Biosynthesis and NMR Analysis of a 73-Residue Domain of a *Saccharomyces cerevisiae* G Protein-Coupled Receptor†

Racha Estephan,‡,§ Jacqueline Englander,‡ Boris Arshava,‡ Karen L. Samples,|| Jeffrey M. Becker,|| and Fred Naider*,‡,§

Department of Chemistry, College of Staten Island and Macromolecular Assemblies Institute of the City University of New York, Staten Island, New York 10314, Ph. D. Program in Biochemistry, The Graduate School and University Center of the City University of New York, New York, New York 10016, and Department of Microbiology and Graduate Program in Genome Science and Technology, University of Tennessee, Knoxville, Tennessee 37996

Received April 19, 2005; Revised Manuscript Received July 8, 2005

ABSTRACT: The yeast *Saccharomyces cerevisiae* α-factor pheromone receptor (Ste2p) was used as a model G protein-coupled receptor (GPCR). A 73-mer multidomain fragment of Ste2p (residues 267−339) containing the third extracellular loop, the seventh transmembrane domain, and 40 residues of the cytosolic tail (E3-M7-24-T40) was biosynthesized fused to a carrier protein. The multidomain fusion protein (designated M7FP) was purified to near homogeneity as judged by HPLC and characterized by mass spectrometry. In minimal medium, 30−40 mg of M7FP were obtained per liter of culture. The 73-residue peptide was released from its carrier by CNBr and obtained in wild-type, ^{15}N, and $^{13}C/^{15}N$ forms. The E3-M7-24-T40 peptide integrated into 1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)] and dodecylphosphocholine micelles at concentrations (200−500 μM) suitable for NMR investigations. HSQC experiments performed in organic solvents and detergent micelles on ^{15}N-labeled E3-M7-24-T40 showed a clear dispersion of the nitrogen-amide proton correlation cross-peaks indicative of a pure, uniformly labeled molecule that assumed a partially ordered structure. NOE connectivities, chemical shift indices, *J*-coupling analysis, and structural modeling suggested that in trifluoroethanol/water (1:1) helical subdomains existed in both the transmembrane and cytoslic tail of the multidomain peptide. Similar conclusions were reached in chloroform/methanol/water (4:4:1). As the cytosolic tail participates in down-regulation of Ste2p, the helical regions in the Ste2p tail may play a role in protein−protein interactions involved in endocytosis.

Membrane proteins currently constitute approximately 30% of the proteins encoded by the human genome (*1*). Despite their large number and physiological significance, relatively little is known about their three-dimensional structures. To date, less than 1% of this group of biomolecules, mostly membrane channels and transporters, has been successfully studied at the atomic level (*2*). The structure of the extracellular portions of some membrane receptors has been determined, but knowledge of the structure of entire membrane receptors is very limited. Among membrane receptors, the largest family is that of G protein-coupled receptors (GPCRs) (*3, 4*).[1] Currently, it is estimated that there are about 1000 GPCRs encoded by the human genome, yet only one high-resolution structure of a GPCR from any organism, that of bovine rhodopsin, elucidated at 2.8 Å resolution, has appeared in the literature (*5*).

GPCRs are integral membrane proteins whose amino acid sequences are characterized by seven hydrophobic segments that are predicted to form transmembrane (TM) α helices, connected by intracellular and extracellular loops (*6*), with the amino and carboxy-termini being located extracellularly and cytoplasmically, respectively (*3, 5*−*11*). GPCRs have been found in a wide range of organisms and mediate cellular responses to a wide variety of physical and chemical extracellular signals (*6*) by activating G proteins, which

† This work was supported by National Institutes of Health Research Grants GM22086, GM22087, and GM22086-27S1 and NSF Grant 0097446. Fred Naider is currently the Leonard and Esther Kurtz Term Professor at the College of Staten Island.

* To whom correspondence should be addressed: Department of Chemistry, College of Staten Island of The City University of New York, 2800 Victory Blvd, Staten Island, NY 10314. Telephone: (718) 982−3896. Fax: (718) 982−3910. E-mail: naider@mail.csi.cuny.edu.

‡ Department of Chemistry, College of Staten Island and Macromolecular Assemblies Institute of the City University of New York.

§ Ph. D. Program in Biochemistry, The Graduate School and University Center of the City University of New York.

|| University of Tennessee.

[1] Abbreviations: GPCRs, G protein-coupled receptors; CD, circular dichroism; DHPC, 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; DMPG, 1,2-dimyristoyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)]; DPC, dodecylphosphocholine; E3-M7-24-T40, a multidomain fragment containing the third extracellular loop, the seventh transmembrane domain, and 40 residues of the cytosolic tail of Ste2p; [θ], mean residue ellipticity; ESI-MS, electrospray ionization mass spectrometry; HPLC, high-performance liquid chromatography; HSQC, heteronuclear single quantum correlation; IPTG, isopropyl-β-D-thiogalactopyranoside; M7FP, TrpΔLE-E3-M7-24-T40 or E3-M7-24-T40 fusion protein; NOESY, nuclear Overhauser effect spectroscopy; PPG, 1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)]; SDS, sodium dodecyl sulfate; SDS−PAGE, SDS−polyacrylamide gel electrophoresis; Ste2p, α-factor pheromone receptor; TFA, trifluoroacetic acid; TFE, 2,2,2-trifluoroethanol; TM, transmembrane; TOCSY, total correlation spectroscopy; TrpΔLE, a portion of the histidine-tagged TrpΔLE1413 polypeptide.

10.1021/bi0507231 CCC: $30.25
Published on Web 08/11/2005

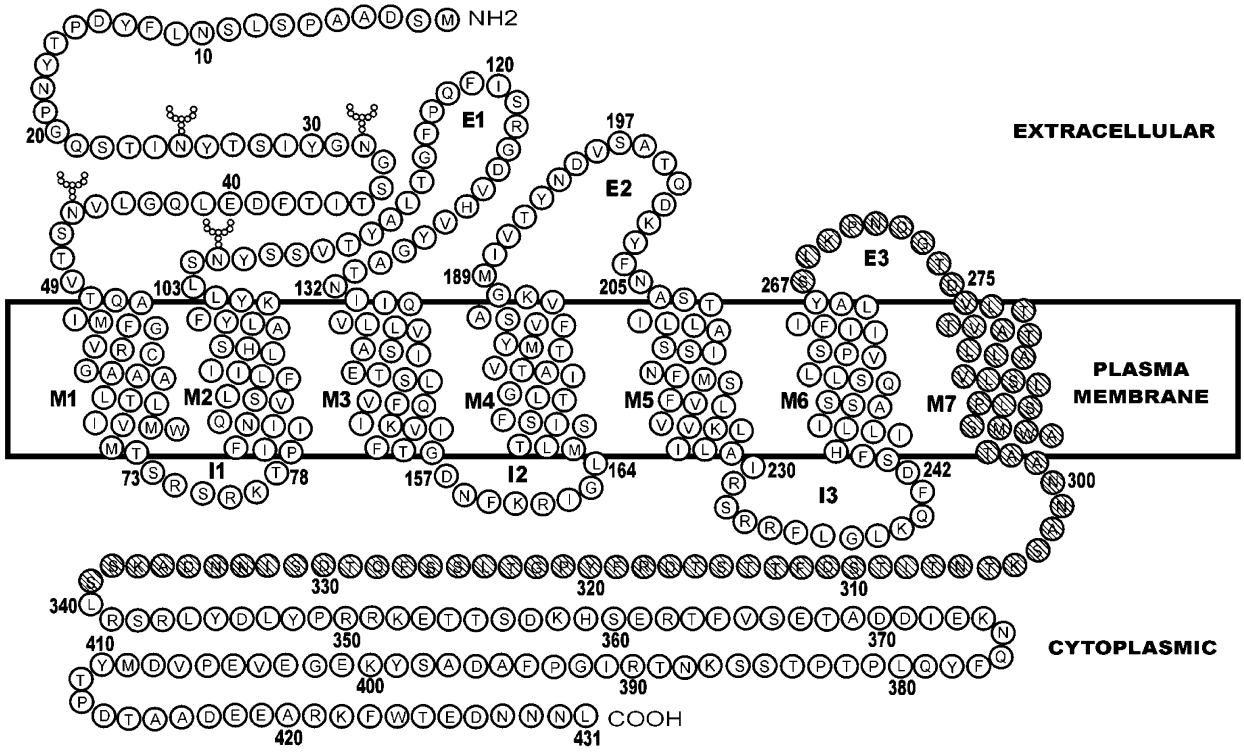

FIGURE 1: Cartoon of the α-factor receptor (Ste2p). Domains are indicated by the following: E, extracellular loops; I, intracellular loops; M, transmembrane domains. Four of the Asn residues are represented as glycosylation positions. The E3-M7-24-T40 region targeted in this study is indicated by hashed residues.

initiate signal transduction pathways ultimately leading to transcriptional activation of specific genes (*3*, *12*). Understanding the structure of GPCRs and the mechanism by which they activate downstream signaling can help to determine the cause of many pathological states and thereby result in possible treatments (*13*).

To gain a better understanding of receptor structure and ligand−receptor interactions, the yeast *Saccharomyces cerevisiae* has been used as a model system because of the abundant biochemical and genetic knowledge and tools that are available (*14*). The α-factor pheromone receptor (Ste2p) encoded by the *STE2* gene of *S. cerevisiae* (*14*, *15*) (Figure 1) is involved in *S. cerevisiae* mating and belongs to Class D (fungal pheromone receptors) of the GPCR family (*16*). Although this class is unique to fungal species, mutational analysis indicates that the α-factor receptor and other GPCRs are structurally and functionally similar (*17*−*19*). The transmembrane segments and the extracellular loops of Ste2p are thought to form a pocket that acts as the ligand-binding domain (*20*), whereas the intracellular receptor surface is known to be involved in G-protein recognition and activation (*21*). GPCR activation involves rearrangement of the intracellular receptor surface, caused by ligand-induced changes in the relative orientation of individual TM helices, that enables the G-protein to interact with previously inaccessible residues on the receptor protein (*21*).

Given the absence of biophysical data on conformational changes in GPCRs that occur upon activation by ligands, studies on structural aspects of Ste2p may provide an understanding of the mechanism of signal transduction by Ste2p and other GPCR family members. Since transmembrane domains 6 and 7 (M6 and M7) of GPCRs interact (*18*, *22*) and are likely involved in the stimulation of cellular responses (*20*), studies on the structure of M6 and M7 and, in particular, on M6 interacting with M7 would have significant ramifications on elucidating Ste2p function. Such studies have been stymied due to difficulties in determining the structural features of membrane proteins by X-ray crystallography (*5*). Moreover, the size of integral membrane proteins and the requirement of membrane-like vesicles for solubilization and stabilization of these proteins prohibit the direct determination of their structure by solution NMR at this time. To overcome difficulties in conducting X-ray and NMR studies on intact receptors, many scientists have begun studying fragments of these molecules to generate atomic level models of domains (*23*, *24*). Since the stabilization of secondary structures is highly dependent on local interactions, it is expected that the conformational features of the receptor domains and of the complementary amino acid sequences in the intact receptor will be similar. Experimental evidence that supports this approach is provided by the fact that reconstituted or coexpressed fragments of Ste2p, bacteriorhodopsin, and rhodopsin assemble into noncovalently linked domains maintaining functions of the full protein (*25*−*29*). Although there is some evidence that long-range effects may have an influence on local secondary structures (*30*), most studies support the idea that studies on GPCR fragments can provide insights into the structure of the intact molecule (*31*). The 3D structure of rhodopsin, calculated from the NMR structures of individual transmembrane regions, was in reasonable agreement with the crystal structure of the full GPCR (*32*−*34*). Biophysical studies on single domains of Ste2p (*35*−*38*) and NMR analysis of the seven transmembrane domains of this receptor indicated significant conformational diversity for these regions of the protein (*24*, *39*).

Table 1: Summary of the Names of the Various Constructs, the Amino Acid Sequence of the Expressed Proteins of the Ste2p Target Regions, and Their Molecular Weights

name of fusion protein	expression plasmid	amino acid sequence[a]	molecular weight (Da)[b]
TrpΔLE-E3-M7-24-T40 or M7FP (Met294Leu)	pREJ02M	TrpΔLE-WMSLKPNQGTDVLTTVATLLAVLSLPLSS**L**WATAANN ASKTNTITSDFTTSTDRFYPGTLSSFQTDSINNDAKSS	21 234
Trx·Tag thioredoxin-E3-M7-24-T40 (Met294Ala)	pKLS01	Trx·Tag thioredoxin-SLKPNQGTDVLTTVATLLAVLSLPLSS**A**WA TAANNASKTNTITSDFTTSTDRFYPGTLSSFQTDSINNDAKSS *LEHHHHHH*	26 022
Trx·Tag thioredoxin-E3-M7-24-T40 (Met294Leu)	pKLS01-1	Trx·Tag thioredoxin-SLKPNQGTDVLTTVATLLAVLSLPLSS**L**WA TAANNASKTNTITSDFTTSTDRFYPGTLSSFQTDSINNDAKSS *LEHHHHHH*	26 064
Trx·Tag thioredoxin-E3-M7-24-T40 (Met294Leu)	pKLS01-2	Trx·Tag thioredoxin-SLKPNQGTDVLTTVATLLAVLSLPLSS**L**WA TAANNASKTNTITSDFTTSTDRFYPGTLSSFQTDSINNDAKSS	24 999

[a] The solid underlined sequences correspond to E3-M7-24-T40 fragment of Ste2p. The LEHHHHHH sequence indicated in italic font is not in the native Ste2p sequence. Bolded residues were M in the native sequence. [b] Molecular weight of the fusion protein.

Virtually all of the previous biophysical investigations on fragments of GPCRs have been limited to fragments containing 30−40 residues, and even for these relatively short peptides, few high-resolution structures in detergent have been published. The reason for the dearth of information on longer regions of GPCRs and for the lack of high-resolution studies in micelles has been due largely to difficulties encountered in obtaining multimilligram quantities of isotopically labeled peptides containing one or more transmembrane domains. Recently, we chemically synthesized a 64-residue multidomain peptide containing the seventh transmembrane domain and 40 residues of the cytosolic terminus of Ste2p and showed that this peptide retained distinct structural characteristics for each of these domains in organic−aqueous medium and various detergents using circular dichroism analyses. Moreover, CD patterns in detergents were nearly identical at peptide concentrations from 50 to 500 μM indicating that NMR analyses were feasible (*40*).

Numerous synthetic attempts to increase the size of this fragment to include the third extracellular loop failed. Therefore, as part of our program to study fragments of Ste2p of increasing size, we turned to biosynthesis. In this communication, we discuss approaches to prepare a ^{15}N- and ^{13}C/^{15}N-labeled multiple domain 73-residue peptides from Ste2p in 10 mg quantities. For expression purposes, comparison was made of the TrpΔLE fusion protein expression system (*41*, *42*) and a thioredoxin fusion protein expression system that had been proposed as an efficient method to biosynthesize transmembrane domains (*43*). The biosynthetic peptide was examined using circular dichroism in organic−aqueous media, detergents, and vesicles. NMR analyses were carried out in trifluoroethanol/water (TFE/H_2O), chloroform/methanol/water ($CDCl_3$/CD_3OH/H_2O), and in detergents. High-quality NMR spectra were obtained for the 73-residue multidomain fragment of Ste2p in the presence of a large excess of detergent. Detailed NMR analyses in TFE/H_2O indicated that the transmembrane region within the 73-residue peptide was helical and that regions of the tail also had helical tendencies. Similar conclusions were reached from an NMR analysis in $CDCl_3$/CD_3OH/H_2O.

MATERIALS AND METHODS

Vectors. The parent plasmid pMD194 containing the wild-type *STE2* gene used to generate Ste2p mutant plasmid pMD602 (Met250Ala) was obtained from Mark Dumont [University of Rochester] (*25*). The plasmid pMMHa, which was used to construct plasmid pREJ02, was obtained as a gift from Dr. Peter Kim [Massachusetts Institute of Technology] (*44*). Vector pREJ02 was constructed as previously described (*45*). This vector utilizes a T7 expression system and expresses a polypeptide fused to a portion of the histidine-tagged TrpΔLE1413 polypeptide (*41*, *42*, *44*) at the N-terminus. The vector pREJ02 contained the DNA sequence of E3-M7-24-T40 [residues Ser267−Ser339 of Ste2p comprising the third extracellular loop (E3), the 24-residue seventh transmembrane domain (M7), and 40 residues of the cytoplasmic tail (T40)]. The pKLS01 plasmid was constructed by polymerase chain reaction amplifying the E3-M7-24-T40 region of the *STE2* gene from the pGA314.Cysless Ste2p.FTHT plasmid (*47*) and inserting it into the vector pET-32b(+) (Novagen, Madison, WI) between the *Nco* I and *Xho* I restriction sites. The resulting clone contained a Trx·Tag, His·Tag, and S·Tag followed by the E3-M7-24-T40 DNA sequence of Ste2p. The correct insertion was confirmed by DNA sequence analysis. The pET-32b(+) vector was designed for cloning and high-level expression of peptide sequences fused with the 109aa Trx·Tag thioredoxin protein (*47*).

The plasmids pREJ02M, pKLS01, pKLS01-1, and pKLS01-2 contained small sequence variations in the DNA sequence of E3-M7-24-T40 of Ste2p and mutations Met294Leu, Met294Ala, Met294Leu, Met294Leu, respectively (Table 1). All plasmids were created through site-directed mutagenesis as previously described (*48*).

Strains. The A232 yeast strain (MAT**a** *ste2*-Δcryl^R ade2-1 his4-580 lys2$_{oc}$ tyrl$_{oc}$ SUP4-3^{ts} leu2 ura3 bar1-1 FUS1::p[FUS1-lacZ TRP1]) transformed with the plasmid DNA harboring the wild-type and mutant receptors was obtained as a courtesy from Mark Dumont (*25*). *Escherichia coli* expression strains BL21(DE3)pLysS and BL21*trxB*(DE3) were purchased as competent cells from Promega (Madison, WI) and Novagen, respectively. *E. coli* DH5α cells used to amplify the engineered plasmids were purchased from Gibco BRL Life Technologies (Grand Island, NY).

Media, Buffers, and Solvents. M9 minimal medium was made as in ref *49*. The ^{15}N-labeled minimal medium was made by substituting ammonium chloride with ^{15}N-labeled ammonium chloride (Spectra Stable Isotopes, Columbia, MD) in M9 medium. The ^{15}N- and ^{13}C-labeled M9 minimal medium was made by substituting ammonium chloride with

^{15}N-labeled ammonium chloride and glucose with uniformly labeled ^{13}C D-glucose (Cambridge Isotope Laboratories, Andover, MA).

Lysis solution contained lysis buffer (50 mM Tris-HCl and 1 mM ethylenediaminetetracetic acid), 1 mM phenylmethylsulfonyl fluoride, and 300 μg/mL lysozyme, pH 8.7. Inclusion body washing buffer contained 1% igepal Ca-630 and 1% deoxycholic acid in lysis buffer. Guanidinium buffer was made of 6 M guanidinium hydrochloride (GuHCl), 10 mM Na_2PO_4, 138 mM NaCl, and 2.7 mM KCl, pH 7.4. Phosphate-buffered saline (1× PBS) contained 137 mM NaCl, 2.7 mM KCl, 10.1 mM Na_2PO_4, and 1.8 mM KH_2PO_4, pH 7.4.

Solvents used for HPLC purification such as acetonitrile and water were purchased from VWR International, Inc. (Bridgeport, NJ). TFA and TFE were purchased from Sigma (St. Louis, MO). Lipids used in CD such as DMPC and DMPG and detergents such as PPG and DPC were purchased from Avanti Polar Lipids (Alabaster, AL). Phosphate buffer contained 10 mM Na_2HPO_4 /NaH_2PO_4 and 0.02% NaN_3, pH 6.4. Deuterated solvents used in NMR such as d_2-TFE was purchased from Cambridge Isotope Laboratories (Andover, MA), and $CDCl_3$ and CD_3OH were purchased from Sigma (St. Louis, MO). Deuterated detergents used in NMR such as d_{38}-DPC or d_{22}-DHPC and d_{25}-SDS were purchased from Avanti Polar Lipids (Alabaster, AL) and Cambridge Isotope Laboratories (Andover, MA), respectively.

Cloning and Protein Expression. Cloning into pMMHa was performed following published procedures (*45, 50*). Cloning into pET-32b(+) vector was performed following the manufacturer's procedures (Novagen). Protein expression using the pREJ02M plasmid was based on a previously published expression procedure (*45*). Protein expression using pKLS01, pKLS01-1, and pKLS01-2 plasmids was based on the manufacturer's procedures (Novagen). Western blot was performed following published procedures (*51*). His-tag detection was carried out following the manufacturer's procedures (Pierce Biotechnology, Inc., Rockford, IL).

Comparison of the Expression of pKLS01, pKLS01-1, pKLS01-2, and pREJ02M. Two colonies of *E. coli* BL21-(DE3)pLysS transformed with pREJ02M and *E. coli* BL21*trxB*(DE3) transformed with pKLS01, pKLS01-1, or pKLS01-2 were inoculated into 5 mL LB medium with the appropriate antibiotics (ampicillin for both cell strains, kanamycin for the BL21*trxB*(DE3) strain, and chloramphenicol for the BL21(DE3)pLysS strain). The cells were then grown with shaking at 225 rpm and 37 °C to late log phase, OD_{600} of 0.6−1 and stored at 4 °C overnight. Two milliliters of each culture were harvested by centrifugation at 5000 rpm for 2 min and 14 000 rpm for 30 s, respectively. Cell pellets were resuspended into 50 mL cultures and incubated to late log phase, OD_{600} of 0.6−1, for 3−6 h. Cells were induced with IPTG (1 mM) for 3−6 h. One-milliliter aliquots were collected every hour for SDS−PAGE analysis of induction of protein expression. A 1 mL aliquot of the BL21(DE3)-pLysS containing pREJ02M and the BL21*trxB*(DE3) containing pKLS01, pKLS01-1, or pKLS01-2 were prepared using the inclusion body isolation protocol and the total cell protein fraction protocol, respectively.

Inclusion Body Isolation and Protein Purification. 1. Large Culture. The isolation of inclusion bodies was performed using a modification of a previously published procedure (*44*). All of the following steps were carried out at 4 °C. After growing and inducing the expression cells with IPTG in rich medium or either ^{15}N- or ^{13}C/^{15}N-labeled M9 mimimal medium, for expressing unlabeled, ^{15}N- or ^{13}C/^{15}N-labeled polypeptides, respectively, cells were harvested by centrifugation at 8000 rpm for 20 min. The wet cell pellet was weighed and then resuspended in lysis solution at a ratio of 5 mL of solution/gram of wet cells. Cell suspension was sonicated for 1−3 min and then the lysate was centrifuged at 18 000 rpm for 20 min. The resulting pellet was sequentially resuspended by sonication in 3 mL of lysis buffer, 3 mL of inclusion body washing buffer, and then 3 mL of distilled H_2O. After each resuspension and sonication, the resuspended pellet was centrifuged as specified above. Finally, the inclusion body pellet was weighed and resuspended by sonication in 4 mL of guanidinium buffer.

HPLC purification was carried out using a Hewlett-Packard/Agilent instruments equipped with a quaternary gradient solvent system. Chromatograms were monitored at 220 nm. Fusion protein purification was carried out using a Vydac 259VHP82215 preparative reversed-phase polymer column (22 mm × 150 mm; 8 μm; 300 Å) with a water jacket at 50 °C and a water−acetonitrile (0.1% TFA) gradient from 30−60% acetonitrile in 80 min at a flow rate of 4 mL/min. Cleaved peptides were purified using a Waters μBondpak preparative C18 reversed-phase column (19 mm × 300 mm; 10 μm; 125 Å) with a water jacket at 50 °C and a water−acetonitrile (0.1% TFA) gradient from 40−80% acetonitrile in 80 min at a flow rate of 4 mL/min. Fusion proteins and cleaved peptides were analyzed using a Vydac 259VHP54 reversed-phase polymer column (4.6 mm × 150 mm; 5 μm; 300 Å) and a Waters Delta Pak C18 reversed-phase column (3.9 mm × 150 mm; 5 μm; 100 Å), respectively, at 50 °C and water−acetonitrile (0.1% TFA) gradient from 30−60% acetonitrile in 20 min at a flow rate of 1 mL/min. All peptides were purified to over 98% homogeneity as judged by RP-HPLC. The final products were assessed by ESI-MS. MS measurements were performed at Hunter College, City University of New York.

2. Small Culture. The 1 mL aliquots of the TrpΔLE fusion proteins, which were collected for SDS−PAGE analysis, were centrifuged at 8000 rpm for 5 min, and the pellets were resuspended completely in 100 μL of lysis solution. Cell suspensions were sonicated for 20−30 s ,and then the lysate solution was centrifuged at 14 000 rpm for 5 min. The resulting pellet was resuspended by sonication in 200 μL of inclusion body washing buffer and then centrifuged as specified above. All of the above steps were perforemed at 4 °C. Finally, the inclusion body pellet was resuspended by sonication in 100 μL of 4× SDS sample buffer. The samples were stored at −20 °C and then boiled for 1−2 min just before running in SDS−PAGE.

Total Cell Protein Fraction. The 1 mL aliquots, which were collected for SDS−PAGE analysis of the induction of protein expression as 109aa Trx•Tag thioredoxin fusion proteins, were centrifuged at 8000 rpm for 5 min. The pellets were resuspended completely in 100 μL of 1× PBS to yield a concentration factor of 10×. Then 100 μL of 4× SDS sample buffer was added. The samples were sonicated and then boiled for 1−2 min and stored at −20 °C.

Protein Cleavage. Cyanogen bromide (CNBr) was used to release the 73-residue transmembrane peptide from the

N-terminal leader peptide (*51*, *52*). The lyophilized pure fusion protein was cleaved with 1M CNBr. Approximately a 500-fold molar excess of CNBr was added to the M7FP solution. The lyophilized pure M7FP fusion protein and the CNBr were dissolved in 50% TFA. The CNBr was dissolved separately in TFA and was then added to the fusion protein. The cleavage reaction of M7FP was carried out in the dark, at room temperature for 4.5 h. To stop the cleavage reaction, the reaction mixture was lyophilized. The lyophilized E3-M7-24-T40 peptide was resuspended in a mixture of 10% TFA, 40% acetonitrile, and 50% H_2O and then purified using a preparative C18 reverse-phase column.

CD Spectroscopy. For CD analysis of E3-M7-24-T40 in organic−aqueous solvents, micelles and vesicles solutions, stock solutions were first prepared in 95% TFE, and then aliquotes of the stock solutions were diluted with the appropriate amount of TFE and H_2O to obtain the final peptide concentration of interest (*40*). The final concentration of the peptides was determined by UV spectroscopy at 280 nm using an extinction coefficient of 6890 M^{-1} cm^{-1} (*40*). Samples in vesicle-containing media were made by adding solutions of the lipids in chloroform to the stock solution of the peptide and drying the resulting mixture under nitrogen, then resuspended in phosphate buffer, and sonicated for 60 min at 50 °C using a W-385 unit (Misonix, Inc., Farmingdale, NY) equipped with a 2.5 in. Cup Horn sonicator operated at 40% output power (∼200 W). Samples in detergent micelle-containing media were made by aliquoting the TFE solution into vials, adding water, lyophilizing, adding detergent in phosphate buffer, and sonicating for 15 min. The final solutions were either ∼50 or ∼400 μM in peptide, and the detergent/lipid-to-peptide ratios were between 100:1 and 200:1.

The circular dichroism (CD) spectra of the peptides were recorded on an AVIV model 62-DS CD instrument (AVIV Associates, Lakewood, NJ). Quartz cuvettes with path length of 1 and 0.2 mm were used for peptides in TFE solution and vesicles, respectively. In both cases, the peptide concentration was 50 μM. Cuvettes with path length of 0.2 and 0.1 mm were used for peptides in detergents (15 or 100 mM) with concentration of 50 and 400 μM, respectively. All spectra were the average of 3−5 scans between 280 and 185 nm at an interval of 1 nm with a 3−5 s integration time at each wavelength. The bandwidth for each measurement was set to 2 nm. CD spectra on blanks, corresponding to the different media except that no protein was dissolved, were collected and subtracted from the spectra containing the protein. Protein concentrations were obtained by UV spectroscopy. CD intensities are expressed as mean residue ellipticities (deg cm^2 $dmol^{-1}$). The percentage of α-helicity was calculated using the method of Wu et al. (*53*) and Chen et al. (*54*).

NMR Sample Preparation. 1. NMR Samples in TFE/H_2O or $CDCl_3$/CD_3OH/H_2O. [^{15}N]-E3-M7-24-T40 (1.9 mg) was first dissolved in 250 μL TFE-d_2, and then 250 μL H_2O (+0.1% TFA) was added to give 500 μL of TFE-d_2/H_2O (1:1 v/v) solution with peptide concentration of 488 μM. Similarly, [^{15}N]-E3-M7-24-T40 (1.9 mg) was first suspended in 222 μL of $CDCl_3$, then 222 μL of CD_3OH and, finally, 55 μL of H_2O (0.1% TFA) were added to give 500 μL of $CDCl_3$/CD_3OH/H_2O (4:4:1 v/v) solution with peptide concentration of 488 μM. The solution clarified upon addition of the aqueous component. Acidic water was used to aid in water suppression in the NMR spectra. We have found that, in the case of the $CDCl_3$/CD_3OH/H_2O solvent, sealing the NMR tube with the torch is essential to maintain the solvent composition during long 3D NMR experiments. The peptide concentrations were based on the measured weights of the samples.

2. NMR Samples in Detergents. 2.1. NMR Samples at Low Peptide Concentrations (0.17−0.2 mM) in SDS, DPC, DHPC, and PPG at High Detergent/Peptide Ratios (∼450:*1).* [^{15}N]-E3-M7-24-T40 (1.6 mg) was dissolved in TFE/H_2O (1:1), then divided into three vials and lyophilized. To each vial (containing 0.53 mg of [^{15}N]-E3-M7-24-T40) was added 350 μL of d_{25}-SDS, d_{38}-DPC, or d_{22}-DHPC (86 mM) detergents in H_2O/D_2O (9:1), containing 0.02% NaN_3; the pH was adjusted to 4.0 using TFA to yield a final peptide concentration of 194 μM and a peptide/detergent ratio of 1:440. In a separate experiment, 350 μL of PPG (86 mM) in H_2O/D_2O (9:1), containing 0.02% NaN_3, pH 4, was added to [^{15}N]-E3-M7-24-T40 (0.46 mg) to yield a peptide concentration of 170 μM and a peptide/detergent ratio of 1:500. The peptide/detergent mixtures were sonicated at 50 °C for 10 min and then transferred to an NMR Shigemi tube (Shigemi, Inc., Allison Park, PA).

2.2 NMR Sample at a High Peptide Concentration (0.5 mM) in PPG at High Detergent/Peptide Ratio (∼400:1). A total of 1.14 mg of [^{15}N]-E3-M7-24-T40 was dissolved in 350 μL of 200 mM d_{38}-DPC (27.3 mg) in H_2O/D_2O (9:1) containing 0.02% NaN_3; the pH was adjusted to 4.0 using TFA to yield a peptide concentration of 475 μM and a peptide/detergent ratio of 1:420. The peptide−detergent mixture was sonicated at 50 °C for 10 min and then transferred to an NMR Shigemi tube.

NMR Spectroscopy. ^{1}H NMR spectra were recorded on a three-channel Varian UNITY INOVA 600 MHz NMR spectrometer (Varian NMR Instrument, Palo Alto, CA) equipped with a *z*-axis pulsed-field-gradient and a Varian 5-mm ^{1}H/^{13}C/^{15}N triple resonance probe. In the case of the [^{15}N]-E3-M7-24-T40 peptide in TFE-d_2/H_2O (1:1 v/v), $CDCl_3$/CD_2OH/H_2O (4:4:1 v/v), d_{25}-SDS, d_{38}-DPC, d_{22}-DHPC, and PPG, heteronuclear single quantum correlation (HSQC), HSQC-nuclear Overhauser effect spectroscopy (NOESY), and HSQC-total correlation spectroscopy (TOCSY) experiments (*55*, *56*) were used. Identification of specific amino acid signals was made using three-dimensional HSQC-TOCSY spectra with contact times of 70 ms to identify spin systems. Sequential assignments of amino acids were determined using three-dimensional HSQC−NOESY spectra. Two-dimensional TOCSY and NOESY spectra were also recorded and were helpful in identifying cross-peaks or connections belonging to proline residues in 3D NOESY. All NOESY experiments were performed with a 300 ms mixing time, and the results were used in modeling calculations.

Structure Refinement. Molecular modeling was performed using the DYANA program (*57*), and all calculations were performed on a Linux workstation. All necessary file conversions from the Varian format and further data processing to the NMRView format were performed using the NMRPipe software (*58*). NOESY cross-peak assignments, integration of cross-peak volumes, and creation of distance restriction files in the DYANA input format were aided by the NMRView program (*59*). The upper NOE distance

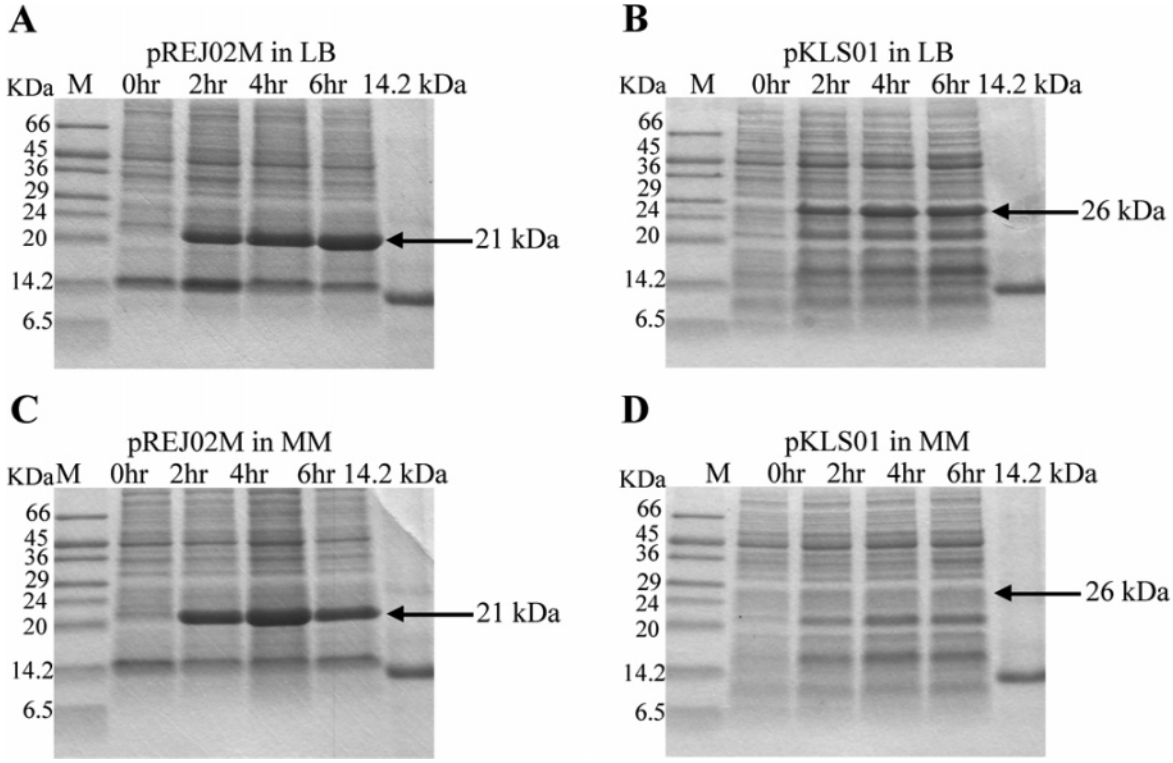

FIGURE 2: Comparison of the expression of E3-M7-24-T40 as TrpΔLE and thioredoxin fusion proteins. SDS−PAGE of E3-M7-24-T40 fusion proteins biosynthesized in rich (LB) or minimal medium (MM). Cells were induced with 1 mM IPTG in LB or M9 MM for up to 6 h, and the fraction containing the target fusion protein was isolated as per Materials and Methods. TrpΔLE−E3-M7-24-T40 (Met294Leu) (A and C) and Trx•Tag thioredoxin−E3-M7-24-T40 (Met294Ala) (B and D) have molecular weights of 21 and 26 kDa, respectively. SDS−PAGE gels were stained with Coomassie dye. Panel A was taken from ref *45*.

constraints were calculated by the NMRView program using the median peak intensity calibration method; all lower distance constraints were set to 1.8 Å. Dihedral angle constraints were generated for Φ torsion angles using NH−CHα coupling constants. The DYANA modeling starts by generating 50−100 molecules with random conformations. We used the standard DYANA protocol that includes simulated annealing for every initial random conformation molecule with 4000 torsion angle dynamics steps. The first 800 of these steps were performed at an initial high temperature that is software-determined, followed by slow-cooling of the molecule. One-half of the refined structures representing those molecules in the ensemble with the lowest target function (lowest sum of NMR distance constraint violations) was selected for generating the output protein database files and for further analysis, including rsmd calculation and visualization by the MOLMOL program (*60*).

RESULTS

Biosynthesis, Isolation, and Purification of Fusion Proteins Corresponding to Domains of Ste2p. An important requirement of NMR studies on peptides corresponding to domains of GPCRs is the availability of milligram quantities of isotopically labeled molecules. To achieve this goal, we constructed a plasmid coding for a fusion protein containing a 73 amino acid residue multidomain peptide (E3-M7-24-T40) consisting of the third extracellular loop (267−275), transmembrane domain 7 (276−299), and 40 amino acid residues of the carboxyl terminus (300−339) of Ste2p (Figure 1). Since the carboxyl terminus of Ste2p is highly hydrophilic, these 40 residues were expected to increase the solubility of the M7 domain (*40*). A previous study found that a construct coding for expression of transmembrane seven and nearly all of the cytosolic tail had no biological activity (*25*). On the basis of this precedent, it is unlikely that E3-M7-24-T40 would be biologically active. The peptide was expressed as a fusion protein using histidine-tagged TrpΔLE as the N-terminus. The TrpΔLE directs the production of the fusion protein into inclusion bodies, which protects the cell from potential protein toxicity and facilitates isolation of a fusion protein. The peptide was fused to the C-terminus of TrpΔLE through a methionine so that it could be released by cyanogen bromide cleavage. This necessitated the mutation of Met294Leu in the M7 domain. This mutation did not impair the function of Ste2p (*61*).

Expression of unlabeled and isotopically labeled fusion proteins was induced with 1mM IPTG for up to 6 h in LB-rich medium and M9 minimal medium which contained ^{13}C-labeled glucose and/or ^{15}N-labeled ammonium chloride. Inclusion bodies expressed from the induced cells were isolated, and the expression of [^{15}N]-M7FP or [^{13}C/^{15}N]-M7FP in M9 minimal medium (data not shown) and unlabeled peptide in rich medium (Figure 2A) and minimal medium (Figure 2C) were analyzed via SDS−PAGE. Western blotting using a hexa-histidine probe confirmed the presence of the fusion proteins (data not shown). The M7FP fusion protein containing the E3-M7-24-T40 region of Ste2p exhibited excellent expression (Figure 2A,C). Maximal expression was observed 4−6 h after induction.

A thioredoxin fusion system has been used in the biosynthesis of transmembrane domains of the CFTR protein (*43*). To evaluate the efficiency of the thioredoxin and TrpΔLE

expression systems, we compared the biosynthesis of M7FP by the pREJ02M and pKLS01 plasmids. pREJ02M encodes the E3-M7-24-T40 region of Ste2p with the Met294Leu mutation, fused with the TrpΔLE peptide (TrpΔLE−E3-M7-24-T40 or M7FP), a 21 kDa peptide, whereas the pKLS01 plasmid encodes the E3-M7-24-T40 region of Ste2p with a Met294Ala mutation and an additional eight amino acids at the C-terminus of the peptide fused with the Trx·Tag thioredoxin (Trx·Tag thioredoxin−E3-M7-24-T40), a 26 kDa peptide (Table 1). Both expression systems were tested in rich and minimal media. Expression of the TrpΔLE−E3-M7-24-T40 fusion protein driven by pREJ02M was significantly higher than the expression of the Trx·Tag thioredoxin−E3-M7-24-T40 fusion protein driven by pKLS01 in rich medium (Figure 2A,B). Moreover, the TrpΔLE fusion protein represented a much higher fraction of the proteins as judged by SDS−PAGE chromatography. Strikingly, the expression of M7FP by pREJ02M remained high in minimal medium (Figure 2C), whereas no expression of fusion protein was driven by pKLS01 in minimal medium (Figure 2D). Two other plasmids, pKLS01-1 and pKLS01-2 coding for variations of the E3-M7-24-T40 domain, one virtually identical to that coded for by pREJ02M (Table 1), were tested and gave similar results to that of pKLS01 (data not shown). On the basis of this observation, in large scale expressions, the system that expressed proteins fused to a portion of the TrpΔLE1413 polypeptide (*41*, *42*, *44*) was used to prepare the M7 region of Ste2p.

After the expression and isolation of the M7FP inclusion bodies, the final pellet was resuspended by sonication in 6M GuHCl and then purified via HPLC (see Materials and Methods). The crude inclusion bodies were highly enriched in the target peptide that was readily isolated to yield a pure fusion protein of greater than 95% homogeneity as judged by reversed phase HPLC, SDS−PAGE, and mass spectrometry (data not shown). Biosynthesis of unlabeled, ^{15}N-labeled, and/or ^{13}C/^{15}N-labeled M7FP yielded about 100 mg and 30−35 mg (>95% homogeneous) of fusion protein/liter of bacterial culture in rich and minimal media, respectively. The molecular weight of all fusion proteins was measured by ESI-MS and determined to be consistent with the calculated values (21 234 kDa calculated and 21 235 kDa experimental for M7FP).

Release of Membrane Peptides Using CNBr Cleavage. To optimize release of the 73-residue peptide from M7FP, the time course of cleavage by CNBr was followed using HPLC. Optimal cleavage of a single transmembrane domain from the fusion protein had been previously found to occur with 1 M CNBr in 70% TFA for 24 h (*52*). For M7FP, maximal yields occurred at about 4.5 h in both 50% TFA and 70% TFA (Figure 3). Similar results were obtained using sublimed CNBr or newly purchased reagent. Interestingly, the HPLC monitoring indicated that maximum product was obtained for E3-M7-24-T40 at about 4.5 h despite the fact that significant fusion protein remained at this time. Further incubation apparently resulted in degradation of the product. We were able to isolate about 4 mg/liter of ^{15}N-labeled E3-M7-24-T40, 2 mg/liter of ^{13}C/^{15}N-labeled E3-M7-24-T40, and nearly 15 mg/liter of this unlabeled peptide from rich medium (Table 2). All final peptides were ≥95% homogeneous as judged by analytical HPLC and had molecular weights consistent with expected values (Table 2, Figure 3C).

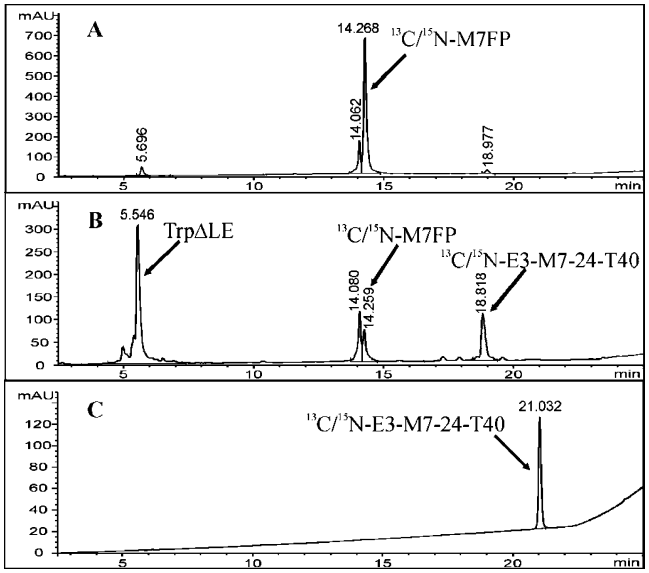

FIGURE 3: HPLC monitoring of the CNBr cleavage products from the [^{13}C/^{15}N]-M7FP fusion protein. The [^{13}C/^{15}N]-M7FP fusion protein was cleaved with 1 M CNBr in 50% TFA for 4.5 h in the dark at room temperature. The cleavage reaction was analyzed at 0 h (A) and 4.5 h (B) by HPLC as described in Materials and Methods. After 4.5 h, the cleavage was stopped as described in Materials and Methods, the cleaved [^{13}C/^{15}N]-E3-M7−24-T40 peptide was purified via HPLC, and fractions were analyzed on an analytical C18 reversed-phase column. (C) Purified [^{13}C/^{15}N]-E3-M7-24-T40 peptide is observed at a retention time of around 21 min. Retention time of the pure [^{13}C/^{15}N]-E3-M7-24-T40 peptide in C is different from A and B because this sample was run on a different HPLC instrument.

Table 2: Summary of Recoveries and Molecular Weights of Peptides from Purification after Release from TrpΔLE Fusion Proteins via CNBr Cleavage

peptides	recovery per liter of fermentation (mg)	calculated molecular weight (Da)	experimental molecular weight (Da)[c]
^{15}N-labeled E3-M7-24-T40	3.9[a]	7759	7757
unlabeled E3-M7-24-T40	15[b]	7671	7671
^{13}C/^{15}N-labeled E3-M7-24-T40	2.0[a]	8093	8079

[a] Fermentation in minimal media (see Materials and Methods). [b] Fermentation in rich media (see Materials and Methods). [c] Molecular weights were determined by ESI-MS.

Biophysical Studies. 1. Circular Dichroism. Previously, we had conducted detailed CD studies on M7-24-T40, a synthetic 64-residue peptide that contained the seventh transmembrane domain and 40 residues of the cytosolic loop of Ste2p (276−339) in organic−aqueous media and various detergents (*40*). CD studies on E3-M7-24-T40[Ste2p(267−339)] in TFE/H_2O mixtures, SDS, DPC, and PPG micelles and DMPC/DMPG (4:1) vesicles provided data with significant similarities to the slightly smaller homologue. Both peptides assumed a partially helical structure in many of the above media (data not shown). Most significantly, at peptide concentrations near 0.5 mM, the 73-residue peptide maintained its helicity in organic−aqueous media and detergent solutions. This suggested that high-resolution NMR studies on this GPCR fragment could be carried out.

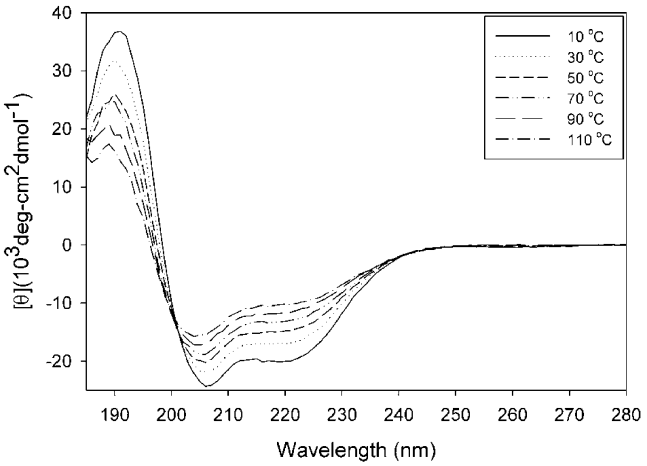

FIGURE 4: CD of E3-M7-24-T40 in TFE/H_2O (1:1) at different temperatures (10−110 °C). The peptide concentration was 20 μM.

The CD of E3-M7-24-T40 in TFE/H_2O (1:1) was measured at various temperatures (Figure 4). Over the range 10−110 °C, a significant decrease in the ellipticity at 208 and 222 nm was observed. However, even at 110 °C, this peptide exhibited a discernible $n \rightarrow \pi^*$ transition at 222 nm, and a $\pi \rightarrow \pi^*$ minimum that had shifted from 208 nm at 10 °C to 203 nm at 110 °C was observed. The ellipticity at 222 nm at 110 °C (−10 070 deg cm^2 dmol^{-1}) indicated 23% helix as compared to 43% measured at 25 °C (Figure 4).

2. NMR Spectroscopy. To evaluate the feasibility of three-dimensional NMR studies on E3-M7-24-T40 in membrane mimetic environments, ^{1}H−^{15}N HSQC spectra were measured in TFE/H_2O (1:1), $CDCl_3$/CD_3OH/H_2O (4:4:1), d_{25}-SDS, d_{38}-DPC, d_{22}-DHPC, and PPG. In both organic−aqueous solvents, HSQC spectra measured on 0.5 mM concentrations of [^{15}N]-E3-M7-24-T40 at 25 °C were well-resolved (Figure 5). In these media, out of 69 expected cross-peaks corresponding to the backbone NHs (73 minus amine terminal serine and three prolines) between 65 and 70 cross-peaks were observed. The expected number of doublets for the side-chain amides of Asn and Gln were also observed.

To screen the detergents for NMR experiments, a set of HSQC spectra was measured using low peptide concentration (∼0.2 mM) in ∼80−90 mM d_{25}-SDS, d_{38}-DPC, d_{22}-DHPC, or PPG at 50 °C (Figure 6A−D). At such high (400−500-fold) detergent excess 67, 64, 65, and 62 cross-peaks were observed for these d_{25}-SDS, d_{22}-DHPC, d_{38}-DPC, and PPG samples, respectively. Using significantly lower detergent excess (∼20:1) we observed very poorly resolved HSQC spectra which were only slightly improved when measured using a TROSY−HSQC pulse sequence (*62*) in place of the standard HSQC pulse sequence (data not shown).

The pattern of HSQC cross-peaks was slightly different in each detergent preparation using high lipid−peptide ratios and 0.2 mM peptide. As judged by HSQC NMR spectra, preparations of E3-M7-24-T40 in SDS and DPC were found to be more stable with time, whereas those measured in PPG and DHPC were less stable.

An attempt was made to increase peptide concentration in the detergent preparations to 0.5 mM to make 3D NMR measurements possible using our 600 MHz instrument. DPC was chosen based on the sample stability observed at 0.2 mM peptide concentration and to avoid using SDS which often denatures proteins and peptides (*63*). The HSQC spectrum obtained for a fresh solution of 0.5 mM 73-residue peptide in 200 mM DPC (Figure 6E) was similar in resolution to that of the 0.2 mM peptide in the same detergent (Figure 6B). However, the 0.5 mM sample was not stable and became turbid and gel-like after one week at 50 °C. The HSQC spectrum of this sample lost about 40% of peak intensity, some peaks broadened and some extra peaks appeared in the spectrum. Similar instability of other peptides in lipid-like environments also has been reported (*64*). In contrast, the lower concentration sample in DPC was stable for several weeks at the same temperature.

Analysis of E3-M7-24-T40 in Organic−Aqueous Media. ^{15}N-edited TOCSY and NOESY spectra were used to make complete assignments of backbone and side-chain protons of the ^{15}N-labeled E3-M7-24-T40 peptide in TFE/H_2O (1:1) and $CDCl_3$/CD_3OH/H_2O (4:4:1) at 25 °C (for chemical shifts in TFE/H_2O and chloroform/methanol/water see Supporting Information).

In TFE/H_2O, most residues in the peptide exhibited sequential NH−NH (d_{NN}) cross-peaks with a notable break at residues 22−24. Blocks of long-range CHα−NH [$d_{\alpha N}(i,i+3)$] cross-peaks were observed for residues 7−19, 38−52, and 60−70 (Figure 7B), and these regions also exhibited numerous $d_{\alpha N}(i,i+2)$ and d_{NN} $(i,i+2)$ cross-peaks. Several $d_{\alpha N}(i,i+4)$ could be detected for residues in these same regions of E3-M7-24-T40. The long-range NOE

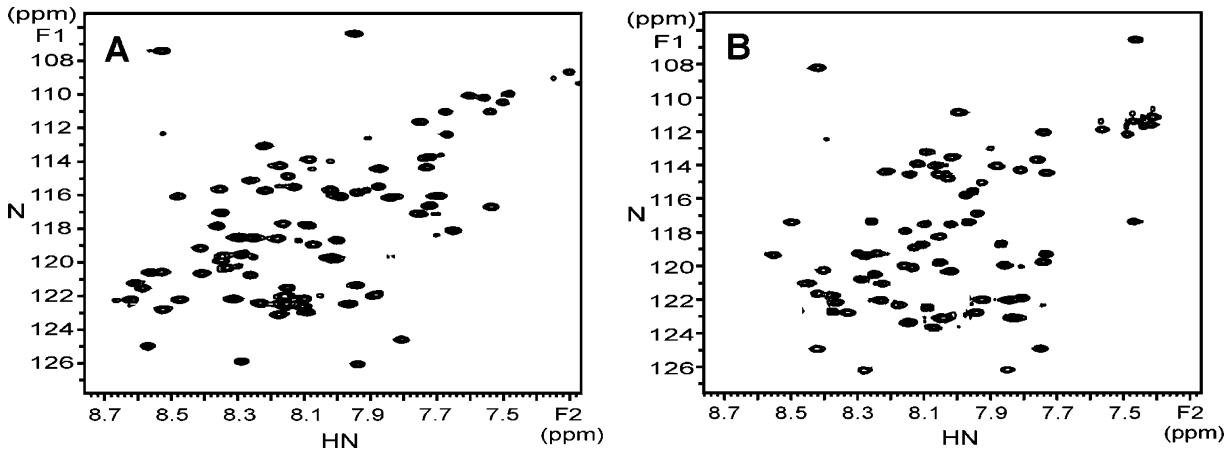

FIGURE 5: ^{1}H-^{15}N HSQC spectra of [^{15}N]-E3-M7-24-T40 in organic−aqueous media. (A) $CDCl_3$/CD_3OH/H_2O (4:4:1); (B) TFE/H_2O (1:1). Backbone NH portion of the spectra are shown. The peptide concentration was 0.5 mM.

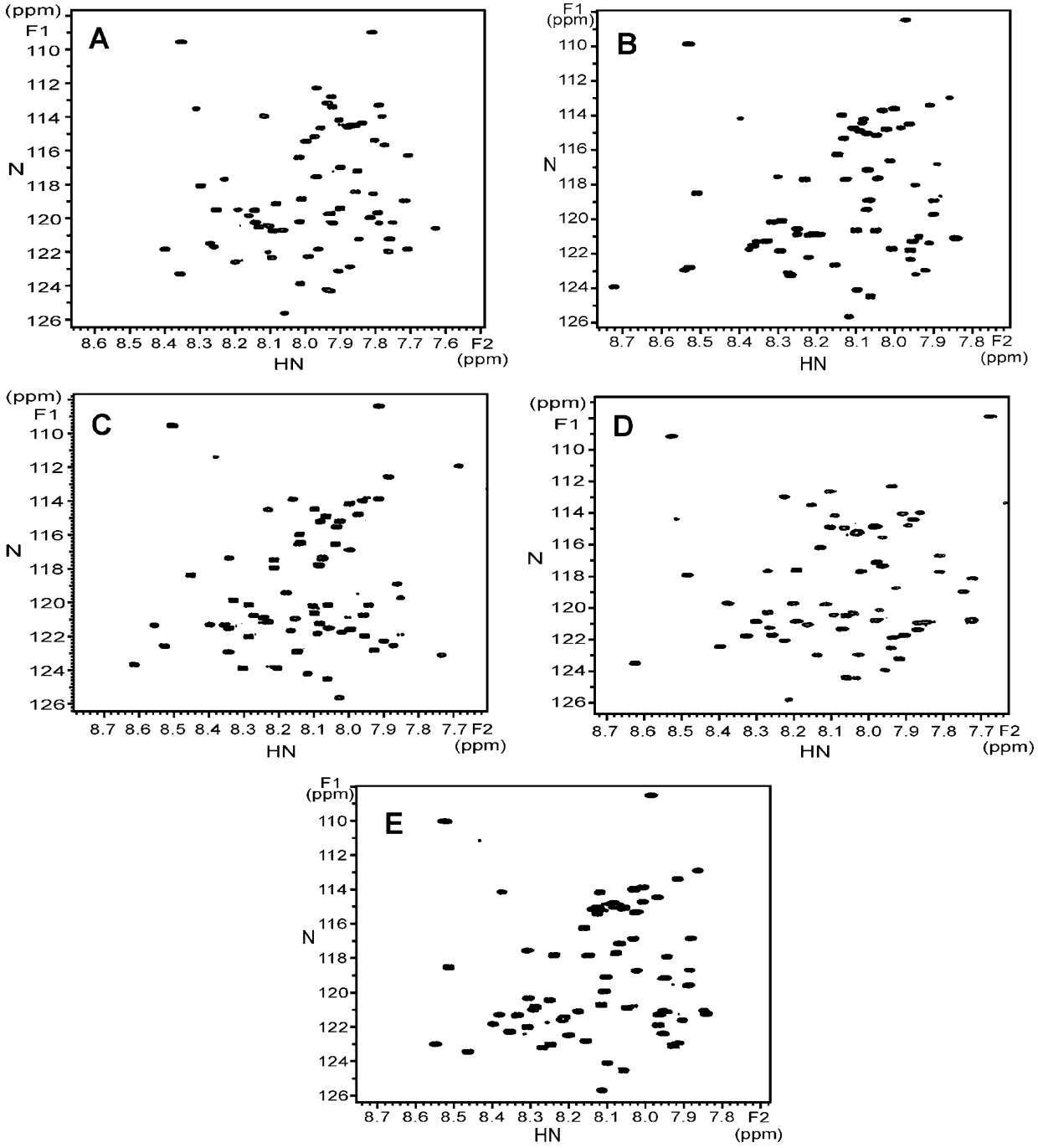

FIGURE 6: ^{1}H-^{15}N HSQC spectra of [^{15}N]-E3-M7−24-T40 in detergents at high detergent/peptide ratios (A) d25-SDS micelles, (B) d38-DPC micelles, (C) d22-DHPC micelles and (D) PPG micelles. The peptide-to-detergent ratio was 1:440 in A, B, & C and 1:500 in D. The peptide concentration was approximately 0.2 mM. (E) 0.5 mM [^{15}N]-E3-M7−24-T40 in 200 mM d38-DPC micelles (1:400 peptide-to-detergent ratio).

connectivity pattern in $CDCl_3/CD_3OH/H_2O$ (Figure 7A) was similar to the one observed in TFE/H_2O (Figure 7B). Low coupling constants ($^{3}J_{HN\alpha}$) from 3 to 5 Hz were measured for 6−23 and 25−43 residues, and most residues between 56 and 65 had $^{3}J_{HN\alpha}$ values below 6 Hz in TFE/H_2O (Figure 7C). The NOE connectivities and low coupling constants support the presence of three helical segments in the 73-residue peptide.

The δCHα chemical shifts for E3-M7-24-T40 in organic−aqueous media were compared to the δCHα random coil values (*65*, *66*) and plotted versus position in the peptide (Figure 8A). Using a deviation of more than ±0.1 ppm from the random coil value as the criterion for possibility of regular secondary structure (*65*), residues 8−22, 25−52, and 59−70 have helical tendencies in both solvents. If one considers the magnitude of the deviation as indicative of the stability of the helix, domain 8−22 forms the most stable helix followed by less stable 25−52 and 59−70. This conclusion is in a good agreement with the NOE long range connectivities and the coupling constant analysis.

Determination of the NMR Structure of E3-M7-24-T40 in TFE/H_2O and $CDCl_3/CD_3OH/H_2O$. A model of the structure of E3-M7-24-T40 in TFE/H_2O (1:1) at 25 °C was determined using a combination of NOE and dihedral angle constraints following the procedures described in Materials and Methods. A total of 547 NOE distance constraints (169 intraresidue,

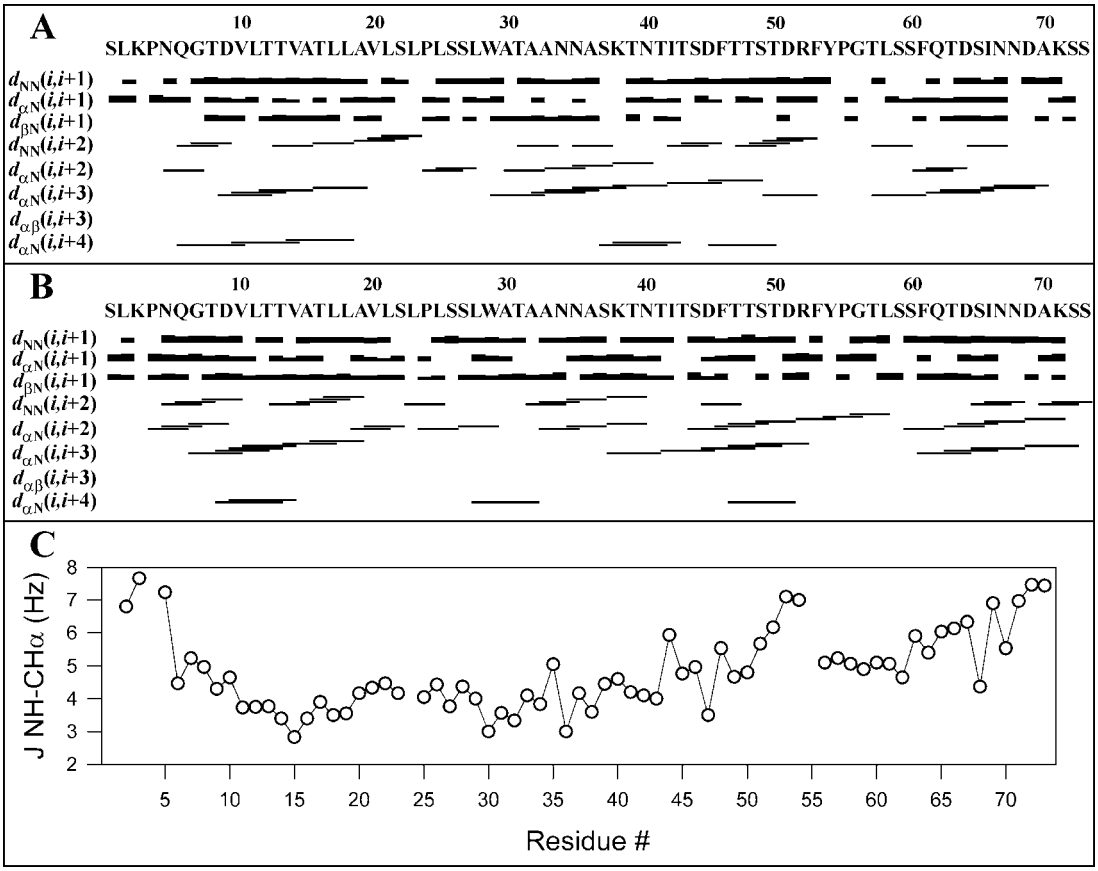

FIGURE 7: Summary of connectivities and coupling constants for E3-M7-24-T40 in $CDCl_3/CD_3OH/H_2O$ (4:4:1) and TFE/H_2O (1:1). (A) Sequential and medium-range NOEs in $CDCl_3/CD_3OH/H_2O$ (4:4:1). (B) Sequential and medium-range NOEs in TFE/H_2O (1:1). (C) *J*-coupling values of NH to CHα ($^3J_{HN\alpha}$) in TFE/H_2O (1:1).

243 sequential, and 135 long range) and 60 Φ dihedral angle constraints were input into the structure calculation for the peptide in TFE/H_2O. Modeling of the peptide structure in $CDCl_3/CD_3OH/H_2O$ was done using only NOE constraints. A total of 488 NOE distance constraints (157 intraresidues, 237 sequential, and 94 long range) were input into the calculation for the peptide in $CDCl_3/CD_3OH/H_2O$. The model of E3-M7-24-T40 in TFE/H_2O indicated three helical domains comprising residues 5−24, 30−52, and 55−70. When the 10 structures with the minimum NOE distance violations were overlaid separately for each of these helical domains, region I (residues 5−24) had a backbone rmsd of 1.9 Å, region II (residues 30−52) had a backbone rmsd of 1.86 Å, and region III (residues 55−70) had a backbone rmsd of 1.5 Å (Figure 9). The backbone rmsd calculated for the entire peptide was very large (∼10 Å).

Using the same approaches, the model obtained for the structure of E3-M7-24-T40 dissolved in $CDCl_3/CD_3OH/H_2O$ (Supporting Information Figure SI-3) exhibited features similar to those found in TFE/H_2O having helical residues near the N-terminus and in the cytosolic tail. However, in general, the cytoplasmic section of the peptide was less ordered and the rmsd values for the two helical regions (30−52 and 55−70) were >4 Å and >3 Å, respectively.

DISCUSSION

The determination of high-resolution structures of membrane proteins remains an elusive goal. One approach to realize this goal utilizes fragments corresponding to different regions of a protein combined with spectroscopic analysis in membrane mimetic solvents. It is possible to question the use of peptide fragments to learn about the structural tendencies of an intact protein. Nevertheless, extensive analysis has indicated that localized interactions often determine the secondary structure of domains of proteins. Moreover, in the case of bacteriorhodopsin (*28*) and rhodopsin (*29*), the structures of peptides corresponding to single transmembrane domains and loops were superimposed to assemble the entire protein and the good agreement obtained with crystal structures of the intact proteins would seem to validate the approach. Finally, both in vitro and in vivo reconstitution studies show that biologically inactive fragments of a GPCR reassemble to form biologically active receptor (*25*−*27*). Together, the above results suggest that, although one must be cautious in making absolute conclusions, it is reasonable to assume that structural information on a peptide fragment (E3-M7-24-T40) corresponding to nearly one-sixth of the residues of Ste2p will be relevant to the intact GPCR.

Although CD has been extensively used in evaluations of membrane peptide structure, NMR is the preferred method to obtain high-resolution information at the residue level. To carry out such NMR investigations on larger fragments of membrane peptides, especially in the presence of detergents, efficient methods to obtain isotopically labeled peptides are required and conditions must be optimized for measurement of well-resolved NMR spectra under membrane mimetic conditions.

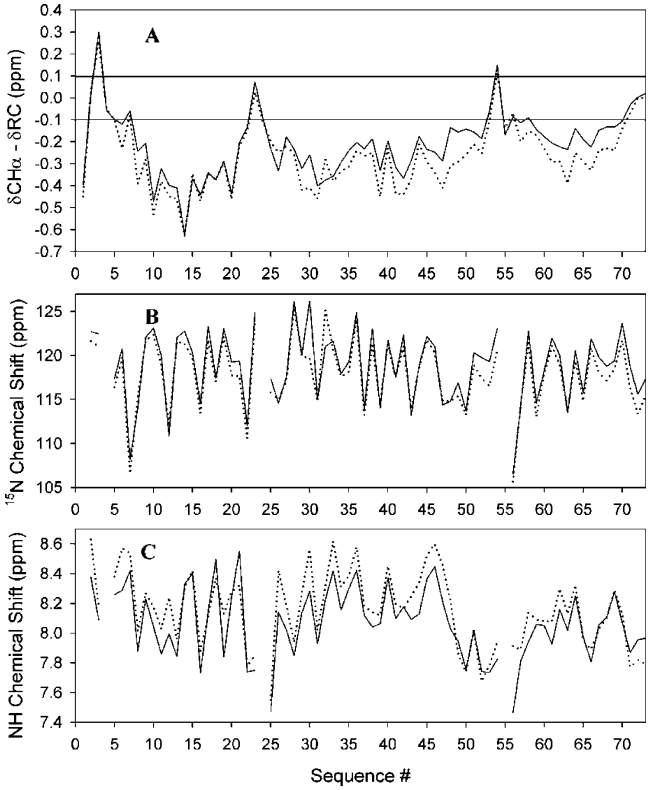

FIGURE 8: Comparison of δCHα values, nitrogen chemical shifts [^{15}N], and amide NH chemical shifts for the E3-M7-24-T40 peptide at 25 °C in both TFE/H_2O [1:1] (represented by a solid line) and $CDCl_3$/CD_3OH/H_2O [4:4:1] (represented by a dotted line). (A) The δCHα values plotted are the difference between the experimental CHα chemical shifts and the random coil (RC) values as per ref *66*. (B) ^{15}N chemical shifts and (C) amide NH chemical shifts. The lines at 0.1 and −0.1 in panel A are the cutoffs used in the determination of the chemical shift index (*66*).

In this communication, we report the successful biosynthesis and isolation of 100 mg quantities of isotopically labeled fusion protein containing a 73-residue multidomain peptide of a GPCR with more than 95% homogeneity. The fusion protein was processed by CNBr to release the receptor domain (Figure 3B), which could then be isolated by reversed phase chromatography (Figure 3C). The peptide, isolated in 10 mg quantities sufficient for a complete structural analysis using NMR spectroscopy, corresponded to part of the cytosolic tail, a transmembrane domain, and an extracellular loop (E3-M7-24-T40) and was labeled with either ^{15}N or ^{13}C/^{15}N.

The biosynthesis of membrane peptides is complicated by their poor water solubility, tendency to aggregate, and toxicity to the host cell. Few reports have appeared for preparation of membrane peptides containing more than 30−40 residues. In one elegant study, peptides corresponding to two predicted transmembrane domains of the cystic fibrosis transmembrane conductance regulator were expressed, released from the thioredoxin carrier, and purified in milligram quantities (*43*). In the present report, we compared the expression levels of E3-M7-24-T40 as driven by plasmids coding for TrpΔLE or thioredoxin fusion proteins. A plasmid coding for the TrpΔLE carrier was originally used to express bovine pancreatic trypsin inhibitor and very recently was used to express a single transmembrane domain of the CB2 receptor (*44*, *67*). We observed that the TrpΔLE system gave higher expression of Ste2p fusions than the thioredoxin system. This was true when three different thioredoxin constructs were prepared, one coding for additional residues at the C-terminus and the other engineered to code for the identical peptide fused to TrpΔLE. Both the level of expression and the percentage of the target protein appeared higher using the TrpΔLE vector (pREJ02M) as compared with pKLS01 and its variants. The expression of the presumed target protein by the thioredoxin system virtually disappeared in minimal medium, whereas in minimal medium, good-to-excellent expression of the Ste2p fusions were obtained using the pREJ vectors.

An expression system using GB-1 fusion proteins in *E. coli* has also been used to biosynthesize a 26-residue peptide predicted to be a transmembrane region of the human sodium proton exchanger, and loss of expression of GB-1 fusion protein constructs was also noted in minimal medium (*68*). This difficulty was overcome by adding about 5% isotopically labeled rich medium to the culture. Our experiments indicate that some optimization will be required in determining which system to choose for the expression of membrane peptides and that the thioredoxin system cannot be considered to be a general vector for all membrane peptides as suggested previously (*43*). Since the expression levels of E3-M7-24-T40 fusions were consistently higher than those for the M6 fusions (unpublished results) it is likely that none of the expression systems will prove optimal for all membrane peptides.

The release of the membrane peptides from the TrpΔLE fusion proteins using CNBr was reasonably efficient. Interestingly, different conditions proved optimal for a single TM domain fusion peptide (*52*) and the multidomain peptide we examined. As judged by HPLC monitoring, maximal yields of E3-M7-24-T40 were obtained after 3.5−4.5 h of cleavage, whereas release of the single transmembrane (M6) domain peptides required 24 h (*52*). When the cleavage of E3-M7-24-T40 proceeded for 24 h, very little product was isolated. In contrast, when the fusion protein containing the M6 domain was cleaved for only 3 h, most of the fusion protein remained unprocessed. Thus, it is advisable to optimize the CNBr cleavage conditions when one seeks to produce isotopically labeled product. We found that the percent acid (50% or 70%) used for the CNBr cleavage had a small influence on the cleavage rate but did not affect the ultimate yield. Similarly, we did not see significant improvements with sublimed CNBr. We do not sublime this reagent when newly purchased material is available because of the inherent danger of putting CNBr into the vapor phase. Overall, we obtained 30−42% yields of 73-residue peptide with >95% homogeneity for the combination of the cleavage and purification steps. When fractions containing product of lower homogeneity were taken into account, the yield was between 50% and 60%. Mass spectrometry analysis confirmed the molecular weight of the product. In the case of isotopically labeled E3-M7-24-T40, percent labeling for the ^{15}N and ^{13}C/^{15}N isomers was calculated to be 98% and 96.7%, respectively. The above experiments indicated that approximately 1−3 L of culture sufficed to obtain 10 mg quantities of labeled E3-M7-24-T40 peptide. Thus, to obtain ^{15}N-labeled product required between $100 and $200 of ^{15}N-labeled starting material. In the case of ^{13}C/^{15}N double labeling, we found that approximately a 25% decrease in

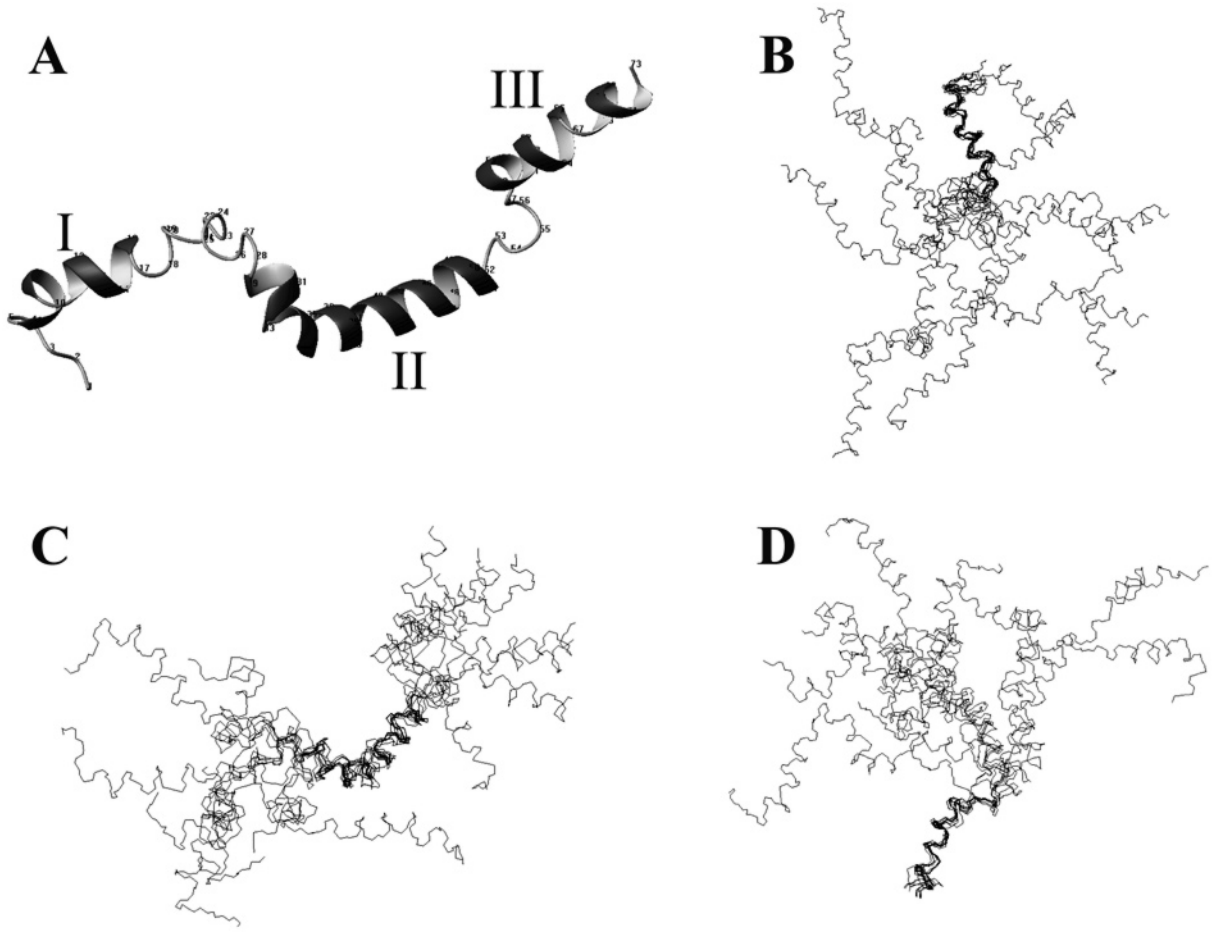

FIGURE 9: NMR-based molecular model of E3-M7-24-T40 in TFE/H_2O (1:1), 25 °C. (A) Representative NMR-derived model of the peptide indicating three helical domains. (B) Overlay of region I (residues 5−24) of 10 minimum NOE distance violation structures with backbone rmsd = 1.9 Å. (C) Overlay of region II (residues 30−52) with backbone rmsd = 1.86 Å. (D) Overlay of region III (residues 55−70) with backbone rmsd = 1.5 Å.

yield occurred when 0.2% glucose was used as compared with the 0.4% glucose normally supplemented. However, this reduction was cost-effective and allowed us to obtain the desired product for an expenditure of about $700 in labeled ammonium sulfate and glucose.

We have previously carried out CD studies to evaluate the secondary structure of peptides corresponding to the sixth and seventh transmembrane domains of Ste2p (*37, 40*). The CD studies on E3-M7-24-T40 indicated that both in TFE/H_2O environments and in the presence of various detergents and lipid vesicles this peptide maintained high helicity. Most importantly, the CD studies provided clear evidence that there is no significant difference in CD spectra of this peptide measured at low (50 mM) and high (500 mM) concentrations (data not shown). Therefore, using isotopically labeled E3-M7-24-T40 high-resolution NMR studies could be conducted in both organic−aqueous media and the presence of various detergents allowing us to compare structural findings under different membrane-mimetic conditions.

HSQC spectra of ^{15}N labeled peptides are useful to obtain a preliminary assessment concerning the presence of secondary or tertiary structure in a peptide/protein, information about the behavior of a peptide in solution, and its suitability for a further NMR investigation. $^{1}H-^{15}N$ HSQC spectra of [^{15}N]-E3-M7-24-T40 were measured in organic−aqueous and detergent media. Spectra measured on 0.5 mM concentrations of [^{15}N]-E3-M7-24-T40 in organic−aqueous media (Figure 5) and 0.2 mM or 0.5 mM peptide (Figure 6) in the presence of 400-fold molar excess of detergent were well-resolved and showed a dispersity consistent with the existence of secondary structure. In contrast, when a 20-fold molar excess of detergent at peptide concentration of 0.5 mM was used, poorly resolved spectra with extensive line broadening were obtained (data not shown). Opella and co-workers previously showed that the peptide/detergent ratio is the critical parameter in getting well-resolved spectra in detergents (*69*). In their study, poor resolution and, in some cases, especially with SDS, peak doubling were observed at low detergent concentrations.

A recent report screened 25 detergents for suitability in determining structural information on membrane proteins and concluded that PPG was most fit as judged using five proteins (*64*). Our results indicated that, for the E3-M7-24-T40 domain of Ste2p, DHPC and PPG were superior to SDS and DPC as judged by line width and peak separation; unfortunately, in these detergents, sample stability was insufficient for long NMR experiments. As indicated in Results, we were also unable to prepare a stable peptide solution in detergent at peptide concentrations high enough to perform 3D NMR measurements using our 600 MHz NMR instrument. We intend to pursue the structure of E3-M7-24-T40 in detergent on an 800 MHz spectrometer equipped with a cryoprobe using 0.2 mM $^{13}C/^{15}N$-labeled peptide in the presence of DPC micelles.

On the basis of the fact that stable solutions with 0.5 mM peptide concentrations could be readily obtained in TFE/H_2O (1:1) and in $CDCl_3/CD_3OH/H_2O$ (4:4:1) and that similar CD spectra (*40*) and similar overall peak spreading in $^{15}N-$HSQC spectra were observed in all media examined (Figures 5 and 6), a complete NMR analysis on E3-M7-24-T40 in organic/water media was undertaken. We chose 50% aqueous TFE because the CD results indicated very similar secondary structures in both 75% and 50% TFE (data not shown) and we wanted to minimize the concentration of the organic cosolvent. Chloroform/methanol/water has also been used as a membrane mimetic solvent, and the EmrE membrane transport protein could be reconstituted in biologically active form from this medium (*70*). Therefore, we also carried out a detailed NMR analysis in chloroform/methanol/water.

Initial insights into the structural tendencies of E3-M7-24-T40 came from analysis of the δCHα values for the E3-M7-24-T40 peptide in TFE/H_2O (1:1) and in $CDCl_3/CD_3OH/H_2O$ (4:4:1) at 25 °C (Figure 8A). One might question the use of random coil values obtained in water in the determination of chemical shift indices in organic aqueous medium. In comparing the chemical shifts in TFE/water with those in chloroform/methanol/water, we observed that despite the differences in these solvents the CHα shifts are very close and, in ~80% of the residues, differed by less than 0.1 ppm (Tables SI-1 and SI-2; Supporting Information). This suggests that the local conformation of the residue is more determining for the CHα chemical shift than the solvent. Moreover, even in a protein, the local dielectric constant of residues in the same helix varies significantly depending on whether these residues are buried or exposed. Others have used chemical shift indexing to assess structure of membrane peptides in organic aqueous media and in detergents (*71*−*73*). We believe that, provided that the conclusions reached are in agreement with those from other NMR parameters, chemical shift indexing provides a useful way to analyze secondary structural tendencies of domains of long peptides.

We observed three regions of uninterrupted chemical shift indices of −1. When the criterion that a helical region should have at least four consecutive residues with chemical shift indices of −1 (*65*, *66*) is used, residues 8−22, 25−52, and 59−70 have helical tendencies. On the basis of the CHα chemical shift data, 55 of 73 residues (75%) of the peptide are helical. This approach overestimates the percent helicity compared to that determined from the CD data (43%) in TFE/H_2O (1:1).

When the long range (i, i+3) NOE connectvities is used as the criterion, three segments of E3-M7-24-T40 have helical tendencies: 5−20, 38−52, and 60−70 in both $CDCl_3/CD_3OH/H_2O$ and TFE/H_2O solutions (Figure 7A,B). Furthermore, in TFE/H_2O, $J_{NH-CH\alpha}$ coupling constants of 3−5 Hz are also found for many residues in these same regions with the J values drifting up near the carboxyl terminus (Figure 7C). Thus, on the basis of three independent criteria—chemical shift indices, Overhauser connectivities, and J-coupling values—we can conclude that for E3-M7-24-T40 in TFE/H_2O and on chemical shift indices and NOE connectivities in $CDCl_3/CD_3OH/H_2O$ the putative transmembrane region (residues 10−33) is helical with a break at the proline-24 residue, that the region following this domain also has significant helical tendencies, and that the residues near the end of the peptide have a small but clear tendency to form helices.

The helical tendency in the cytosolic tail was confirmed by the NMR derived model of the E3-M7-24-T40 peptide that indicated three helical domains [residues 5−24, 30−52, and 55−70] (Figure 9 and Supporting Information). The backbone rmsd calculated for the entire peptide was very large (~10 Å). This is likely due to the fact that significant segmental mobility is available in the complete 73-residue peptide in organic−aqueous solvents and that this peptide does not assume one 3D structure in these membrane mimetics. However, when 10 structures modeled for the TFE/H_2O solution with the minimum NOE distance violations were overlaid separately for each of the three helical domains, much smaller backbone rmsd values were calculated for these regions of the multidomain peptide (Figure 9). The poorer convergence observed in the model derived from $CDCl_3/CD_3OH/H_2O$ data is likely due to the availability of fewer long-range constraints for use in the calculation. Nevertheless, the similar outcomes from studies in two quite different media suggests that the essential structural features may be an inherent characteristic of this domain of Ste2p.

The conclusion that the multidomain E3-M7-24-T40 receptor peptide has helices with differential stabilities is supported by the CD studies at different temperatures (Figure 4). About one-third of the helical residues appear to randomize as the temperature is raised to 110 °C. The remaining residues likely correspond to the transmembrane region of this peptide and form an extraordinarily stable secondary structure.

A previous CD investigation from our laboratory concluded that a 40-residue peptide corresponding to Ste2p-[300−339] had a partial helical structure in TFE/H_2O (3:1) but was primarily disordered in TFE/H_2O (1:1) (*40*). Similar CD results were obtained on a 23-residue peptide corresponding to Ste2p[350−372] (*35*). An NMR investigation revealed that the same 23-residue peptide was disordered in H_2O (*36*). In contrast to the previous CD and NMR data (*35*, *36*, *40*) of the tail fragments alone, the present NMR results on the multidomain fragment show that when attached to the transmembrane domain many residues in the tail of Ste2p from position 300 to 372 have helical tendencies. Our hypothesis is that in the context of the transmembrane domain there is some constraint on the conformational freedom of the tail residues resulting in an increased propensity to assume helical conformations. If this is true, the subtraction method previously used (*40*) to determine the helicity of the transmembrane domain must be reevaluated. More insights into this question will come after the structure of E3-M7-24-T40 in detergents is determined.

It is known that the cytosolic tail of Ste2p plays an essential role in desensitization of this GPCR by endocytosis (*74*). Many studies have shown that GPCR endocytosis requires phosphorylation by kinases and additional protein−protein interactions (*75*). Since protein−protein interactions often involve helical surfaces (*76*, *77*), it is reasonable to hypothesize that the helical tendencies found for domains of the cytosolic tail of Ste2p reflect the biologically relevant structure of these regions of the receptor. It is also noteworthy that the X-ray structure of rhodopsin indicated that residues following the seventh transmembrane domain of this GPCR are helical (*5*). This X-ray structure was determined for

crystals in a mixed detergent/lipid medium. Our results represent the first NMR data suggesting that a cytosolic tail region of Ste2p has some preferred secondary structure.

The relevance of structures determined in aqueous trifluoroethanol (TFE) to the biological state of a protein has been discussed almost since the introduction of this solvent in conformational analysis of peptides (*78*, *79*). Many studies on peptides and protein fragments indicate a reasonably good correlation between the TFE-induced structure of the peptides and the protein structure (*80*). Our experience with transmembrane peptides from Ste2p is that titration of TFE solutions of these peptides with water resulted in CD spectra that correlated well with those measured in detergents or lipid bilayers. It is significant that, although the HSQC spectra of E3-M7-24-T40 in TFE/H_2O and $CDCl_3$/CD_3OH/H_2O appear different (Figure 5), systematic comparison of nitrogen-15 (^{15}N) chemical shifts and δCHα values in both solvents shows that they were almost identical (Figure 8B; Supporting Information tables). The amide NH proton chemical shifts in both solvents reveal only minor differences (Figure 8C). These similarities allow us to conclude that the conformation of the E3-M7-24-T40 peptide in TFE/H_2O and $CDCl_3$/CD_3OH/H_2O will be very close. Unlike TFE/H_2O mixtures, $CDCl_3$/CD_3OH/H_2O has not been considered to be a helix-inducing medium. Indeed, a high-resolution structure of the *E. coli* EmrE transport protein in $CDCl_3$/CD_3OH/H_2O (6:6:1) has been taken to represent the native structure of this molecule (*70*). Therefore, we believe that the structural preferences for the 73-residue domain of Ste2p determined in this paper may be relevant to the functional state of this GPCR. The ability to measure well-dispersed and highly resolved spectra of E3-M7-24-T40 in detergents (Figure 6) and organic−aqueous media (Figure 5) and the ability to model its structure should permit comparison of high-resolution structures of this multidomain peptide under a broad spectrum of conditions. Such structures will provide further insights into the biological relevance of studies in organic−aqueous media.

In conclusion the successful biosynthesis of an isotopically labeled 73-residue multidomain peptide from a GPCR on a 10 mg scale has been achieved. The results show that biosynthesis of peptide fragments of membrane receptors needs to be optimized to obtain efficient synthesis and that no one expression system or CNBr release condition will likely be useful with all receptor domains. The multidomain peptide we studied exhibited well-dispersed HSQC spectra in several membrane mimetic solvents, including a variety of detergents at high detergent/peptide ratios. The NMR results in TFE/H_2O and in $CDCl_3$/CD_3OH/H_2O indicate that residues in the cytosolic tail region of Ste2p may have helical tendencies. We suggest that these helical tendencies may play a role in protein−protein interactions involved in the desensitization of this GPCR.

ACKNOWLEDGMENT

We thank Dr. Hsin Wang for assistance in setting up the NMR experiments. The NMR Facility is operated by the College of Staten Island and the CUNY Institute for Macromolecular Assemblies.

SUPPORTING INFORMATION AVAILABLE

Information relating to the nitrogen-15 and proton NMR chemical shifts of the E3-M7-24-T40 peptide in TFE/H_2O (1:1) [Table SI-1] and $CDCl_3$/CD_3OH/H_2O (4:4:1) [Table SI-2] and an NMR-based molecular model of E3-M7-24-T40 in $CDCl_3$/CD_3OH/H_2O (Figure SI-3). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Wallin, E., and von Heijne, G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, *Protein Sci. 7*, 1029−1038.
2. Lundstrom K. (2004) Structural genomics on membrane proteins: mini review, *Comb. Chem. High Throughput Screening 7*, 431−439.
3. Dohlman, H. G., Thorner, J., Caron, M. G., and Lefkowitz, R. J. (1991) Model systems for the study of seven-transmembrane-segment receptors, *Annu. Rev. Biochem. 60*, 653−688.
4. Fraser, C. M., Lee, N. H., Pellegrino, S. M., and Kerlavage, A. R. (1994) Molecular properties and regulation of G-protein-coupled receptors, *Prog. Nucleic Acids Res. Mol. Biol. 49*, 113−156.
5. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) Crystal structure of rhodopsin: a G protein-coupled receptor, *Science 289*, 739−745.
6. Baldwin, J. M. (1993) The probable arrangement of the helices in G protein-coupled receptors, *EMBO J. 12*, 1693−1703.
7. Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P., and Lanyi, J. K. (1999) Structure of bacteriorhodopsin at 1.55 Å resolution, *J. Mol. Biol. 291*, 899−911.
8. Baldwin, J. M., Schertler, G. F., and Unger, V. M. (1997) An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors, *J. Mol. Biol. 12*, 144−164.
9. Herzyk, P., and Hubbard, R. E. (1998) Combined biophysical and biochemical information confirms arrangement of transmembrane helices visible from the three-dimensional map of frog rhodopsin, *J. Mol. Biol. 281*, 741−754.
10. Cramer, W. A., Engelman, D. M., Von Heijne, G., and Rees, D. C. (1992) Forces involved in the assembly and stabilization of membrane proteins, *FASEB J. 6*, 3397−3402.
11. Herzyk, P., and Hubbard, R. E. (1995) Automated method for modeling seven-helix transmembrane receptors from experimental data, *Biophys. J. 69*, 2419−2442.
12. Dohlman, H. G., and Thorner, J. W. (2001) Regulation of G protein-initiated signal transduction in yeast: paradigms and principles, *Annu. Rev. Biochem. 70*, 703−754.
13. Spiegel, A. M., and Weinstein, L. S. (2004) Inherited diseases involving G proteins and G protein-coupled receptors, *Annu. Rev. Med. 55*, 27−39.
14. Blumer, K. J., and Thorner, J. (1991) Receptor-G protein signaling in yeast, *Annu. Rev. Physiol. 53*, 37−57.
15. Nakayama, N., Miyajima, A., Arai, K., and Matsumoto, K. (1985) Nucleotide sequences of *STE2* and *STE3*, cell type-specific sterile genes from *Saccharomyces cerevisiae*, *EMBO J. 4*, 2643−2648.
16. Horn, F., Vriend, G., and Cohen, F. E. (2001) Collecting and harvesting biological data: the GPCRDB and NucleaRDB information systems, *Nucleic Acids Res. 29*, 346−349.
17. Parrish, W., Eilers, M., Ying, W., and Konopka, J. B. (2002) The cytoplasmic end of transmembrane domain 3 regulates the activity of the *Saccharomyces cerevisiae* G-protein-coupled alpha-factor receptor, *Genetics 160*, 429−443.
18. Dube, P., DeCostanzo, A., and Konopka, J. B. (2000) Interaction between transmembrane domains five and six of the alpha-factor receptor, *J. Biol. Chem. 275*, 26492−26499.
19. Eilers, M., Hornak, V., Smith, S. O., and Konopka, J. B. (2005) Comparison of class A and D G protein coupled receptors: common features in structure and activation, *Biochemistry 44*, 8959−8975.
20. Konopka, J. B., Margarit, S. M., and Dube, P. (1996) Mutation of Pro-258 in transmembrane domain 6 constitutively activates the G protein-coupled alpha-factor receptor, *Proc. Natl. Acad. Sci. U.S.A. 93*, 6764−6769.

21. Wess, J. (1997) G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition, *FASEB J. 11*, 346−354.
22. Dube, P., and Konopka, J. B. (1998) Identification of polar region in transmembrane domain 6 that regulates the function of the G protein-coupled α-factor receptor, *Mol. Cell. Biol. 18*, 7205−7215.
23. Mierke, D. F., and Giragossian, C. (2001) Peptide hormone binding to G-protein-coupled receptors: structural characterization via NMR techniques, *Med. Res. Rev. 21*, 450−471.
24. Naider, F., Arshava, B., Ding, F. X., Arevalo, E., and Becker, J. M. (2001) Peptide fragments as models to study the structure of a G-protein coupled receptor: the alpha-factor receptor of *Saccharomyces cerevisiae*, *Biopolymers 60*, 334−350.
25. Martin, N. P., Leavitt, L. M., Sommers, C. M., and Dumont, M. E. (1999) Assembly of G protein-coupled receptors from fragments: identification of functional receptors with discontinuities in each of the loops connecting transmembrane segments, *Biochemistry 38*, 682−695.
26. Ridge, K. D., Lee, S. S., and Yao, L. L. (1995) In vivo assembly of rhodopsin from expressed polypeptide fragments, *Proc. Natl. Acad. Sci. U.S.A. 92*, 3204−3208.
27. Marti, T. (1998) Refolding of bacteriorhodopsin from expressed polypeptide fragments, *J. Biol. Chem. 273*, 9312−9322.
28. Katragadda, M., Alderfer, J. L., and Yeagle, P. L. (2001) Assembly of a polytopic membrane protein structure from the solution structures of overlapping peptide fragments of bacteriorhodopsin, *Biophys. J. 81*, 1029−1036.
29. Yeagle, P. L., Choi, G., and Albert, A. D. (2001) Studies on the structure of the G-protein-coupled receptor rhodopsin including the putative G-protein binding site in unactivated and activated forms, *Biochemistry 40*, 11932−11937.
30. Hunt, J. F., Earnest, T. N., Bousche, O., Kalghatgi, K., Reilly, K., Horvath, C., Rothschild, K. J., and Engelman, D. M. (1997) A biophysical study of integral membrane protein folding, *Biochemistry 36*, 15156−15176.
31. Albert, A. D., and Yeagle, P. L. (2000) Domain approach to three-dimensional structure of rhodopsin using high-resolution nuclear magnetic resonance, *Methods Enzymol. 315*, 107−115.
32. Katragadda, M., Chopra, A., Bennett, M., Alderfer, J. L., Yeagle, P. L., Albert, A. D. (2001) Structures of the transmembrane helices of the G-protein coupled receptor, rhodopsin, *J. Pept. Res. 58*, 79−89.
33. Yeagle, P. L., Danis, C., Choi, G., Alderfer, J. L., and Albert, A. D. (2000) Three-dimensional structure of the seventh transmembrane helical domain of the G-protein receptor, rhodopsin, *Mol. Vision 6*, 125−131.
34. Chopra, A., Yeagle, P. L., Alderfer, J. A., and Albert, A. D. (2000) Solution structure of the sixth transmembrane helix of the G-protein-coupled receptor, rhodopsin, *Biochim. Biophys. Acta 1463*, 1−5.
35. Reddy, A. P., Tallon, M. A., Becker, J. M., and Naider, F. (1994) Biophysical studies on fragments of the alpha-factor receptor protein, *Biopolymers 34*, 679−689.
36. Arshava, B., Liu, S. F., Jiang, H., Breslav, M., Becker, J. M., and Naider, F. (1998) Structure of segments of a G protein-coupled receptor: CD and NMR analysis of the *Saccharomyces cerevisiae* tridecapeptide pheromone receptor, *Biopolymers 46*, 343−357.
37. Xie, H., Ding, F. X., Schreiber, D., Eng, G., Liu, S. F., Arshava, B., Arevalo, E., Becker, J. M., and Naider, F. (2000) Synthesis and biophysical analysis of transmembrane domains of a *Saccharomyces cerevisiae* G protein-coupled receptor, *Biochemistry 39*, 15462−15474.
38. Ding, F., Xie, H., Arshava, B., Becker, J. M., and Naider, F. (2001) ATR-FTIR study of the structure and orientation of transmembrane domains of the *Saccharomyces cerevisiae* α-mating factor receptor in phospholipids, *Biochemistry 40*, 8945−8954.
39. Arshava, B., Taran, I., Xie, H., Becker, J. M., and Naider, F. (2002) High-resolution NMR analysis of the seven transmembrane domains of a heptahelical receptor in organic−aqueous medium, *Biopolymers 64*, 161−176.
40. Naider, F., Ding, F. X., VerBerkmoes, N. C., Arshava, B., and Becker, J. M. (2003) Synthesis and biophysical characterization of a multidomain peptide from a *Saccharomyces cerevisiae* G protein-coupled receptor, *J. Biol. Chem. 278*, 52537−52545.
41. Miozzari, G. F., and Yanofsky, C. (1978) Translation of the leader region of the *Escherichia coli* tryptophan operon, *J. Bacteriol. 133*, 1457−1466.
42. Kleid, D. G., Yansura, D., Small, B., Dowbenko, D., Moore, D. M., Grubman, M. J., McKercher, P. D., Morgan, D. O., Robertson, B. H., and Bachrach, H. L. (1981) Cloned viral protein vaccine for foot-and-mouth disease: responses in cattle and swine, *Science 214*, 1125−1129.
43. Therien, A. G., Glibowicka, M., and Deber, C. M. (2002) Expression and purification of two hydrophobic double-spanning membrane proteins derived from the cystic fibrosis transmembrane conductance regulator, *Protein Expression Purif. 25*, 81−86.
44. Staley, J., and Kim, P. (1994) Formation of a native-like subdomain in a partially folded intermediate of bovine pancreatic trypsin inhibitor, *Protein Sci. 3*, 1822−1832.
45. Naider, F., Estephan, R., Englander, J., Suresh Babu, V. V., Arevalo, E., Samples, K., and Becker, J. M. (2004) Sexual conjugation in yeast: a paradigm to study G-protein-coupled receptor domain structure, *Biopolymers 76*, 119−128.
46. Akal-Strader, A., Khare, S., Xu, D., Naider, F., and Becker, J. M. (2002) Residues in the first extracellular loop of a G protein-coupled receptor play a role in signal transduction, *J. Biol. Chem. 277*, 30581−30590.
47. LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F., and McCoy, J. M. (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm, *Biotechnology 11*, 187−193.
48. Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection, *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
49. Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins, *J. Biomol. NMR 20*, *71*−75.
50. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
51. Walker, J., and Gaastra, W. (1988) *Methods in Molecular Biology, New Protein Techniques*, 3rd ed., pp 427−440, Humana Press Inc., Clifton, New Jersey.
52. Arevalo, E., Estephan, R., Madeo, J., Arshava, B., Dumont, M., Becker, J. M., and Naider, F. (2003) Biosynthesis and biophysical analysis of domains of a yeast G protein-coupled receptor, *Biopolymers 71*, 516−531.
53. Wu, C. S., Ikeda, K., and Yang, J. T. (1981) Ordered conformation of polypeptides and proteins in acidic dodecyl sulfate solution, *Biochemistry 20*, 566−570.
54. Chen, Y. H., Yang, J. T., and Chau, K. H. (1974) Determination of the helix and β form of proteins in aqueous solution by circular dichroism, *Biochemistry 13*, 3350−3359.
55. Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *J. Am. Chem. Soc. 114*, 10663−10665.
56. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques, *J. Biomol. NMR 4*, 845−858.
57. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA, *J. Mol. Biol. 273*, 283−298.
58. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX Pipes, *J. Biomol. NMR* 6, 277−293.
59. Johnson, B. A., and Blevins, R. A. (1994) A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603−614.
60. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55.
61. Martin, N. P., Celic, A., and Dumont, M. E. (2002) Mutagenic mapping of helical structures in the transmembrane segments of the yeast α-factor receptor, *J. Mol. Biol. 317*, 765−788.
62. Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997) Attenuated T-2 relaxation by mutual cancellation of DD coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12366−12371.
63. Vinogradova, O., Sonnichsen, F., Sanders, C. R., II (1998) On choosing a detergent for solution NMR studies of membrane proteins, *J. Biomol. NMR 11*, 381−386.
64. Krueger-Koplin, R. D., Sorgen, P. L., Krueger-Koplin, S. T., Rivera-Torres, I. O., Cahill, S. M., Hicks, D. B., Grinius, L., Krulwich, T. A., and Girvin, M. E. (2004) An evaluation of

detergents for NMR structural studies of membrane proteins, *J. Biomol. NMR 28*, 43−57.
65. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, *Biochemistry 31*, 1647−1651.
66. Wishart, D. S., and Sykes, B. D. (1994) Chemical shifts as a tool for structure determination, *Methods Enzymol. 239*, 363−392.
67. Zheng, H., Zhao, J., Wang, S., Lin, C.-M., Chen, T., Jones, D. H., Ma, C., Opella, S., and Xie, X.-Q. (2005) Biosynthesis and purification of a hydrophobic peptide from transmembrane domains of G-protein-coupled CB2 receptor, *J. Pept. Res. 65*, 450−458.
68. Lindhout, D. A., Thiessen, A., Schieve, D., and Sykes, B. D. (2003) High-yield expression of isotopically labeled peptides for use in NMR studies, *Protein Sci. 12*, 1786−1791.
69. McDonnell, P. A. and Opella, S. J. (1993) Effect of detergent concentration on multidimensional solution NMR spectra of membrane proteins in micelles, *J. Magn. Reson., Ser. B 102*, 120−125.
70. Schwaiger, M., Lebendiker, M., Yerushalmi, H., Coles, M., Gröger, A., Schwarz, C., Schuldiner, S., and Kessler, H. (1998) NMR investigation of the multidrug transporter EmrE, an integral membrane protein, *Eur. J. Biochem. 254*, 610−619.
71. Franzoni, L., Nicastro, G., Pertinhez, T. A., Oliveira, E., Nakaie, C. R., Paiva, A. C. M., Schreier, S., and Spisni, A. (1999) Structure of two fragments of the third cytoplasmic loop of the rat angiotensin II AT_{1A} receptor, *J. Biol. Chem. 274*, 227−235.
72. Déméné, H., Granier, S., Muller, D., Guillon, G., Dufour, M.-N., Delsuc, M.-A., Hibert, M., Pascal, R., and Mendre, C. (2003) Active peptidic mimics of the second intracellular loop of the V_{1A} vasopressin receptor are structurally related to the second intracellular rhodopsin loop: a combined ^{1}H NMR and biochemistry study, *Biochemistry 42*, 8204−8213.
73. Zhang, L., DeHaven, R. N., and Goodman, M. (2002) NMR and modeling studies of a synthetic extracellular loop II of the κ opioid receptor in a DPC micelle, *Biochemistry 41*, 61−68.
74. Hicke, L., Zanolari, B., and Riezman, H. (1998) Cytoplasmic tail phosphorylation of the alpha-factor receptor is required for its ubiquitination and internalization, *J. Cell Biol. 141*, 349−358.
75. Gainetdinov, R. R., Premont, R. T., Bohn, L. M., Lefkowitz, R. J., and Caron, M. G. (2004) Desensitization of G protein-coupled receptors and neuronal functions, *Annu. Rev. Neurosci. 27*, 107−144.
76. Eilers, M., Patel, A. B., Liu, W., and Smith, S. O. (2002) Comparison of helix interactions in membrane and soluble alpha-bundle proteins, *Biophys. J. 82*, 2720−2736.
77. Gimpelev, M., Forrest, L. R., Murray, D., and Honig, B. (2004) Helical packing patterns in membrane and soluble proteins, *Biophys. J. 87*, 4075−4086.
78. Tamburro, A. M., Scatturin, A., Rocchi, R., Marchiori, F., Borin, G., and Scoffone, E. (1968) Conformational-transitions of bovine pancreatic ribonuclease S-peptide, *FEBS Lett. 1*, 298−300.
79. Goodman, M., Naider, F., and Toniolo, C. (1971) Circular dichroism studies of isoleucine oligopeptides in solution, *Biopolymers 10*, 1719−1730.
80. Sonnichsen, F. D., Van Eyk, J. E., Hodges, R. S., and Sykes, B. D. (1992) Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide, *Biochemistry 31*, 8790−8798.

BI0507231