3.99	2.29	2.55				
Ser254	7.96	4.19	3.99	4.07				
Leu255	7.48	3.59	1.58	1.84		0.88		
Leu256	7.85	4.16	1.84	1.69		0.95		
Val257	8.07	3.77	2.34		0.96			1.14
Pro258		4.19	1.92	2.28		3.61		3.71
Ser259	7.40	4.26	3.87	4.17				
Ile260	8.01	3.71	2.05		0.94			
Ile261	8.01	3.59	1.96		0.92			
Phe262	8.22	4.12	3.25	3.29				
Ile263	8.31	3.66	2.04		0.98			
Leu264	8.60	4.01	1.92			0.88		
Ala265	8.69	3.93	1.44					
Tyr266	8.62	3.88	2.71	2.89		6.92	6.70	
Ser267	8.24	4.10	3.91	3.97				
Leu268	7.92	4.05	1.52	1.78	1.65	1.29		0.84
Lys269	7.73	3.99	1.79	1.79	1.40	1.57		1.41
Lys	7.76	4.31	1.93		1.77	1.46		
Lys	7.69	4.42	3.87	3.93				
Ser	8.75	4.28	1.69		1.39	1.44		

* Chemical shifts are given in parts per million (ppm) relative to the H_2O proton resonance.

Table 18.

^{1}H Resonance Assignments* of M6-35L in CD_3Cl:CD_3OH:H_2O (4:4:1) at 25 °C

Residue	NH	αCH	β_1CH	β_2CH	γCH	δCH	εCH	Other
Lys								
Lys	8.74	4.29	1.75	1.68	1.45	1.39	2.93	
Lys	8.40	4.17	1.62	1.62	1.35	1.26	2.90	
Phe241	8.14		3.11	2.97				
Asp242	8.17	4.67	2.86	2.79				
Ser243	8.43	4.07	3.86	3.86				
Phe244	8.04	4.24	3.17	3.06		7.03		
His245	8.17	4.09	3.26	3.26		7.14		
Ile246	7.84	3.66	1.97		0.81			1.16, 0.93
Leu247	7.96	3.94	1.59			0.77		
Leu248	4.09	4.09				0.95		
Ile249	7.81	3.69	1.95		1.71			0.91
Met250	8.61	4.17	2.15		2.53			2.73
Ser251	8.44	4.20	4.15	4.15				
Ala252	8.24	4.08	1.57					
Gln253	8.47	3.98	2.32		2.54			
Ser254	8.09	4.16	3.91					
Leu255	7.95	4.08	1.80		1.52	0.84		
Leu256	8.02	4.09	1.95			0.95		
Val257	8.08	3.54	2.19		0.98			1.11
Leu258	8.26	3.98	1.80		1.61	0.88		
Ser259	8.10	4.22						
Ile260	7.95	3.70	2.09		0.94			
Ile261	8.18	3.60	2.06		0.92			
Phe262	8.48	4.18	3.26	3.27		7.19		
Ile263	8.42	3.70	2.06		0.99			
Leu264	8.71	4.02	1.95	1.87	1.58	0.88		
Ala265	8.76	3.95	1.46					
Tyr266	8.63	3.92	2.95	2.81		6.95	6.70	
Ser267	8.27	3.98	4.11	3.92				
Leu268	7.96	4.02			1.67			
Lys269	7.75	4.00	1.80		1.58	1.42	2.87	
Lys	7.92	4.05	1.73		1.66	1.55	2.86	
Lys	7.76	4.32	1.94	1.78	1.62	1.47	2.88	
Ser	7.70	4.43	3.94	3.88				

* Chemical shifts are given in parts per million (ppm) relative to the H_2O proton resonance.

Figure 52 shows the sequential NN (i, i+1) assignments of the crosspeaks observed in the amide region of the NOESY spectrum of M6-35P.

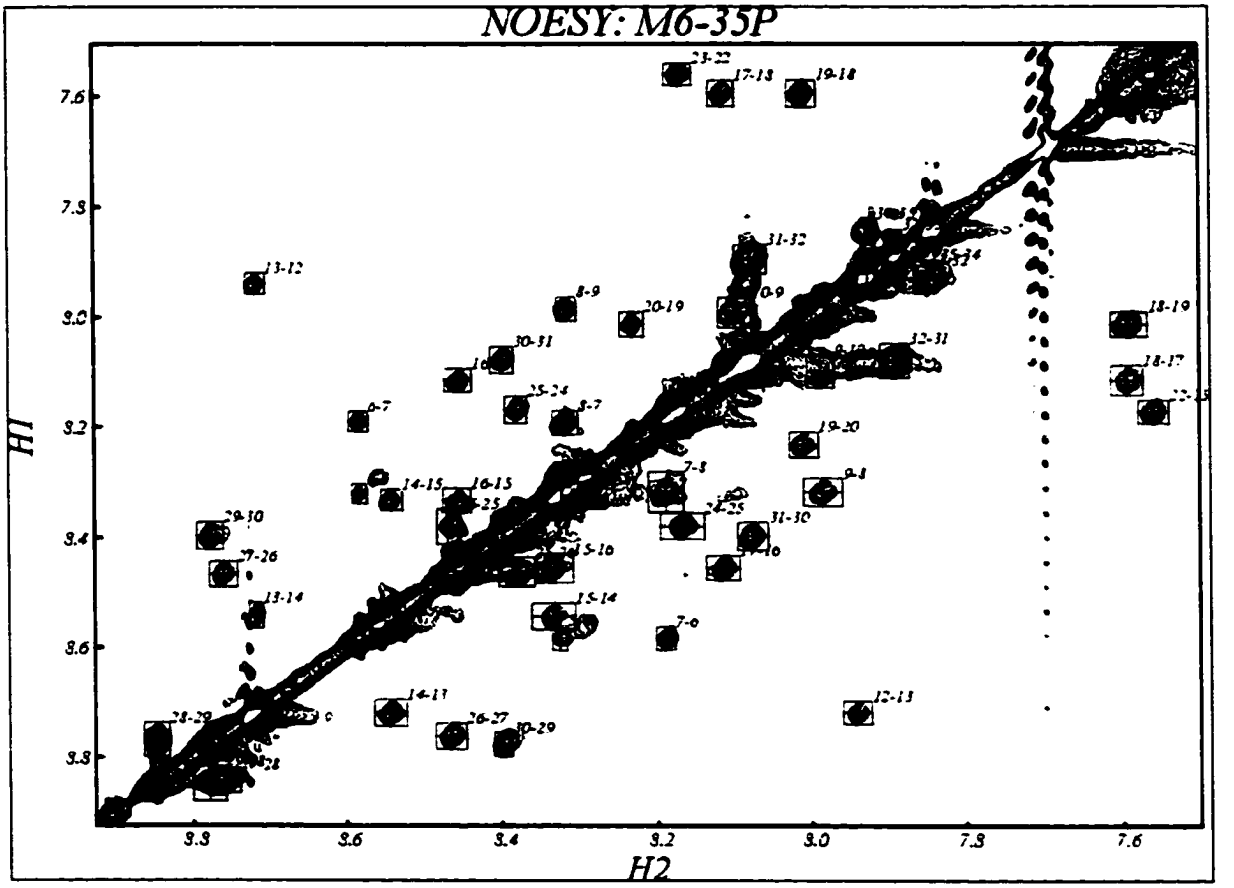

Figure 52: 2D-NOESY spectrum of M6-35P in $CDCl_3$:CD_3OH:H_2O (4:4:1). The spectrum depicts the sequential NN (i, i+1) assignments in the amide region. The sequence of the peptide is:
$^{1}K^{2}K^{3}F^{4}D^{5}S^{6}F^{7}H^{8}I^{9}L^{10}L^{11}I^{12}M^{13}S^{14}A^{15}Q^{16}S^{17}L^{18}L^{19}V^{20}P^{21}S^{22}I^{23}I^{24}F^{25}I^{26}L^{27}A^{28}Y^{29}S^{30}L^{31}K^{32}K^{33}K^{34}S^{35}$

Similar assignments were made in the M6-35L peptide. The M6-35L peptide was more difficult to assign in this region because the mutation Pro258Leu introduced one more leucine into the overlapped spectrum of leucines present in this membranous

peptide.

Determination of the $^3J_{HN\alpha}$ (J coupling constant that relates to the ϕ angle) was carried out by measuring the distance between the positive and negative phase lines of the HN_α cross peak in the DQF-COSY experiment. The measurement was done omitting residues where the apparent splitting was distorted due to antiphase cancellation. Wishart's method (Gesell et al., 1997; Wishart et al., 1992) was used to rapidly identify secondary structure based on the chemical shifts differences of the α protons and the random coil values. Figures 53 and 54 show the chemical shift difference of M6-35P and M6-35L, respectively.

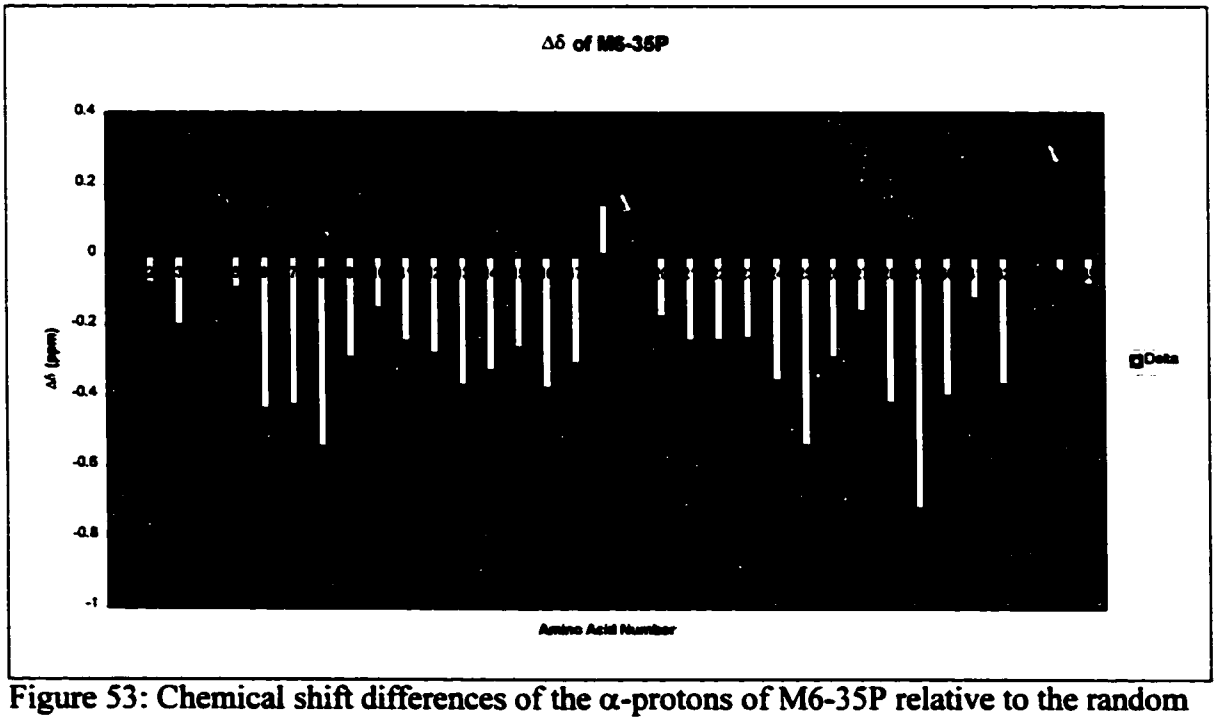

Figure 53: Chemical shift differences of the α-protons of M6-35P relative to the random coil values. The measurements were made in $CDCl_3$:CD_3OH:H_2O (4:4:1) and the random coil values are those reported in water. The number in the X-axis represents the position of the amino acid in the peptide sequence. The chemical shift values of residues 1, 4, and 33 could not be determined during the assignments.

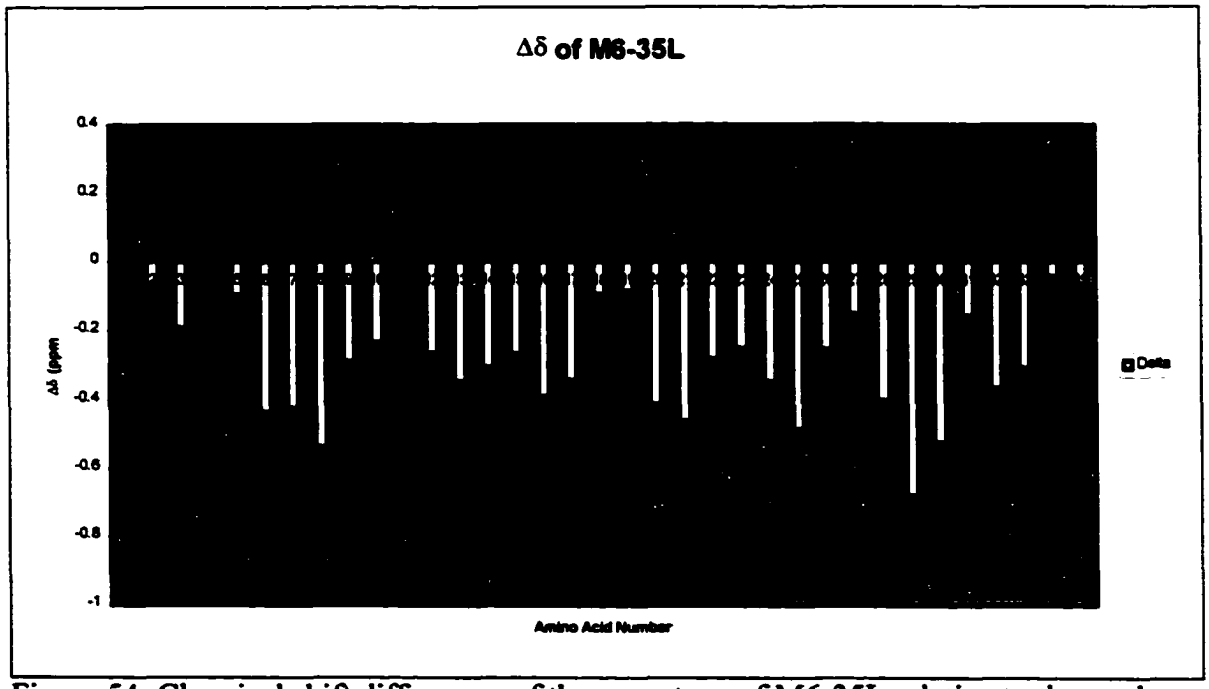

Figure 54: Chemical shift differences of the α-protons of M6-35L relative to the random coil values. The measurements were made in $CDCl_3$:CD_3OH:H_2O (4:4:1) and the random coil values are those reported in water. The number in the X-axis represents the position of the amino acid in the peptide sequence. The chemical shift values of residues 1, 4, and 11 could not be determined during the assignments.

These chemical shift differences that can be converted to indices predict that both peptides are mostly helical in the solvent system used. This agrees with the fact that sequential NN (i, i+1) cross peaks found were consistently stronger than αN (i, i+1) cross peaks, an observation which also correlates to a helical structure (Clore et al., 1989). It is important to note that the reference random coil chemical shifts taken for the

calculations correspond to the values given by Wishart and others (1992) for residues in water. To date no random coil reference values have been reported in $CDCl_3$:CD_3OH:H_2O solvent mixtures.

The M6-35L peptide showed consistently strong and regular long-range connectivities αN (i, i+3) at the C-terminus of the peptide (Figure 55).

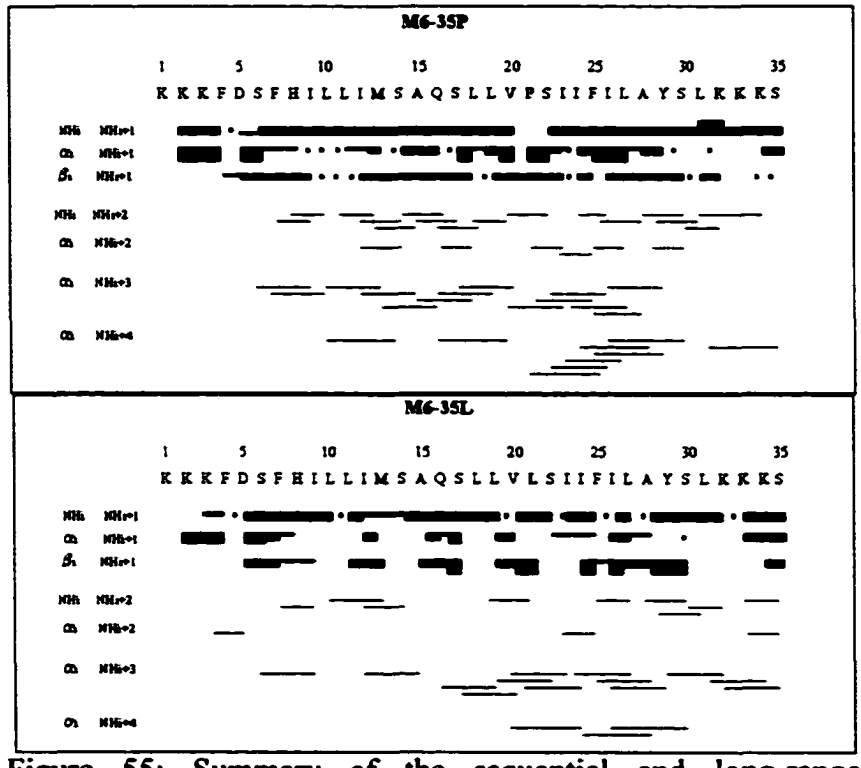

Figure 55: Summary of the sequential and long-range connectivities of M6-35P (top panel) and M6-35L (bottom panel). The thickness of the bars in the (i, i+1) connectivities represents the intensity of the observed crosspeaks. The asterisks represent the NOE connectivities obscured by peak overlap.

Fewer long-range connectivities were observed at the amine side of this peptide. In contrast the M6-35P peptide showed regular long-range connectivities throughout the

peptide (excluding residues at the flexible ends). Significantly, there are more long-range connectivities around residue 21 in M6-35L than there are in M6-35P. This can be attributed to the flexibility around the proline residue. Both peptides show some overlapping peaks around 8.1 ppm. The overlapping was caused by the degeneracy in the leucine amide NH resonances 10, 11, and 31 (according to their position in the peptide sequence). Comparison of the amide chemical shift values of the M6-35P and M6-35L peptides revealed a significant shift in the resonances of S^{22} and L^{18} (0.7 and 0.5 ppm, respectively). Amide resonances of residues F^{25}, I^{24}, L^{19}, S^{17}, and Q^{16} also displayed changes in their chemical shift of greater than 0.1 ppm (Figure 56).

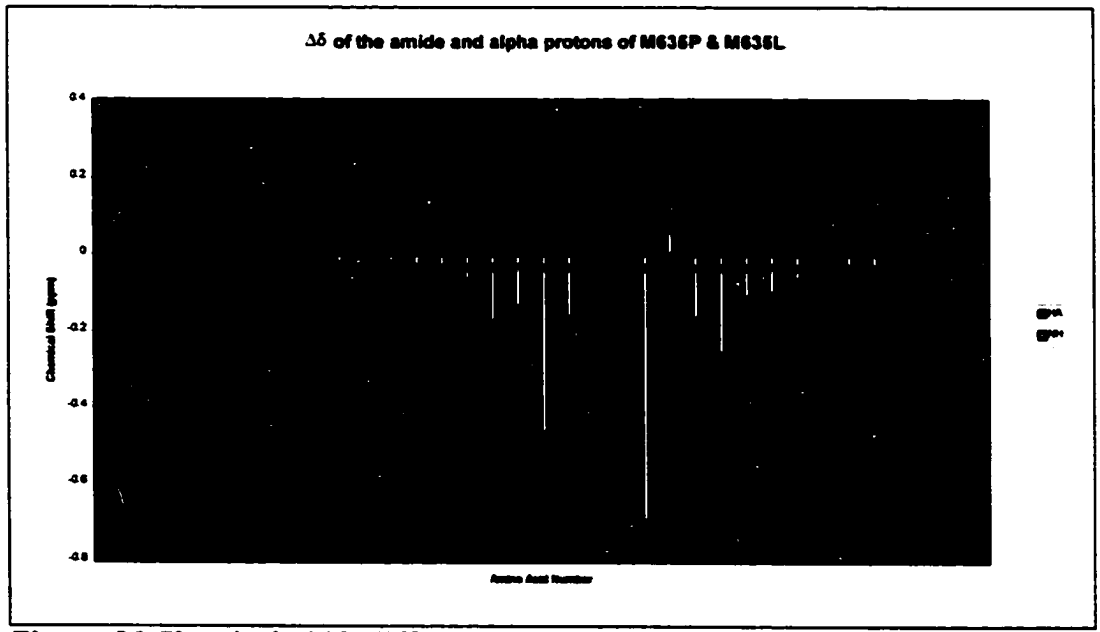

Figure 56:Chemical shift difference between αCH and NH resonances of M6-35P and M6-35L. The difference in the amide and α-protons of M6-35P with respect to M6-35L are plotted against the amino acid sequence.

The other amide resonances of these peptides were usually within 0.05 ppm of

each other. The α protons of V^{20}, L^{19}, and L^{18} also manifest changes in their chemical shift in the range of 0.07-0.24 ppm. These changes indicate that the substitution of a leucine for proline residue perturbed the structure from residue 16 to residue 25 of the peptide. The chemical shift differences are also manifested to some extent in the differences in the long-range connectivities around residue 21 in both peptides. Interestingly, very few long-range connectivities were observed below residue 15 in M6-35L. The reason for this will be considered in the Discussion.

Studies on a ^{15}N labeled transmembrane peptide

In order to obtain a good 3D structure for a long polypeptide using solution structure NMR spectroscopy it is essential to make complete assignments of all protons. This is quite challenging for large proteins and for membrane proteins that are highly helical and composed predominantly of hydrophobic amino acid residues. To overcome this problem heteronuclear NMR experiments are quite useful. We therefore expressed an M6 peptide with ^{15}N labels and carried out NMR studies on this model compound.

The sixth transmembrane domain fused to the TrpΔLE leader protein was expressed in minimal media containing $^{15}NH_4Cl$ as a sole source of nitrogen. The resulting ^{15}N labeled fusion protein was cleaved with CNBr and the ^{15}N-M6 isolated and purified as outlined previously. Solution NMR spectroscopy was carried out dissolving the ^{15}N-M6 in $CDCl_3$:CD_3OH:H_2O (4:4:1) to a concentration of 0.7 mM and analyzing the correlations between directly bonded ^1H and ^{15}N atoms through 2D heteronuclear

single quantum coherence (HSQC) experiments (Bodenhausen et al., 1980). 3D-NMR experiments such as TOCSY-HSQC (Marion et al., 1989; Zhang et al., 1994) and NOESY-HSQC (Ikura et al., 1990; Zhang et al., 1994) were also performed on the solution of the labeled peptide. Figure 57 shows the HSQC spectrum of ^{15}N-M6.

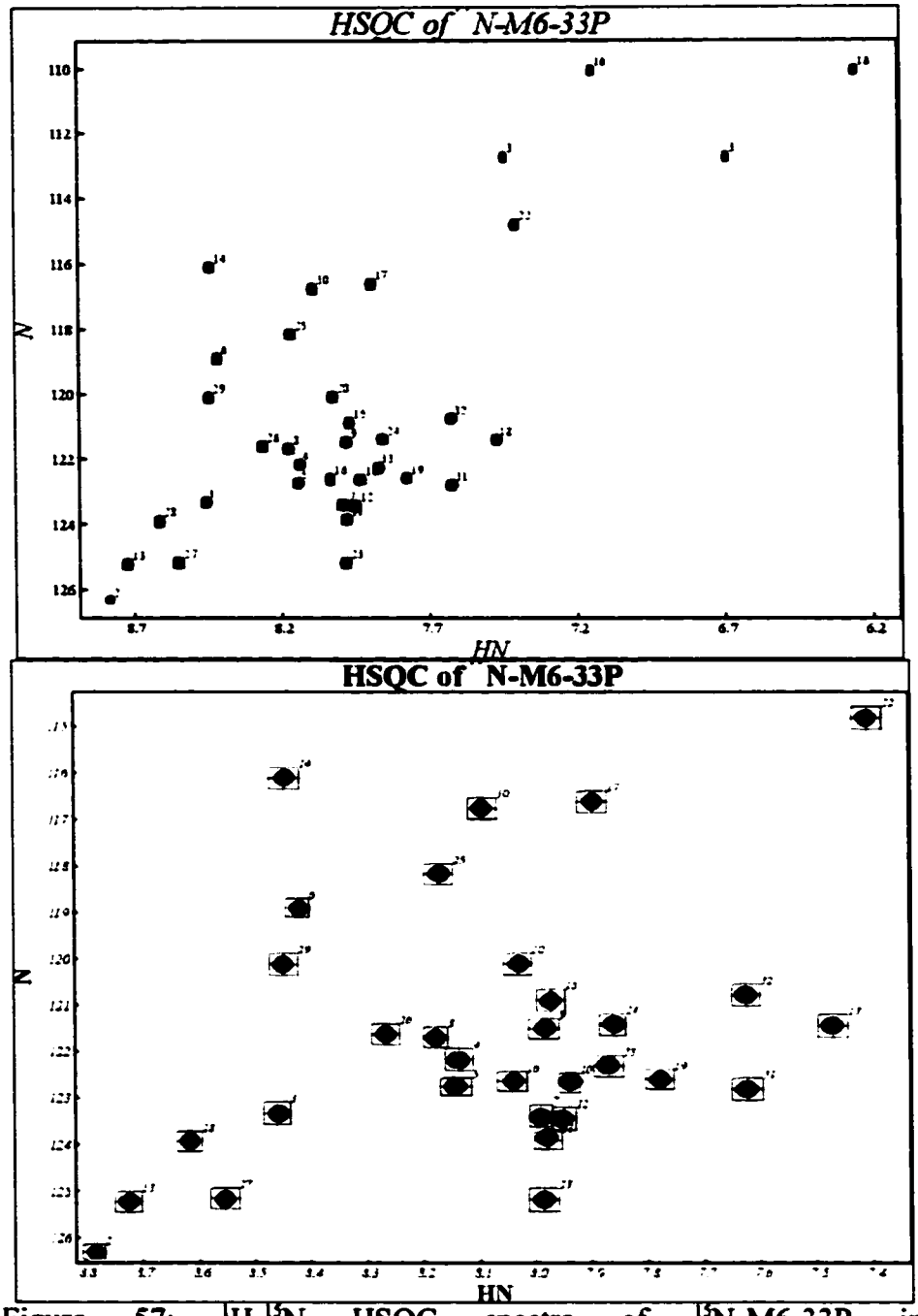

Figure 57: ^{1}H-^{15}N HSQC spectra of ^{15}N-M6-33P in $CDCl_3$:CD_3OH:H_2O (4:4:1). The top panel is the full spectrum depicting the glutamine side chains of residues 3 and 16 in the upper right corner. The bottom panel is the cropped spectra depicting only the backbone correlation peaks. The numbers on each peak represent the assignments of each correlation to each amino acid in the peptide sequence.

The spectrum is extremely well resolved. Thirty-one backbone NH correlations and four side chain amide (NH_2) correlations are observed. The results strongly indicate that labeled membrane peptides can be produced biosynthetically and subjected to multidimensional NMR analysis. The 3D spectra will be analyzed in the future.

NMR Modeling of the sixth transmembrane domain

Structure calculation of the M6-35P and M6-35L was performed using the distance constraints found by NMR. NMR distances computed from the NOESY cross peaks were input into the Crystallography & NMR System (CNS) software to calculate a series of related structures that satisfy the constraints found by NMR (Brunger et al., 1998). The structures are calculated using ab initio simulated annealing starting from an extended template structure and constraining the conformations with the interproton distance estimates, hydrogen bonds, and coupling-constant-derived dihedral angle restraints. In M6-35P, of a total of 225 NOEs, 125 corresponded to inter residue Overhauser effects. For M6-35L, 126 NOEs constituted inter residue correlations in addition to the 101 intraresidue NOEs observed during the NMR analyses. The results of the annealing yielded 20 structures of each peptide with a RMSD of approximately 5.5 Å and 6.5 Å for M6-35P and M6-35L, respectively. Figures 58 and 59 show one of the structures generated for each peptide with the least number of dihedral angle violations.

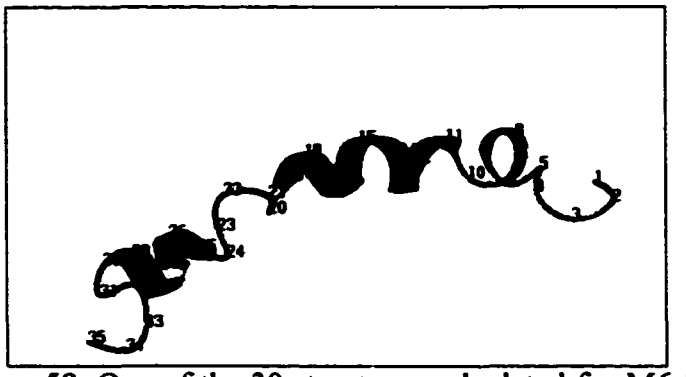

Figure 58: One of the 30 structures calculated for M6-35P by the CNS software. The numbers represent the position of the amino acids in the sequence. $K^{1}K^{2}K^{3}F^{4}D^{5}S^{6}F^{7}H^{8}I^{9}L^{10}L^{11}I^{12}M^{13}S^{14}A^{15}Q^{16}S^{17}L^{18}L^{19}V^{20}\mathbf{P}^{21}S^{2}{}^{2}I^{23}I^{24}F^{25}I^{26}L^{27}A^{28}Y^{29}S^{30}L^{31}K^{32}K^{33}K^{34}S^{35}$

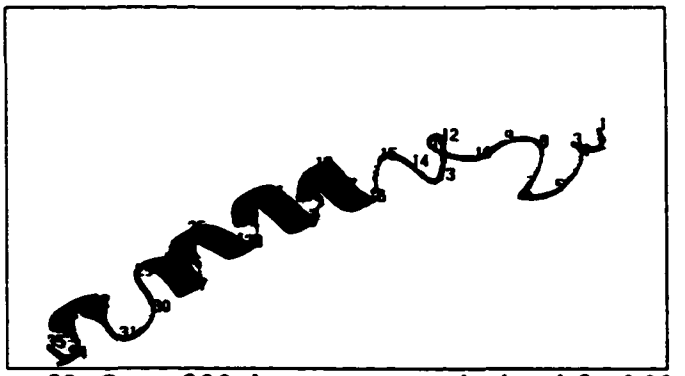

Figure 59: One of 30 the structures calculated for M6-35L by the CNS software. The numbers represent the position of the amino acids in the sequence. $K^{1}K^{2}K^{3}F^{4}D^{5}S^{6}F^{7}H^{8}I^{9}L^{10}L^{11}I^{12}M^{13}S^{14}A^{15}Q^{16}S^{17}L^{18}L^{19}V^{20}L^{21}S$ $^{22}I^{23}I^{24}F^{25}I^{26}L^{27}A^{28}Y^{29}S^{30}L^{31}K^{32}K^{33}K^{34}S^{35}$

Close examination of the different structures generated by the annealing showed that the low resolution obtained may be a consequence of the low number of distance constraints used for the calculations. Inclusion of the $^{3}J_{HN\alpha}$ constants in the calculations did not improve the results of the calculation. In addition, more dihedral angles were

observed that fell into the forbidden regions of the Ramachandran map. Due to the low number of distance constraints per residue, the structures generated for each peptide did not provide a high resolution average structure.

Chapter IV: Discussion

Biological assays

The mutations generated in Ste2p with the intention to replace Met^{218} and Met^{250} lead to a set of receptor mutants that were partially functional compared to the wild type as judged by the different assays explained in the results section. The halo assay that measures the growth arrest response of the cells to pheromone induction over a relatively long period of time, showed that the single mutation Met218Leu and Met250Ala and the double mutation Met218Leu, Met250Ala functioned similarly to the wildtype receptor but exhibited lower sensitivity to α-factor. Sommers and Dumont (1997) previously reported that the Met218Thr mutation exhibited a normal pheromone response. In the present case, the response of the Met218Leu showed a slight decrease in sensitivity. Sommers et al. (1997) stated that the Met218Thr mutation showed suppressor properties since it repaired the loss of function caused by the Glu143Lys and Thr144Pro mutations but it could not suppress the weak loss of function of Ile142Asn. They suggested that the "partial allele specificity" indicated the existence of some physical interaction of the methionine of the fifth transmembrane domain and some residues in the third transmembrane region. More recent studies conducted by Dube and others (2000) identified specific interactions between Val^{223} and Leu^{247} from the fifth and sixth transmembrane regions. It is likely that substitution of Met^{218} will have

an effect on the interaction of the fifth transmembrane domain with some other transmembrane region of Ste2p.

Fusion of the *LacZ* gene to the *FUS*1 yeast gene generated the *FUS*1-*LacZ* gene that did not alter the functionality of the *FUS1* gene in response to the mating signals but allowed the qualitative and quantitative measurement of the levels of response to the pheromone signal in a short period of time (2 h). The β-gal experiments conducted on the wild type and all the mutated receptors showed responses that were qualitatively similar to the ones obtained in the growth arrest assays. The mutant receptors were less sensitive to α-factor than wild-type Ste2p but the mutations did not cause loss of function of the mating receptor.

The decrease in response or sensitivity to the signal of the pheromone could be due to erroneous localization of the receptor or to deficient receptor-ligand binding. Therefore, binding assays were conducted. These experiments concluded that in all cases (except that the reconstructed A1314 strain harboring the Met218Leu mutation was not characterized [see Results]) the strains containing the mutant receptors were able to express and localize the receptor protein within the cell membrane. In general, the binding affinity of the receptor to the ligand decreased in all the mutant receptors with respect to the wild type. However, the K_D only decreased from about 4 nM for wild-type receptor to about 10 nM for the double mutation Met218Leu, Met250Leu, which showed the lowest affinity among all the mutated receptors. The binding experiments show a good agreement with the results of the halo assay and β-gal assays,

since the double mutation Met218Leu, Met250Leu consistently showed the lowest response.

The biological assays performed on the mutated receptors indicate that the substitution of the methionines located in the fifth and sixth transmembrane domains resulted in receptors that maintained both signaling and binding activities. The Met218Leu, Met250Ala double mutation had significantly better function compared to the Met218Leu, Met250Leu homolog and was used for expression of peptides for biophysical analyses.

Expression, purification and characterization of the fusion proteins

Expression of the M5I3M6 region of the Ste2p in a commercial plasmid, pRSETB, did not provide successful expression of the membrane protein in *E. coli.* BL21(DE3)pLysS is a tight bacterial system that lacks the *lon* and *ompT* proteases. It also keeps low levels of T7 RNA polymerase in order to minimize basal expression of toxic heterologous proteins and only promotes the expression of the target protein upon induction with IPTG (Studier et al., 1990). Despite the aforementioned properties of the bacterial strain, expression using the pRSETB plasmid was unsuccessful, as seen in Figure 15, which shows no difference in expression for transformed cells and non-transformed cells. Plasmid instability experiments demonstrated that most of the cells lost the plasmid coding for M5I3M6 before the induction with IPTG despite the

presence of high concentrations of ampicillin in the media.

The use of a pET derived plasmid that contained the TrpΔLE leader peptide gave better results for the expression of the target membrane protein. The TrpΔLE peptide, derived from the tryptophan operon protein, contained a ribosome-binding site that could serve as a site for initiation of translation (Miozzari et al., 1978). The TrpΔLE was previously used to successfully express globular proteins such as bovine pancreatic trypsin inhibitor (Staley et al., 1994) and the HIV "Vpu" membrane protein (Marassi et al., 1999). The reason that the vector containing the TrpΔLE showed better results than the pRSET derived vector above is that in addition to having a strong initiation site for translation, the TrpΔLE peptide triggers the formation of protein aggregates (inclusion bodies), thus minimizing the toxicity of the heterologous protein and secluding the expressed protein from the deleterious action of endogenous proteases (Uhlén et al., 1990). It was observed that transformation of BL21(DE3)pLysS strain with vectors that express TrpΔLE-M5I3M6 or TrpΔLE-M6 (Table 5) significantly reduced the size of the transformed colonies compared with the non-transformed cells or even with the cells transformed with the TrpΔLE expression vector. This effect was also noticed in some pET vectors (Studier et al., 1990). However, this phenotype does not interfere with the growth rate. If the expression of the membrane peptides fused to the TrpΔLE leader peptide had not been satisfactory, another alternative could have been the use of an expression system that efficiently exports the protein to the extracellular medium (Das, 1990).

Since the target membrane region of Ste2p was designed to be fused to the C-terminus of TrpΔLE through a methionine, the template plasmid used to clone the respective DNA sequences into the pMMHa plasmid was the mutated Ste2p plasmid pMD604 (Table 4) that contained the Met218Leu and Met250Ala mutations. This construct allowed release of the membrane portion from the leader polypeptide by chemical cleavage using CNBr. It is well known that CNBr cleaves at methionine residues (Walker et al., 1998). Although SDS polyacrylamide gel electrophoresis did not clearly show typical bands of overexpression of a protein relative to the endogenous proteins (Figure 19), the protein was detected by a Western blot type method using the histidine tag (Figure 23). The lack of staining properties in the gel electrophoresis even under different staining conditions was a constant feature observed with the fusion proteins. This effect seemed to be related to the small number of polar residues present in the hydrophobic sequences, as it is known that polar residues interact with the Coomassie dye and that amino groups in the side chains of certain residues form silver complexes (in the case of silver staining) (Patel et al., 1988).

Purification of the fusion protein containing the M5I3M6 or the M6 peptide was carried out using affinity chromatography and/or HPLC techniques. Both methods required stringent conditions. For example in order to solubilize the proteins high concentrations of urea or GuHCl were necessary. Solutions in these chaotropic solvents were used for affinity chromatography experiments. Alternatively, high temperatures (60 °C) and strong organic solvents such as 2-propanol:1-butanol in the mobile phase

were used in high-performance liquid chromatography. Purification of the fusion proteins from the inclusion bodies using nickel affinity chromatography provided low yields of the purified protein and the purity of the protein was never higher than 85%. The low yield obtained in the affinity chromatography may be a consequence of interference of the membrane part of the fusion protein with the binding of the His-tag to the immobilized Ni^{+2} ions. Passing the product from the first affinity column through a second column did not improve the yield of the purified fusion protein.

High performance liquid chromatography was carried out using 2-propanol:1-butanol/H_2O (0.1 % TFA) as a gradient system. These conditions were previously tried by Glover and coworkers (1999) in order to purify large quantities of insoluble membrane peptides. HPLC purification of the samples was performed as quickly as possible in order to avoid possible esterification reactions of serines and threonines or deamination of glutamine and asparagine residues due to the high temperature conditions and the presence of TFA in the solvent gradient system (Thomas et al., 1990). Due to the size of the fusion proteins and even of the TrpΔLE alone, HPLC columns containing 300 Å pore size were used as recommended by Hancock and Harding (1984). HPLC purification of the inclusion bodies containing the TrpΔLE-M5I3M6 or the TrpΔLE-M6 provided better yields than the Ni^{2+} affinity chromatography in terms of the amount and the quality of the pure proteins.

Despite the repeated purifications, the TrpΔLE-M5I3M6 showed multiple bands in SDS-PAGE (Figure 26). The presence of the extra bands was determined to be a

consequence of aggregation of the fusion protein with other molecules of the same species and perhaps also due to different folded forms that the fusion protein adopted during the running within the polyacrylamide gel (Figure 27). The same behavior during the gel electrophoresis was observed with the TrpΔLE-M6 fusion protein (Figures 28 and 29). Characterization of the TrpΔLE-M5I3M6 and TrpΔLE-M6 by amino acid analysis (Figures 36 and 37) gave evidence that the putative purified proteins contained the proper proportion of most of the amino acids except for serines and threonines, which exhibit degradation during the extensive hydrolysis required for these membrane proteins (Arshava et al., 1998). The definitive verification of the structure of the peptides came from the mass spectrometry results that showed that the fusion proteins obtained through a bacterial system had the correct mass and no posttranslational modifications (Figures 33, 34, and 35).

Spectroscopic analyses

Conformational analysis of the TrpΔLE-M5I3M6 and TrpΔLE–M6 fusion proteins were carried out using CD, FT-IR, and NMR techniques. Circular dichroism experiments were conducted in different membrane environments ranging from organic solvents such as TFE, TFE-water mixtures, SDS micelles, and DMPC lipid vesicles. To study the secondary structure of the fusion proteins in micellar environments, the concentration of SDS was adjusted to be 0.5% (17.8 mM), which is above the critical micelle concentration (CMC) of SDS of about 8.2 mM (~0.23%) (Neugenbauer, 1990).

The molar protein:detergent ratio was 1:571 for TrpΔLE-M5I3M6 and 1:239 for TrpΔLE, or more importantly the molar protein:micelle ratio was calculated to be 1:9.5 and 1:4 for TrpΔLE-M5I3M6 and TrpΔLE, respectively. It is reasonable to believe that the CMC of SDS was somewhat lower than the number indicated above since SDS was dissolved in PBS and it is known that salt concentration lowers the CMC of SDS, therefore the molar protein:micelle ratio was probably lower than indicated by the above ratios (Neugenbauer, 1990; Baleja, 2001). The detergent concentration used in this study was the same as previously used in the laboratory by Xie and others (2000) to study single transmembrane peptides. It was a little lower than the molar protein:detergent ratio (1:800) used to study the outer membrane protein A (OmpA) (Kleinschmidt et al., 1999).

CD spectra of the TrpΔLE and TrpΔLE-M5I3M6 in SDS micelles (Figure 43) exhibited the typical pattern of an α-helical structure in the wavelength range of 200-240 nm. The split $\pi\pi^*$ and the enhanced $n\pi^*$ transitions at 208 and 222 nm, respectively (Johnson, 1990; Kallenbach et al., 1996) were clearly noticeable. The lack of information in the 180-200 nm UV-region was due to the high ionic strength that resulted in high absorption and thus unacceptable noise levels under the experimental conditions. Based on the qualitative information gained from the spectra of TrpΔLE-M5I3M6 and TrpΔLE in SDS micelles it can be concluded that both polypeptides form helical structures in this environment.

CD experiments on TrpΔLE and both fusion proteins at different concentrations

of TFE in water show patterns typical of α-helical structures (Figures 38, 39 and 41). In these solvent mixtures three typical absorption bands at 190, 208 and 222 nm corresponding to the exciton splitting of the $\pi\pi^*$ transition band and the enhanced $n\pi^*$ transition band, respectively, were observed (Woody, 1995).

The decrease in concentration of TFE led to a decrease in helical content in the three different polypeptides. However, even at 10% TFE each of the proteins still shows a helical profile. With the assumption that the TrpΔLE and the hydrophobic regions constitute independent domains in the fusion proteins, the circular dichroism spectrum of the TrpΔLE can be subtracted from that of TrpΔLE-M5I3M6 or of TrpΔLE-M6 to gain information about the M5I3M6 and M6 peptides. This mathematical procedure is based on the general principle that a CD spectrum of a protein can be analyzed as a linear combination of basis spectra (Venyaminov et al., 1996). Figures 40 and 42 show the spectra of M5I3M6 and M6, respectively, after the subtractions. Both membrane peptides manifest CD profiles of helical structures. The red shift of the dichroic absorption at 190 nm is more pronounced in the M5I3M6 than in the M6 peptide consistent with the number of membrane segments present in the double domain compared to the single transmembrane domain. The ratio of the dichroic absorptions $[\theta_{190}]:[\theta_{220}]$ is approximately 2:1 in all solvent mixtures, a characteristic feature of α-helical proteins. Qualitatively, it is noticeable that both peptides show different degrees of helical content at different concentrations of TFE. However, M6 is significantly less structured in 10% TFE, likely a consequence of the behavior of TrpΔLE leader protein

under these conditions.

Thermal denaturation experiments were conducted using the samples dissolved in 30% TFE in order to work with samples that contain relatively high amounts of helical structure, and to be within the range of TFE concentration where helix induction is free from the effect of TFE on the enthalpy change per residue upon helix formation (ΔH). This effect was observed to be null in the range of 0 to 30-35% TFE (Luo et al., 1997). Figures 44, 45, and 47 show the spectra of TrpΔLE, TrpΔLE-M5I3M6, and TrpΔLE-M6 in 30% TFE at different temperatures from 25 to 85 °C. It can be observed that all the samples show the typical α-helical pattern and the amount of helical content decreased upon increasing the temperature. However, complete denaturation was not observed at the highest temperature examined. If a helix-coil transition were to occur an S-shaped temperature profile would be expected. A sharp loss in tertiary structure would enhance this sigmoidal transition. Since the temperature profiles for the ellipticity at 222 nm were linear and parallel, it is concluded that there is no evidence of globular folding involving the TrpΔLE and the hydrophobic domains and that the conformation of the leader peptide is independent of the conformation of the Ste2p segments under study.

The biological environment of an integral membrane protein is generally composed of a mixture of phospholipids, fatty acids, glycolipids, lipopolysaccharides, and steroids. Despite the molecular complexity of the natural environment, individual or mixtures of phospholipids have been used to mimic the natural membrane environment, because they reliably resemble the natural milieu (Sanders et al., 2000; Baleja, 2001).

Lipid bilayers, lipid micelles and lipid vesicles are three important types of lipid environment used to study the properties of membrane proteins. Lipid vesicles have been extensively utilized to study the biophysical properties of different membrane peptides, for instance domains of the cystic fibrosis transmembrane conductance regulator (Wigley et al., 1998), mellitin (Sharon et al., 1999), the transmembrane of influenza hemagglutinin (Tatulian et al., 2000) and single transmembrane domains of Ste2p (Xie et al., 2000).

The procedure followed to prepare lipid vesicles was similar to the one outlined by Ong and others (1981). The vesicles used were composed of DMPC phospholipids. These vesicles were stable and formed rapidly in the presence of TrpΔLE and both fusion proteins. Figure 49 shows the results of the CD experiments of each of the proteins in DMPC vesicles at three different temperatures. The CD profile of TrpΔLE has a distinct minimum near 208 nm in contrast to the CD profiles of the fusion proteins, which have a broad trough between 225 and 205 nm. Nevertheless, the overall shapes reflect a significant amount of α-helix content. Both fusion proteins, TrpΔLE-M5I3M6 and TrpΔLE-M6, formed more stable helices than did TrpΔLE at elevated temperatures. Indeed the helical structure of the TrpΔLE peptide appears to break down at 75 °C. Since the TrpΔLE initially had a more helical structure than the fusion proteins, it is possible that the anchoring of TrpΔLE by the membrane peptides to the DMPC vesicles forced the leader peptide to adopt a different structure compared to what it adopts by itself. A CD spectrum represents contributions to the optical activity from

the average of the ensemble of structures present in solution. Unfortunately, it was not possible to accurately determine the protein concentration in the vesicle samples. Therefore, it was not possible to determine the structure of the membrane regions in these protein constructs. The conclusion that the reconstitution of the fusion proteins in lipid vesicles involved the insertion of the membrane region in the hydrophobic environment relied on the observation that the vesicles were significantly more stable in the presence of the fusion proteins and that the leader peptide by itself had a very distinctive CD temperature profile compared to that of the fusion proteins under the same experimental conditions. Finally, the fusion proteins had almost no water solubility in the absence of the vesicles.

Quantitative analysis of the circular dichroism spectra carried out using model polypeptides as basis spectra yielded the percentage of α-helical structure calculated from the dichroic absorption at 222 and 208 nm (Tables 10 and 11). Some of the calculated values for the fusion proteins exceed 100%. This obvious overestimation of the amount of helix probably results because the value for an infinite helix was devised from soluble polypeptides in a more polar environment. The quantitative deconvolution confirms that for the three proteins, helical content decreases upon decrease in the concentration of TFE. This finding is consistent with the hypothesis that the fluorinated alcohol dehydrates the protein backbone and is less effective at lower TFE concentration.

Park and coworkers (1992) argue that the local environment affects the transition

bands of the chiroptical chromophore, and that the low dielectric constant inside the membrane is different from that experienced by helices in globular proteins. For this reason they assigned the values of the maximum ellipticities of the 208 and 222 nm bands to be –50,000 and –60,000 deg cm^2 $dmol^{-1}$, respectively. Tables 12 and 13 show the helical content values calculated based on the maximum absorption given by Park and others (1992). As expected the calculated percentage helical structure were lower for each protein using the values of Park et al. (1992). The question of whether the maximum absorption values describe by Park and coworkers are more accurate than the values used for globular proteins when applied to membrane proteins is still an unsettled issue. Wu et al. (1981) found that the CD magnitudes of poly(L-glutamic acid) at pH 4 were not affected when studied in the presence of SDS at the same pH. However, it was not demonstrated with certainty whether the molecules of poly(L-glutamic acid) were indeed immersed in the detergent micelles. It is known that SDS by itself is fairly transparent in the circular dichroism wavelength range (Johnson, 1990). Therefore, the difference in dichroic strength should be due to the electronic transition bands of the amide bonds in a particular dielectric medium.

The backpropagation neural network was the third method used to calculate the percent helical structure of the TrpΔLE, of both fusion proteins and of the hydrophobic peptides using mathematical subtraction. This method, unlike the model peptides described above, calculates the secondary structure composition based on defined regions of the spectrum and not based on rotational strength at two specific

wavelengths. The method works based on different layers called the neuron layers with specific functions, such as the input layer, the hidden layer and the output layer all connected in a neural type of network (Böhm et al., 1992). The software used for deconvolution of the CD spectra relies on a training set of CD spectra, which dated from the year 1997. For the samples in various concentrations of TFE in water, application of the neural network to the CD in the region 190-260 nm predicted values that ranged from 27 to 93 % α-helical structure (Table 14). The percentage of helicity found with the neural network for the samples in SDS were only 10 and 7% lower than the percentages calculated based on the model peptides, but in general both methods predict low helical content that could be related to the lack of precise concentration of the samples.

Comparison of the three strategies to calculate the helical content of the samples in TFE (Table 16) shows that the model based peptides based on maximum values used for membrane peptides always yielded lower α-helical percentages. The neural network, except for a few values in the row of the TrpΔLE-M5I3M6 and TrpΔLE-M6, in general predicted the highest percentages among the three calculations. Taken together these results provide evidence for the helical tendency of both the M6 peptide and the double domain polypeptide M5I3M6. Although under the same conditions the M6 synthetic peptides studied in the laboratory (Arshava et al., 1998; Xie et al., 2000) showed a somewhat higher helical content than M6 in the fusion protein, in general the results of M6 from the subtraction method are reasonable. Thus the 35-40% helical content found

for the double domain M5I3M6 in 100% TFE is a good first estimation for the secondary structure of this polypeptide. This value is only 15-20% lower than the helicity calculated from hydropathy analysis on this region of Ste2p.

FT-IR is a spectroscopic technique that provides information on the secondary structure of peptides derived from the analysis of the amide groups of the peptides. The results of the FT-IR experiments carried out showed that the proteins in DMPC multilayers exhibit a strong amide I band at 1658 cm^{-1} that is characteristic of an α-helical structure. Unfortunately, it was not possible to quantitate the amount of H/D exchange that occurred in each of the polypeptides. Interestingly, the TrpΔLE-M6 and the TrpΔLE-M5I3M6 fusion proteins did not form stable complexes in DMPC:DMPG vesicles, therefore FT-IR experiments on the fusion peptides could not be conducted in these vesicles.

A previous FT-IR study on a chemically synthesized M6-31 peptide in DMPC bilayers (Ding et al., 2001b) showed that it adopts a predominant β-sheet structure. Another FT-IR study of the sixth transmembrane containing 6 extra flanking lysines (M6-35P) in DMPC:DMPG bilayers (Xie et al., 2000) indicated that this peptide adopted mainly a β-structure under the conditions of their experiments. Previous ATR-FTIR studies in DMPC vesicles of the sixth transmembrane domain containing only 18 residues (M6-18) (Valentine et al., 20001) showed that the 18-mer peptide assumed mainly an α-helical structure with a small amount of β-sheet. The fusion protein TrpΔLE-M6 is similar to M6-35P because it also contains extra flanking lysines. The

FT-IR results of the TrpΔLE and TrpΔLE-M6 showed a complex series of bands in the amide I region, making it impossible to deconvolute both spectra and determine the specific absorption due to the biosynthetic M6-33 peptide. Moreover, FT-IR cannot be used to measure helix content of proteins, since turns or loops contribute in the spectral region near 1650 cm^{-1} and contributions around 1640 cm^{-1} have been assigned to random conformations (Hutchison et al., 1998). Furthermore, a band at 1630 cm^{-1} was present in both TrpΔLE and in TrpΔLE-M6 protein. Therefore, this band cannot be directly attributed to M6 and it cannot be used as an indication of the presence of β-structure in M6 as reported in the aforementioned cases. Thus the FT-IR results do not allow conclusive determination of the secondary structure adopted by the fusion proteins in DMPC multilayers.

The preliminary IR experiments suggest that when more material becomes available it will be worthwhile to see if accurate data on H/D exchange can be obtained. This would provide experimental evidence for the location of the fusion protein with respect to the lipid. In particular it may allow a distinction between the location of the TrpΔLE leader and the membrane domain with respect to the multilayer.

NMR spectroscopy is a versatile and powerful technique that along with X-ray crystallography constitutes a principal method for structure determination of proteins at atomic resolution. Due to the inherent difficulty in carrying out X-ray and NMR experiments on membrane proteins, fewer than 1% of the structures deposited in the protein data bank represent membrane proteins (Marassi et al., 1998; Baleja, 2001).

Despite the technical limitations of performing such studies, several solution NMR studies of membrane proteins have been carried out in micelles, bicelles and lipid bilayers. The latter investigations utilized solid-state NMR techniques (Marassi et al., 1998). NMR experiments on TrpΔLE, TrpΔLE-M5I3M6 and TrpΔLE-M6 in 50% TFE did not provide any structural information about the three proteins. Although CD results showed that the three proteins contain significant amounts of helical structure in this solvent, the quality of the NMR spectra was very poor. This could be understood as an intrinsic problem of large proteins and their slow motion on the NMR timescale, which causes large line broadening substantially decreasing the quality of the spectra. The relatively small dispersion of the α-helical amide resonances and the presence of many random coil resonances that also resonate in the same region, results in large regions of overlapping peaks. This further hampered the process of peak assignment. Finally, the NMR experiment was conducted at a 20-fold higher concentration than the CD studies. Therefore, it is possible that inter-protein association occurred, further complicating the spectral analysis. No attempts to investigate the structure of the fusion proteins in micelles were made since it is known that line width in TFE/H_2O mixtures is usually narrower than in micellar systems due to the increase of the effective molecular mass of the detergent-protein complex (Opella, 1997; Vinogradova et al., 1997).

As a preamble to study the isotopically labeled ^{15}N-M6, chemically synthesized M6-35P and M6-35L (Xie et al., 2000) were studied by NMR in the presence of the organic solvent mixture $CDCl_3$:CD_3OH:H_2O (4:4:1). As mentioned in the Results

section this solvent system had been previously used to study the c subunit of the F_1F_0 ATP synthase (Rastogi et al., 1999) and the integral membrane colicin E1 immunity protein (Taylor et al., 2000). The results of the NMR assignments of M6-35P and M6-35L and the chemical shift index plot show that both peptides form partially helical structures in $CDCl_3$:CD_3OH:H_2O (4:4:1). Previous NMR studies of M6-18 in TFE/H_2O (Arshava et al., 1998) showed that this membranous peptide adopted an α-helical structure from the proline residue towards the C-terminus of the peptide, with the N-terminal residues showing disorder. For M6-35P in the $CDCl_3$:CD_3OH:H_2O mixture, long-range connectivities were found to be strong and regular towards the carboxy terminus of the peptide and continued until residue number five from the amino terminus with a disruption around Pro21. In contrast M6-35L showed continuous long-range (α_i to NH_{i+3}) connectivities from residue 15 through the carboxyl terminus. The failure to observe many long-range connectivities at the N-terminus is partially due to the poorer chemical shift dispersion for the Leu21 containing homolog compared to the Pro21 homolog of the peptide. This decreased dispersion resulted from overlap that occurred with the Leu21 residue and other Leu residues in the molecule. Furthermore, because of time constraints the NOESY experiment was only run for 10 h and the signal to noise ratio was not high enough to distinguish all of the long-range NOEs.

Unfortunately, the number of distance constraints identified in both peptides was insufficient to determine high-resolution structures. As a step to improve the NMR data experiments were performed in the same solvent system on the ^{15}N-M6 peptide

obtained biosynthetically. The 2D HSQC experiment showed excellent resolution. Every expected amide and the glutamine side-chains were clearly observed (Figure 56). Due to time constraints no structure determination was achieved based on the ^{15}N-M6 peptide. However, data on these peptides was obtained using 3D-TOCSY-HSQC and 3D-NOESY-HSQC experiments. These data are currently being analyzed and will hopefully result in a high resolution structure for this transmembrane domain in $CDCl_3$:CD_3OH:H_2O (4:4:1). Finally, the general applicability of mixed aqueous organic solvents for structure determination of membrane domains is still not proven. Therefore, the biological relevance of such structures must be viewed with caution.

Chapter V: Conclusions

The objective of this dissertation was to develop biosynthetic approaches that would permit expression of milligram quantities of receptor domains for biophysical analyses and to initiate conformational analysis on these materials. The specific target that was selected initially was the M5I3M6 domain of Ste2p. This was later extended to include the single transmembrane domain M6. Analysis of the results indicates that successful biosynthesis and purification of fusion proteins TrpΔLE-M5I3M6 and TrpΔLE-M6 in milligram quantities was achieved. These fusion proteins were engineered so that the wild-type Met residues in transmembranes five and six were replaced with Leu and Ala residues to allow CNBr release of the product. Bioassays on the mutated receptors demonstrated that such replacements gave functional molecules. Finally, TrpΔLE-M6 protein was isotopically labeled and cleaved to release a uniformly ^{15}N labeled sixth transmembrane domain peptide.

Since the release of the hydrophobic peptides from the leader peptide would significantly lower the solubility of these peptides, studies of the secondary structure of the membrane peptides were carried out while they were still covalently attached to the TrpΔLE polypeptide. Circular dichroism experiments on TrpΔLE-M6 and TrpΔLE-M5I3M6 fusion proteins in TFE indicated that the respective transmembrane domains remained helical when bound to the leader peptide. The M5I3M6 also adopted a helical structure in SDS micelles. A less well defined structure was obtained in lipid vesicles,

but insertion of the membrane peptide portion of the fusion protein into the lipid multilayers was clearly observed.

In general both the IR and NMR experiments conducted on the fusion peptides were inconclusive, although the former experiment indicated that helical structures were present in both molecules. The NMR experiment on ^{15}N-M6 gave a very well resolved HSQC spectrum in $CDCl_3$:CD_3OH:H_2O providing impetus for further studies on the transmembrane domains in this medium. Infrared spectroscopy results of the fusion proteins in DMPC vesicles also indicated some helical structure of the constructs in this particular environment. However, this technique did not allow determination of the secondary structure of the fusion peptides in this milieu or localization of the leader peptide relative to the lipid bilayers.

Future Work

In order to understand the mechanism by which the α-factor receptor internalizes the signal of the pheromone ligand, high-resolution structures of Ste2p are needed both for its resting state and its activated state. Based on the conclusions of this dissertation and the current state of my research a number of experiments need to be completed and several other experiments can be suggested. These are included as follows.

Biosynthesis of the isotopically labeled sixth transmembrane domain containing the natural occurring proline at position 258 was achieved in this project. ^{15}N-HSQC experiments were reported herein and 3D-NMR spectroscopy has been performed in $CDCl_3$:CD_3OH:H_2O (4:4:1). The spectral resolution seemed to be optimum to fully assign all intra-residue and inter-residue connectivities. The full interpretation of these spectra needs to be completed so that enough distance constraints can be identified to generate a high resolution model for this peptide. Once this is accomplished similar studies should be carried out in DPC micelles. This detergent has been reported to have desirable properties for NMR studies (Baleja, 2000; Sanders et al., 2000; Kallick et al., 1995) and is a model for the lipid-like interior of the membrane. A comparison of the structure obtained in the two membrane mimetic environments would give insights into the "true" conformational tendencies of M6 and also validate $CDCl_3$:CD_3OH:H_2O (4:4:1) as a medium to investigate membrane peptides. I also suggest that to

complement the structural studies, analyses of the flexibility of this peptide should be performed using NMR relaxation experiments. After solving the structure of the M6 peptide as above I suggest that it be investigated in the presence of synthetic unlabeled peptides such as M5 and/or M7 to examine the effect of neighboring peptides on the structure of M6. Finally all the above approaches can be extended to a biosynthetic mutant M6 where Pro258 is mutated to Leu in order to compare the structural differences between the wild type transmembrane peptide and the peptide harboring the mutation that causes the constitutive activation of the receptor.

The double domain of Ste2p was also successfully expressed and purified as a fusion protein in the laboratory. Further attempts to cleave, isolate and purify the M5I3M6 peptide should be performed following the procedures developed in this thesis. Solubilization and deaggregation methods developed for the fusion proteins should be used to facilitate the studies of M5I3M6. In particular, solubilization of the peptide with mixtures of TFE/HFIP/H_2O should be done in order to deaggregate the hydrophobic molecules, after lyophilization of the cleavage reaction mixture. Ultimately uniformly labeled peptide should be biosynthesized in order to obtain a high-resolution structure of this important domain of the Ste2p receptor.

Bibliography

Abel, M. G., Zhang, Y., Lu, H., Naider, F. & Becker, J. (1998). Structure-function analysis of the *Saccharomyces cerevisiae* tridecapeptide pheromone using alanine-scanned analogs. *J. Peptide Res.* **52**, 95-106.

Arshava, B., Liu, S., Jiang, H., Breslav, M., Becker, J. & 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.

Baenziger, J. & Méthot, N. (1995). Fourier transform infrared and hydrogen/deuterium exchange reveal an exchange-resistant core of α-helical peptide hydrogens in the nicotinic acetylcholine receptor. *J. Biol. Chem.* **270**, 29129-29137.

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

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

Baleja, J. (2001). Structure determination of membrane-associated proteins from nuclear magnetic resonance data. *Anal. Biochem.* **288**, 1-15.

Blumer, K., Reneke, J., 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.

Blumer, K. & Thorner, J. (1991). Receptor-G Protein Signaling in Yeast. *Annu. Rev. Physiol.* **53**, 37-57.

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

Böhm, G., Muhr, R. & Jaenicke. R. (1992). Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.* **5**, 191-195.

Bourne, H. R., Meng, E. C. (2000). Rhodopsin Sees the Light. *Science* **289**, 733-734.

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

Breslav, M., Becker, J. & Naider, F. (1997). Dithioketal formation during synthesis of Bpa containing peptides. *Tetrahedron Lett.* **38**, 2219-2222.

Breslav, M., McKinney, A., Becker, J. & Naider, F. (1996). Preparation of radiolabeled peptides via an iodine exchange reaction. *Anal. Biochem.* **239**, 213-217.

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

Bukusoglu, G. & Jenness, D. (1996). Agonist-specific conformational changes in the yeast alpha-factor pheromone receptor. *Mol. Cell. Biol.* **16**, 4818-4823.

Carlson, S., Chatterjee, T., Murphy, K. & Fisher, R. (1998). Mutation of a putative amphipathic α-helix in the third intracellular domain of the platelet-activating factor receptor disrupts receptor/G protein coupling and signaling. *Mol. Pharmacol.* **53**, 451-458.

Cavanagh, J. Fairbrother, J., Palmer III, G., Skelton, N. (1996). Protein NMR spectroscopy. Principles and Practice. pp. 300-304, Academic Press. San Diego,

California.

Cetin, R., Anborg, P., Cool, R. & Parmeggiani, A. (1998). Functional role of the noncatalytic domains of elongation factor Tu in the interaction with ligands. *Biochemistry* **37**, 486-495.

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

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

Clore, G. & Gronenborn, A. (1989). Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. *Crit. Rev. Biochem. Mol. Biol.* **24**, 479-564.

Cole, G. & Reed, S. (1991). Pheromone-induced phosphorylation of a G protein β subunit in *S. cerevisiae* is associated with an adaptive response to mating pheromone. *Cell* **64**, 703-716.

Corbin, J., Méthot, N., Wang, H., Baenziger, J. & Blanton, M. (1998). Secondary

structure analysis of individual transmembrane segments of the nicotinic acetylcholine receptor by circular dichroism and Fourier transform infrared spectroscopy. *J. Biol. Chem.* **273**, 771-777.

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

Czerski, L., Vinogradova, O. & Sanders, C. (2000). NMR-Based amide hydrogen-deuterium exchange measurements for complex membrane proteins: development and critical evaluation. *J. Magn. Reson.* **142**, 111-119.

Das, A. (1990). Overproduction of proteins in *Escherichia coli*: vectors, hosts, and strategies. *Methods Enzymol.* **182**, 93-112.

David, N., Gee, M., Andersen, B., Naider, F., Thorner, J. & Stevens, R. (1997). Expression and Purification of the *Saccharomyces cerevisiae* α-Factor Receptor (Ste2p), a 7-Transmembrane-segment G-Protein Coupled Receptor. *J. Biol. Chem.* **272**, 15553-15561.

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

277-293.

Demmers, J., Haverkamp, J., Heck, A., Koeppe II, R. & Killian, J. (2000). Electrospray ionization mass spectrometry as a tool to analyze hydrogen/deuterium exchange kinetics of transmembrane peptides in lipid bilayers. *Proc. Natl. Acad. Sci. USA* **97**, 3189-3194.

De Planque, M., Kruitjtzer, J., Liskamp, R., Marsh, D., Greathouse, D., Koeppe II, R., de Kruijff, B. & Killian, J. (1999). Different membrane anchoring positions of tryptophan and lysine in synthetic transmembrane α-helical peptides. *J. Biol. Chem.* **274**, 20839-20845.

Ding, F., Lee, B., Hauser, M., Davenport, L., Becker, J. & Naider, F. (2001a). Probing the binding domain of the *Saccharomyces cerevisiae* α-mating factor receptor with fluorescence ligands. *Biochemistry* **40**, 1102-1108.

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

Divall, G. (1984). *Methods in Molecular Biology, New Protein Techniques* **1**, 105-111,

Humana Press, Clifton, New Jersey.

Dohlman, H., Song, J., Ma, D., Courchesne, W. & Thorner, J. (1996). Sst2, a negative regulator of pheromone signaling in the yeast *Saccharomyces cerevisiae*: Expression, localization, and genetic interaction and physical association with Gpa1 (The G-protein α subunit). *Mol. Cell. Biol.* **16**, 5194-5209.

Dohlman, H. G. & Thorner, J. (1997). RGS proteins and signaling by heterotrimeric G proteins. *J. Biol. Chem.* **272**, 3871-3874.

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

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

Dosil, M., Giot, L., Davis, C. & Konopka, J. (1998). Dominant-negative mutations in the G protein-coupled α-factor receptor map to the extracellular ends of the transmembrane segments. *Mol. Cell. Biol.* **18**, 5981-5991.

Dosil, M., Schandel, K., Gupta, E., Jenness, D. & Konopka, J. (2000). The C terminus

of the *Saccharomyces cerevisiae* α-factor receptor contributes to the formation of preactivation complexes with its cognate G protein. *Mol. Cell. Biol.* **20**, 5321-5329.

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

Dubendorff, J. & Studier, F. (1991). Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with *lac* repressor. *J. Mol. Biol.* **219**, 45-59.

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

Dube, P., Konopka J. (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.

Fasman G. (1996). Determination between transmembrane helices and peripheral helices by deconvolution of circular dichroism spectra of membrane proteins, pp. 381-412. Circular dichroism and the conformational analysis of biomolecules, edited by G.

D. Fasman. Plenum Press. New York.

Gazit, E. & Shai, Y. (1995). The assembly and organization of the α5 and α7 Helices from the pore-forming domain of *Bacillus thuringiensis* δ-endotoxin. *J. Biol. Chem.* **270**, 2571-2578.

Gesell, J., Zasloff, M. & Opella, S. (1997) Two-dimensional ^{1}H NMR experiments show that the 23-residue magainin antibiotic peptide is an α-helix in dodecylphosphocholine micelles, sodium dodecylsulfate micelles, and trifluoroethanol/water solution. *J. Biomol. NMR* **9**, 127-135.

Glover, K., Martini, P., Vold, R. & Komives, E. (1999). Preparation of insoluble transmembrane peptides: glycophorin-A, prion 110-137, and FGFR 368-397. *Anal. Biochem.* **272**, 270-274.

Gounarides, J., Xue, C., Becker, J. & Naider, F. (1994). NMR Investigation of cyclo7,10[C7,X9,C10,Nle12] Analogues of the alpha-factor from *Saccharomyces cerevisiae*. *Biopolymers* **34**, 709-720.

Greenfield, N. & Fasman, G. (1969). Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **8**, 4108-4116.

Hammond, J. & Kruger, N. (1988). *Methods in Molecular Biology, New Protein Techniques* **3**, 25-32, Humana Press, Clifton, New Jersey.

Hancock, W. & Harding D. (1984). Review of separation conditions. Handbook of HPLC for the separation of amino acids, peptides and proteins. **Vol. II**. 303-312, CRC Press, Inc. Boca Raton, Florida.

Haris, P. & Chapman, D. (1995). The conformational analysis of peptides using Fourier Transform IR spectroscopy. *Biopolymers* **37**, 251-263.

Henry, G. & Sykes, B. (1994). Methods to Study Membrane Protein Structure in Solution. *Methods Enzymol.* **239**, 515-535.

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

Herzyk, P. & 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., Zanolari, B. & Riezman, H. (1998). Cytoplasmic tail phosphorylation of the

alpha-factor receptor is required for its ubiquitination and internalization. *J. Cell. Biol.* **141**, 349-358.

Hirschman, J., De Zutter, G., Simonds, W. & 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.

Hutchison, R., Betts, S., Yocum, C. & Barry, B. (1998). Conformational changes in the extrinsic manganese stabilizing protein can occur upon binding to the photosystem II reaction center: an isotope editing and FT-IR study. *Biochemistry* **37**, 5643-5653.

Ikura, M., Marion, D., Kay, L., Shih, H., Krinks, M., Klee, C. & Bax, A. (1990). Heteronuclear 3D NMR and isotopic labeling of calmodulin. Towards the complete assignment of the 1H NMR spectrum. *Biochem. Pharmacol.* **40**, 153-160.

Iwata, S., Lee, J. W., Okada, K., Lee, J. K., Iwata, M., Rasmussen, B., Link, T. A., Ramaswamy, S. & Jap, B. K. (1998). Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex. *Science* **281**, 64-71.

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

Jiang, Y., Breslav, M., Khare, R., McKinney, A., Becker, J. & Naider, F. (1995). Synthesis of alpha-factor analogues containing photoactivatable and labeling groups. *Int. J. Peptide Protein Res.* **45**, 106-115.

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

Johnson, Jr, W. (1990). Protein secondary structure and circular dichroism: a practical guide. *Proteins: Struct. Funct. Genet.* **7**, 205-214.

Jones D., Ball, E., Sharpe, S., Barber, K. & Grant, C. (2000). Expression and membrane assembly of a transmembrane region from Neu. *Biochemistry* **39**, 1870-1878.

Kallal, L. & 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.

Kallenbach, N. Lyu, P & Zhou, H. (1996). CD spectroscopy and the helix-coli transition in peptides and polypeptides, pp. 201-259. Circular dichroism and the conformational analysis of biomolecules, edited by G. D. Fasman. Plenum Press. New York.

Kallick, D., Tessmer, M., Watts, C. & Li, C-Y. (1995). The use of dodecylphosphocholine micelles in solution NMR. *J. Magn. Reson.* **109**, 60-65.

Kataoka, M. & Ueki, T. (1981). X-ray diffraction studies on chromatophore membrane from photosynthetic bacteria. I. Diffraction pattern of the photoreaction unit isolated from Rhodospirillum rubrum chromatophore and some characteristics of the structure. *J Biochem.* **89**, 71-78.

Kleinschmidt, J., Wiener, M. & Tamm, L. (1999). Outer membrane protein A of *E. coli* folds into detergents micelles, but not in the presence of monomeric detergent. *Protein Science* **8**, 2065-2071.

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

Konopka, J., Jenness, D. & Hartwell, L. (1988). The C-terminus of the *S. cerevisiae* α-pheromone receptor mediates and adaptive response to pheromone. *Cell* **54**, 609-620.

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

Kostenis, E., Conklin, B. & 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.

Krittanai, C. & Johnson, W. (1997). Correcting the circular dichroism spectra of peptides for contributions of absorbing side chains. *Anal. Biochem.* **253**, 57-64.

Kunkel, T., Roberts, J. & Zakour R. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 488-492.

Lambright, D., Sondek, J., Bohm, A., Skiba, N., Hamm, H. & Siegler, P. (1996). The 2.0 A crystal structure of a heterotrimeric G protein. *Nature*, **379**, 311-319.

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

Linden, J. (1982). Calculating the dissociation constant of an unlabeled compound from the concentration required to displace radiolabel binding by 50%. *J. Cyclic Nucleotide Res.* **8**,163-172.

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

Liu, J., Schöneberg, T., Rhee M. & 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.

Liu, S., Henry, L., Lee, B., Wang, S., Arshava, B., Becker, J. & Naider, F. (2000). Position 13 analogs of the tridecapeptide mating pheromone from *Saccharomyces cerevisiae*: design of an iodinatable ligand for receptor binding. *J. Pept. Res.* **56**, 24-34.

Luecke, H., Schobert, B., Richter, H., Cartailler, J. & Lanyi, J. (1999). Structure of Bacteriorhodopsin at 1.55 Å Resolution. *J. Mol. Biol.* **291**, 899-911.

Luneberg, J., Widmann, M., Dathe, M. & Marti, T. (1998). Secondary structure of bacteriorhodopsin fragments. External sequence constraints specify the conformation of transmembrane helices. *J. Biol. Chem.* **273**, 28822-28830.

Luo, P. & Baldwin, R. (1997). Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from

trifluoroethanol/water mixtures back to water. *Biochemistry* **36**, 8413-8421.

Marassi, F., Ma, C., Gratkowski, H., Straus, S., Strebel, K., Oblatt-Montal, M., Montal, M. & Opella, S. (1999). Correlation of the structural and functional domains in the membrane protein Vpu from HIV-1. *Proc. Natl. Acad. Sci. U S A* **96**, 14336-14341.

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

Marassi, F., Ramamoorthy, A. & 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. USA* **94**, 8551-8556.

Marepalli, H., Antohi, O., Becker, J. & Naider, F. (1996). Solution structures of *i* to *i*+3 cyclized model peptides: building blocks mimicking specific conformations. *J. Am. Chem. Soc.* **118**, 6531-6539.

Marion, D., Driscoll, P., Kay, L., Wingfield, P., Bax, A., Gronenborn, A. & Clore, G. (1989). Overcoming the overlap problem in the assignment of 1H NMR spectra of larger proteins by use of three-dimensional heteronuclear 1H-15N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence

spectroscopy: application to interleukin 1 beta. *Biochemistry* **28**, 6150-6156.

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.

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

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

Miller, J. (1972). Experiments in Molecular Genetics. Cold Spring Harbor Laboratory. Cold Spring Harbor. New York.

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

Mittard, V., Morelle, N., Brutscher, B., Simorre, J., Marion, D., Stein, M., Jacquot, J.,

Lirsac, P. & Lancelin, J. (1995). ^{1}H, ^{13}C, ^{15}N-NMR resonance assignments of oxidized thioredoxin h from the eukaryotic green alga *Chlamydomonas reinhardtii* using new methods based on two-dimensional triple-resonance NMR spectroscopy and computer-assisted backbone assignment. *Eur. J. Biochem.* **229**, 473-485.

Moffat, B. & Studier, F. (1987). T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* **49**, 221-227.

Nakafuku, M., Obara, T., Kaibuchi, K, Miyajima, I., Miyajima, A., Ioth, H., Nakamura, S., Arai, K., Matsumoto, K. & Kaziro, Y. (1988). Isolation of a second yeast *Saccharomyces cerevisiae* gene (*GPA2*) coding for guanine nucleotide-binding regulatory protein: studies on its structure and possible functions. *Proc. Natl. Acad. Sci. USA* **85**, 1374-1378.

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

Neer, E. (1995). Heterotrimeric G proteins: organizers of transmembrane signals. *Cell* **80**, 249-257.

Neugenbauer, J. (1990). Detergents: an overview. *Methods Enzymol.* **182**, 239-253.

Ong, R., Marchesi, V. & Prestegard, J. (1981). Small unilamellar vesicles containing Glycophorin A. Chemical characterization and proton nuclear magnetic resonance studies. *Biochemistry* **20**, 4283-4292.

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

Opella, S., Kim, Y. & McDonnell, P. (1994) Experimental nuclear magnetic resonance studies of membrane proteins. *Methods Enzymol.* **239**, 536-560.

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

Park, K., Perczel, A. & Fasman G. (1992). Differentiation between transmembrane helices and peripheral helices by deconvolution of circular dichroism spectra of membrane proteins. *Protein Sci.* **1**, 1032-1049.

Parker, F. (1973). Applications of Infrared Spectroscopy in Biochemistry, Biology and Medicine. Plenum Press. New York.

Patel, K., Easty, D. & Dunn, M. (1988). Detection of proteins in polyacrylamide gels using an ultrasensitive silver staining technique. Chapter 13. 159-167. Methods in Molecular Biology. Vol. 3. Edited by J. M. Walker. Humana Press. Clifton, New Jersey.

Pervushin, K., Orekhov, V., Popov, A., Musina, L. & Arseniev, A. (1994). Three-dimensional structure of (1-71) bacterioopsin solubilized in methanol/chloroform and SDS micelles determined by ^{15}N-^{1}H heteronuclear NMR spectroscopy. *Eur. J. Biochem.* **219**, 571-583.

Provencher, S. & Glöckner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**, 33-37.

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

Rance, M., Sørensen, O., Bodenhausen, G., Wagner, G., Ernst, R. & Wütrich, K. (1983). Improved spectral resolution in cosy 1H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* **117**, 479-485.

Rao, H., Yang, W., Joshua, H., Becker, J. & Naider, F. (1995). Studies on conformational consequences of i to i + 3 side-chain cyclization in model cyclic tetrapeptides. *Int. J. Peptide Protein Res.* **45**, 418-429.

Rastogi, V. & Girvin, M. (1999). Structural changes linked to proton translocation by subunit c of the ATP synthase. *Nature* **401**, 263-268.

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

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

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

Ridge, K., Lee, S. & Yao, L. (1995). *In Vivo* assembly of Rhodopsin from Expressed Polypeptide Fragments. *Proc. Natl. Acad. Sci. USA* **92**, 3204-3208.

Ridge, K., Ngo, T., Lee, S. & 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.

Runyan Garrison, T., Zhang, Y., Pausch, M., Apanovitch, D., Aebersold, R. & Dohlman, H. G. (1999). Feedback phosphorylation of an RGS protein by MAP kinase in yeast. *J. Biol. Chem.* **274**, 36387-36391.

Sambrook, J., Fritsch, E. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sanders, C. & Oxenoid, K. (2000). Customizing model membranes and samples for NMR spectroscopic studies of complex membrane proteins. *Biochim. Biophys. Acta* **1508**, 129-145.

Scholtz, J., Qian, H., York, E., Stewart, J. & Baldwin R. (1991). Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water. *Biopolymers* **31**, 1463-1470.

Scopes R. K., Protein Purification, Principles and Practice, Springer-Verlag, New York, 1982.

Schena, M., Picard, D. & Yamamoto, K. (1991). Vectors for constitutive and inducible gene expression in yeast. *Methods Enzymol.* **194**, 389-398.

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

Sen, M. & 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. & 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.

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

Sharon, M., Oren, Z., Shai, Y. & Anglister, J. (1999). 2D-NMR and ATR-FTIR study of the structure of a cell-selective diastereomer of mellitin and its orientation in phospholipids. *Biochemistry* **38**, 15305-15316.

Siebert, F. (1995). Infrared spectroscopy applied to biochemical and biological problems. *Methods Enzymol.* **246**, 501-526.

Siesler, H. & Holland-Moritz, K. (1980). *Infrared and Raman Spectroscopy of polymers*, 53-56 Marcel Dekker, New York.

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

Sommers, C. & Dumont, M. (1999). Genetic approaches for studying the structure and function of G protein-coupled receptors in yeast, in *Structure-function analysis of G protein-coupled receptors*, edited by Jurgen Wess; John Wiley and Sons, New York.

Soulié, S., Neumann, J., Berthomieu, C., Møller, J., le Marie, M. & Forge, V. (1999). NMR conformational study of the sixth transmembrane segment of sarcoplasmic reticulum Ca^{+2}-ATPase. *Biochemistry* **38**, 5813-5821.

Soulimane, T., Buse, G., Bourenkov, G., Bartunik, H., Huber, R. & Than M. (2000). Structure and mechanism of the aberrant ba(3)-cytochrome c oxidase from thermus thermophilus. *EMBO J.* **19**, 1766-1776.

Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H. & Gouaux, J. (1996). Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science* **274**, 1859-1866.

Sprague, G. & Thorner, J. (1992). The Molecular and Cellular Biology of the Yeast *Saccharomyces* Gene expression, Cold Spring Harbor Lab. Press, New York.

Sprang S. (1997). G Protein mechanisms: insights from structural analysis. *Annu. Rev. Biochem.* **66**, 639-678.

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

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

Studier, F., Rosenberg, A., Dunn, J. & Dubendorff, J. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60-89.

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

Tatulian, S. & Lukas K. (2000). Secondary structure, orientation, oligomerization, and lipid interactions of the transmembrane domain of influenza hemagglutinin. *Biochemistry* **39**, 496-507.

Taylor, R., Zakharov, S., Bernard Heymann, J., Girvin, M. & Cramer, W. (2000). Folded state of the integral membrane colicin E1 immunity protein in solvents of mixed polarity. *Biochemistry* **39**,12131-12139.

Thomas, T. & McNamee, M. (1990). Purification of membrane proteins. *Methods Enzymol.* **182**, 499-521.

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

Unger, V., Hargrave, P., Baldwin, J. & Schertler, G. (1997). Arrangement of rhodopsin transmembrane alpha-helices. *Nature* **389**, 203-206.

Valentine, K., Liu, S., Marassi, F., Veglia, G., Opella, S., Ding, F., Wang, S., Arshava, B., Becker, J. & Naider, F. (2001). Structure and topology of a peptide segment of the sixth transmembrane domain of the *Saccharomyces cerevisiae* α-factor receptor in

phospholipid bilayers. *Biopolymers* **59**, 243-256.

van de Ven, F., van Os, J., Aelen, J., Wymenga, S., Remerowski, M., Konings, R. & 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.

Venyaminov, S. & Yang, J. (1996). Determination of protein secondary structure, pp. 69-107. Circular dichroism and the conformational analysis of biomolecules, edited by G. D. Fasman. Plenum Press. New York.

Vinogradova, O., Badola, P., Czerski, L., Sönnichsen, F. & Sanders, C. (1997). *Escherichia coli* diacylglycerol kinase: A case study in the application of solution NMR methods to an integral membrane protein. *Biophys. J.* **72**, 2688-2701.

Vold, R., Prosser, R. & Deese, A. (1997). Isotropic solutions of phospholipids bicelles: a new membrane mimetic for high-resolution NMR studies of polypeptides. *J. Biomol. NMR* **9**, 329-335.

Von Heijne, G. (1991). Proline kinks in transmembrane alpha-helices. *J. Mol. Biol.* **218**, 499-503.

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

Wall, M., Coleman, D., Lee, E., Iñiguez-Lluhi, J., Posner, B., Gilman, A. & Sprang, S. (1995). The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2. *Cell* **83**, 1047-1058.

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

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

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

Wishart, D., Sykes, B. & Richards, F. (1992). The chemical shift index: A fast and

simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **31**, 1647-1651.

Woody, R. (1995). Circular dichroism. *Methods Enzymol.* **246**, 34-71.

Wu, C., Ikeda, K. & Yang, J. (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*, Wiley, New York.

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

Xue, C., McKinney, A., Lu, H., Jiang, Y., Becker, J. & Naider, F. (1996). Probing the functional conformation of the tridecapeptide mating pheromone of *Saccharomyces cerevisiae* through study of disulfide-constrained analogs. *Int. J. Peptide Protein Res.* **47**, 131-141.

Yang, W., McKinney, A., Becker, J. & Naider, F. (1995). Systematic analysis of the *Saccharomyces cerevisiae* alpha-factor containing lactam constraints of different ring

size. *Biochemistry* **34**, 1308-1315.

Yansura, D. & Henner, D. (1990). Use of *Escherichia coli trp* promoter for direct expression of proteins. *Methods Enzymol.* **185**, 54-60.

Yau, W., Wimley, W., Gawrisch, K. & White, S. (1998) The Preference of Tryptophan for Membrane Interfaces. *Biochemistry* **37**, 14713–14718.

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

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

Zhang, O., Kay, L., Olivier, J. & Forman-Kay, J. (1994). Backbone ^{1}H and ^{15}N 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.

Zhou, P., Lugovskoy, A. & Wagner, G. (2001). A solubility-enhancement tag (SET) for

NMR studies of poorly behaving proteins. *J. Biomol. NMR* **20**, 11-14.