Transmembrane Domain of Cystic Fibrosis Transmembrane Conductance Regulator: Design, Characterization, and Secondary Structure of Synthetic Peptides m1−m6[†]

W. Christian Wigley,[‡] S. Vijayakumar,[‡,§] Jeffrey D. Jones,[‡] Clive Slaughter,[||] and Philip J. Thomas*[,‡]

Departments of Physiology and Biochemistry and The Howard Hughes Medical Institute, The University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235-9040

Received September 15, 1997; Revised Manuscript Received November 5, 1997

ABSTRACT: Mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) give rise to cystic fibrosis (CF), the most common genetic disease in the Caucasian population. CFTR is organized into five putative domains, including two that are predicted to be transmembrane and consist of six membrane-spanning segments each. CFTR mediates regulated anion transport across the apical membrane of epithelial cells. The pore through which CFTR transports its solutes is thought to be formed by some combination of the amino-terminal membrane-spanning segments. Although these sequences are predicted to be α-helical in secondary structure, to date, no direct structural evidence has been presented testing this hypothesis. Here, we present the biophysical characterization of six peptides (m1−m6) representing the predicted amino-terminal membrane-spanning domain of CFTR. The peptides can be incorporated into liposomes and are soluble in SDS micelles and trifluoroethanol (TFE). FTIR and CD spectroscopy indicate all six peptides adopt a stable, predominantly α-helical secondary structure in these environments. In contrast, peptide m6 undergoes a shift from α-helix to β-sheet when dissolved in 20% methanol. Additionally, the peptides show an increase in β-sheet in TFE, a known inducer of α-helices, relative to that seen in the nativelike environments. These results have implications for the folding of this complex membrane protein and suggest that the possible functional role of m6 is manifested through a shift in secondary structure.

Cystic fibrosis (CF)[1] is an autosomal recessive disorder that affects about 1 in 2000 Caucasians and results from mutations in the gene for the cystic fibrosis transmembrane conductance regulator (CFTR) (*1*). CFTR is a 169 kDa, cAMP-dependent protein kinase (PKA) regulated anion channel found in the apical membranes of epithelial cells (*2*, *3*). A number of studies indicate CFTR directly mediates chloride translocation (*4*−*6*). In addition, recent evidence has been presented implicating the protein as having a role in ATP release (*7*) and as a regulator of other channels including ENaC, ROMK, and ORCC (8−10).

CFTR is a member of the traffic-ATPase or ATP-binding cassette (ABC) supergene family of membrane transporters. Over a hundred members of this family have been identified to date including active, ATP-driven transporters for anions (bacterial phosphate and arsenate transporter systems), pteridines *(Drosophila melanogaster* white), lipids (MDR2), and peptides (*Saccharomyces cerevisiae* STE6, *Escherichia coli* OppD, and class 1 MHC−TAP1 and -2) (*11*). CFTR consists of 1480 amino acids organized into functional domains common to this class of proteins (*1*). It is predicted from primary structure and hydropathy analysis that CFTR includes two membrane-spanning domains, TMD1 and TMD2, each containing six transmembrane segments designated m1−m12, and two nucleotide binding domains, NBD1 and NBD2 (*1*). In addition, CFTR contains a cytoplasmic regulatory domain, R, which is required, in part, for the PKA sensitivity of channel gating (*12*).

Not only is channel gating regulated by PKA-catalyzed phosphorylation of serine residues in the R-domain, but it depends upon ATP binding and hydrolysis at the nucleotide-binding domains (*13*−*18*). This nucleotide hydrolysis may satisfy energy requirements associated not only with channel gating but possibly other crucial, functionally significant conformation changes. Although the predicted transmem-

[†] This work was supported by research grants from the National Institutes of Health−NIDDK (DK49835), the Cystic Fibrosis Foundation (Thomas95GO), and the American Cancer Society (IRG-142L).

* Corresponding author: Tel (214) 648-8723; Fax (214) 648-9268; email thomas07@utsw.swmed.edu.

[‡] Department of Physiology.

[§] Present address: Department of Medicine, College of Physicians and Surgeons of Columbia University, New York, NY 10032.

[||] Department of Biochemistry and The Howard Hughes Medical Institute.

[1] Abbreviations: CF, cystic fibrosis; CFTR, cystic fibrosis transmembrane conductance regulator; TMD, transmembrane domain; NBD, nucleotide binding domain; R, regulatory domain; kDa, kilodalton; cAMP, adenosine 3′,5′-cyclic monophosphate; PKA, adenosine 3′,5′-cyclic monophosphate-dependent protein kinase; ATP, adenosine 5′-triphosphate; PC, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; PS, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-[phospho-L-serine]; MEKE, micellar electrokinetic electrophoresis; HPLC, high-performance liquid chromatography; UV, ultraviolet; CD, circular dichroism; ATR-FTIR, attenuated total reflectance Fourier transform infrared; TFE, 2,2,2-trifluoroethanol; HFIP, 1,1,1,3,3,3-hexafluoro-2-propanol; β-OG, octyl-β-D-glucopyranoside; TFA, trifluoroacetic acid; SDS, sodium dodecyl sulfate; dd-H_2O, deionized/distilled water.

S0006-2960(97)02293-9 CCC: $15.00
Published on Web 01/20/1998

brane domains have been less well characterized than those domains residing in the cytosol, recent studies addressing the topology of membrane-integrated CFTR (*19*) support the initial identification of the TMD sequences (*1*). Mutational analyses have provided compelling evidence that portions of these putative membrane-spanning sequences are required for channel function and ion selectivity (*4*, *5*, *20*). In addition, truncation mutations of the C-terminal TMD from CFTR indicate only the first six transmembrane-spanning sequences are required to form a functional Cl^- channel (*21*). Additional insight into the structure of the channel is provided by substituted-cysteine accessibility studies, which are consistent with α-helical secondary structure for m1, both α-helical and extended structure for m6, and a channel diameter of 6 Å (*6*, *22*).

In contrast to the large body of mutational data, little direct structural information pertinent to CFTR and the other ABC transporters exists. This is due, in part, to the lack of availability of these proteins in quantities sufficient for direct biophysical investigations of structure and function. Thus, predictions of secondary structure and their possible implications for the roles of CFTR membrane-spanning segments have yet to be experimentally tested. An understanding of these structures would appear essential to the development of testable models of CFTR structure.

Model transmembrane peptides have been used to investigate structure and function in several integral membrane proteins (*23*−*29*). Peptides representing the transmembrane regions of the voltage-gated sodium channel have been shown to adopt α-helical structures in detergent micelles and TFE (*29*). Studies of peptide models of the membrane-spanning sequences of bacteriorhodopsin (*24*, *25*), and glycophorin A (*26*−*28*), membrane proteins of known three-dimensional structure, indicate that the fold of these proteins is largely dictated by interactions within the membrane and can occur in the absense of the remainder of the protein. Recently, peptides corresponding to two of the CFTR amino-terminal transmembrane sequences have been shown to spontaneously insert into bilayers and form anion-selective channels (*30*). These and other studies have demonstrated that transmembrane sequences can assume membrane-integrated conformations and even form nativelike interpeptide interactions, yielding functional model structures. In order to characterize the transmembrane domains of CFTR, the structures of chemically synthesized peptides corresponding to each of the six predicted transmembrane segments of the amino-terminal transmembrane domain were studied spectroscopically. Here we report an extensive biophysical investigation of the secondary structures of these six peptides (m1−m6), as well as a possible structural role for m6 manifested through a conformational shift.

EXPERIMENTAL PROCEDURES

Materials. Synthetic peptides were generated by the UTSW−Howard Hughes Medical Institute Biopolymers Facility. Protein concentration was determined using assay reagents from Bio-Rad. 2,2,2-Trifluoroethanol (TFE) was from Fisher Biotech (Fair Lawn, NJ). 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP) was from Aldrich (Milwaukee, WI). Phospholipids were obtained from Avanti Polar Lipids, Inc. (Birmingham, AL). Reagent grade octyl-β-D-glucopyranoside (β-OG) was from Calbiochem (La Jolla, CA). All other reagents were of the highest quality commercially available.

Peptide Synthesis. Synthetic peptides corresponding to the six putative membrane-spanning regions of TMD1 were generated on a Rainin Symphony multiplex peptide synthesizer using (fluorenylmethoxy)carbonyl (Fmoc) chemistry. Fmoc-protected amino acids were used with an *N*,*N*-dimethylformamide (DMF) solvent system. Amino acids were activated with a mixture of 100 mM HBTU [2-(1-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate] and 0.4 M 4-methylmorpholine in DMF. Suitable side-chain protecting groups that were easily removed during cleavage with a solution containing 86% (v/v) trifluoroacetic acid (TFA), 4.5% (v/v) water, 4.5% (v/v) anisole, and 4.5% (v/v) thioanisole were selected. The synthesized peptides were stored lyophilized and desiccated as TFA salts at −70 °C prior to use.

Mass Spectrometry. Synthetic peptides were analyzed by negative ion fast atom bombardment mass spectrometry and electrospray mass spectrometry using a 30−250 quadrupole or Quattro II triple quadrupole instrument (Micromass, Altrincham, England).

Micellar Electrokinetic Capillary Electrophoresis. Purity of the synthetic peptides was assessed by MEKE using an Applied Biosystems Model 270A-HT capillary electrophoresis instrument. Synthetic peptides were dissolved (200 μg/mL) in 0.01% HFIP, 20 mM sodium borate, pH 9.0, and 25 mM SDS. Each peptide solution (5 μL) was injected hydrodynamically and monitored by absorption at 220 nm.

Isocratic High-Performance Liquid Chromatography. Synthetic peptide purity was further assessed under isocratic conditions on a Beckman 4.6 mm × 150 mm 5 μm C-18 column. Peptides were dissolved in 100% formic acid (FA) and then diluted 10-fold with 55% 2-propanol (IPA) to a final concentration of 1 mg/mL in 10% FA/50% IPA. Aliquots (10 μL) were chromatographed using a mobile phase of 10% FA/50% IPA at a flow rate of 0.5 mL/min. Peak detection was at 270 nm, due to the lack of tryptophan in some of the peptides and to avoid interfering formate absorption.

Tricine−SDS−Polyacrylamide Gel Electrophoresis. The purity and oligomeric state of synthetic peptides dissolved in SDS was investigated by gel electrophoresis. The buffer system of Schägger and von Jagow (*31*), was employed except that the SDS concentration in the gel was 0.5%, above the critical micellar concentration. Briefly, peptides dissolved in 0.5% (17 mM) SDS and 5 mM phosphate buffer, pH 7.2, were diluted 2-fold with 2× SDS loading buffer containing the tracking dye Serva Blue G. Samples were either loaded directly or boiled 5 min and then loaded onto a 16.5% polyacrylamide−Tricine−SDS gel and then separated at a constant 120 V. Due to their small sizes, peptides were fixed in the gel by soaking in a solution containing 50% methanol and 40% glacial acetic acid for 30 min prior to staining. Protein was visualized by staining with either Coomassie blue or silver.

Reconstitution of Peptides into SDS Micelles. In order to mimic the native environment of the membrane-spanning segments, each synthetic peptide was reconstituted into SDS micelles. Briefly, approximately 0.5 mg of lyophilized peptide was dissolved in 0.4 mL of 5.0% SDS and 50 mM sodium phosphate, pH 7.2. This solution was sonicated three

times at 50% duty cycle for 5 s each at a low to moderate power setting using a Branson sonifier, Model 450, fitted with a microtip probe. The resulting solution was stored at room temperature while complete solubilization took place, within 72−96 h with periodic vortexing. The peptide/SDS solution was then diluted 10-fold with deionized distilled water (dd-H_2O) to a final concentration of 125 μg/mL of peptide in 0.5% SDS and 5 mM sodium phosphate, pH 7.2. The final SDS concentration is 17 mM, which is greater than its critical micellar concentration of 8.2 mM at this temperature.

Liposome Preparation and Peptide Incorporation. Protein-reconstituted liposomes were prepared by the β-OG dilution−dialysis method as previously described (*32*, *33*) with the following changes. Samples were prepared in a liposome buffer containing 10 mM Tris-HCl, 1 mM EDTA, and 1 mM dithiothreitol, pH 7.4. A lipid stock solution containing 30% 1-palmitoyl-2-oleoyl-sn-glycero-3-[phospho-L-serine] (PS) and 70% 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (PC) was prepared in chloroform. Ten micromoles of lipid stock was dried under nitrogen and subsequently lyophilized for 2 h. Lyophilized peptide (0.2 μmol) was added, followed by 1.0 mL of liposome buffer supplemented with 60 mM β-OG, yielding a protein−lipid molar ratio of 1:50. The resulting mixed micelles were bath-sonicated for 5 min and stored under nitrogen at room temperature. Periodic vortexing over the next 72 h was followed by dilution with liposome buffer to a β-OG concentration of 10 mM (below the β-OG cmc). Solutions were next exhaustively dialyzed to remove the β-OG monomer from the aqueous phase. Concentration of liposome solutions to a final volume of about 1 mL was performed using Centricon-10 ultrafiltration units. Finally, liposomes were supplemented with NaN_3 to 1 mM and stored at room temperature prior to analysis.

Negative-Staining Electron Microscopy. Liposome preparations diluted to 1 mM with liposome buffer were applied to 200-mesh copper grids in 5 μL aliquots, where they were allowed to fix for 1 min. Grids were washed twice with dd-H_2O and then stained with 1.0% uranyl acetate. After staining, grids were washed twice with dd-H_2O and allowed to dry. Liposomes were examined and photographed using a Jeol 100SX electron microscope (data not shown).

Ultraviolet Spectroscopy. UV absorbance spectra were collected using a Shimadzu UV-2101PC scanning spectrophotometer at 25 °C from 320 to 190 nm. Absorbance was determined every 0.5 nm with a band pass of 2.0 nm. For peptide quantification by UV absorption, molar extinction coefficients were calculated as previously described (*34*).

Circular Dichroism Spectroscopy. CD spectra were collected on an Aviv 62DS with a 3 s averaging time and a band pass of 1.5 nm. All spectra are the smoothed average of three consecutive runs scanning wavelengths from 260 to 190 nm. Sample spectra were obtained in a 0.1 cm Suprasil cuvette at 25 °C. Fractional helicities were calculated as described (35) assuming $[\theta]_{222}$ for 0% and 100% α-helix content are 2000 and 30 000 deg·cm^2/dmol, respectively (*36*).

Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy. ATR-FTIR spectra were collected at 25 °C utilizing a Nicolet Protégé 460 FTIR spectrometer equipped with a ZnSe ATR crystal from SpectraTech and a liquid nitrogen-cooled MCT detector. The optical bench was continuously purged with the boil-off from a liquid nitrogen dewar to reduce water vapor. All spectra collected were the average of 250 interferograms and were corrected for solvent signals by the subtraction of a solvent blank (without peptide) spectrum. Accuracy of solvent subtractions was monitored by minimization of the water absorbance peak at 2125 cm^{-1}. No additional water vapor absorbances were subtracted. Spectra were collected at a resolution of 4.0 cm^{-1} using Happ−Genzel apodization, which resulted in no appreciable loss of resolution. Data processing and analysis of the amide I region of the ATR-FTIR spectra was performed without smoothing using the OMNIC software package, version 3.1a, provided by the instrument manufacturer.

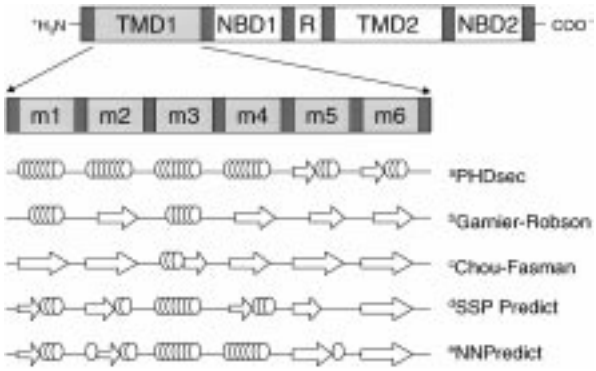

FIGURE 1: Organization and predicted secondary structures of CFTR TMD1 subdomains. A schematic representing the five domains of CFTR and the six putative transmembrane subdomains (m1−m6) of TMD1 is presented. Results from each of five secondary structure prediction algorithms are shown below the enlarged TMD1 graphic. (Regions predicted to assume α-helical conformations are indicated by ovals, and those predicted to assume β-sheet conformations are indicated by arrows.) [a]Rost and Sander (*38*, *39*). [b]Garnier et al. (*67*). [c]Chou and Fasman (*40*). [d]Solovyev and Salamov (*68*). [e]Kneller et al. (*69*) and McClelland and Rumelhart (*70*).

RESULTS

Peptide Design and Synthesis. The domain organization and the predicted secondary structures of the putative TMD1 transmembrane segments can be seen at the top of Figure 1. According to the CFTR model proposed by Riordan and coworkers (*1*), TMD1 was predicted to span the membrane six times with NBD1 residing in the cytoplasm. In the present study, sequence boundaries for the six model synthetic peptides were selected according to consensus sequences between the predictions of Eisenberg (*37*, *1*), PHDhtm (*38*, *39*), and Engelman, Steitz, and Goldman (*26*). Peptides (Table 1) were designed to include all residues predicted by a consensus of algorithms to be in the membrane as well as flanking charged residues that might aid in synthesis and solubility and anchor the peptides in the membrane environments. These are extremely hydrophobic peptides with average hydrophobicities ranging from 0.8 to as high as about 2 (Table 1).

We have applied several secondary structure prediction algorithms to the six peptide sequences and the results are graphically summarized in Figure 1. The apparent lack of consensus among results generated by these programs is not surprising as they use very different criteria for their predictions, often relying on structural basis sets lacking representative membrane protein contributions. However,

Table 1

ID	predicted transmembrane amino acids			synthetic peptide residues	peptide sequences[a]	molecular mass (Da)	p*I*[b]	Φ[c]
	Eisenberg	PHDhtm	Engelman					
m1	81−102	78−92	82−100	80−104	R F M F Y G I F L Y L G E V T K A V Q P L L L G R	2933	9.77	0.818
m2	118−138	122−137	125−143	117−138	R S I A I Y L G I G L C L L F I V R T L L L	2461	9.37	1.968
m3	195−215	198−212	198−218	193−216	E G L A L A H F V W I A P L Q V A L L M G L I W	2663	5.36	1.462
m4	221−241	220−239	218−239	220−242	Q A S A F C G L G F L I V L A L F Q A G L G R	2353	8.25	1.386
m5	308−328	307−332	308−330	307−331	S S A F F F S G F F V V F L S V L P Y A L I K G I	2758	8.86	1.458
m6	330−350	335−350	334−350	329−350	K G I I L R K I F T T I S F C I V L R M A V	2523	11.06	1.404

[a] Underlined amino acids in the peptide sequences represent predicted transmembrane residues with confirmed missense mutations giving rise to cystic fibrosis. [b] Isoelectric point. [c] Average hydrophobicity.

there does appear to be at least partial agreement with regard to the secondary structure peptides m3 and m6 are predicted to assume. Four of the five algorithms predict α-helix for m3, with the fifth (*40*) predicting a partial α-helix. Likewise, all the programs predict *β*-sheet for m6 with the exception of PHDsec (*38*, *39*), which predicts both *β*-sheet and α-helix.

Characterization of Synthetic Peptides. As seen for peptide m1 in Figure 2, peptides were characterized by electrospray mass spectrometry, MEKE, and isocratic HPLC. Mass spectrometry confirmed the identity of each of the model synthetic peptides, as all six peptides had the expected mass values within experimental error. The synthetic peptides were also analyzed by fast atom bombardment mass spectrometry with similar results (data not shown). Purity of the model peptides was characterized by MEKE (Figure 2B). One major peak was obtained for each peptide under the conditions described. The purity of each peptide was estimated to be greater than 90%. The peptides were also estimated to be greater than 90% pure by isocratic HPLC (Figure 2C). Again, one major peak was obtained for each peptide by monitoring absorbance of eluate from a C-18 column at 270 nm.

Oligomeric State of Peptides in SDS. As seen in Figure 3, peptides dissolved in SDS were separated by Tricine−SDS−PAGE. Although a significant amount of band broadening took place as expected with these small synthetic peptides, the monomeric form of each peptide is visible near the bottom of their corresponding lanes with the exception of m2. The peptides m1 and m3 are easily discernable at molecular weights indicative of their presence only in monomeric states. Peptide m2 runs almost entirely at a molecular weight consistent with the formation of a dimeric species with a slight amount of apparently nonspecific aggregation present at higher molecular weights. However, the band present at the apparently dimeric molecular weight may correspond to an m2 monomer which runs anomalously under micelle-forming conditions.

In addition, peptides m6 and, to a slightly lesser extent, m5 also appear to self-associate and form specific higher molecular weight multimers. Peptide m5 runs primarily as an apparent 10-mer and m6 is detected at a molecular weight that would indicate an oligomeric species containing in excess of 20 peptides. Peptide m6 also forms a species whose molecular weight is in agreement with formation of a homodimer, but at a level much less than the apparent monomer or the much higher molecular weight oligomer. Although m4 is detected as a monomer, a portion of the peptide appears to aggregate nonspecifically in a manner similar to that of m2.

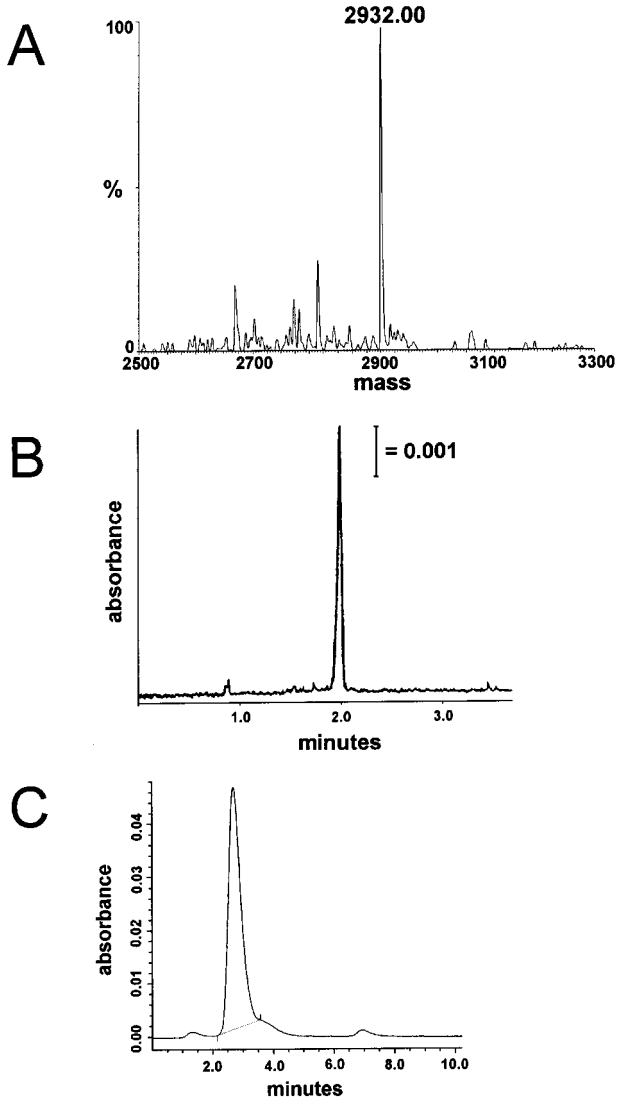

FIGURE 2: Characterization of synthetic peptides. The six synthetic peptides representing the predicted transmembrane spans of TMD1 were characterized by a series of analytical techniques as described in Experimental Procedures. Data from the characterization of peptide m1 is presented as a representative example and includes (A) electrospray mass spectroscopy; (B) micellar electrokinetic electrophoresis; and (C) isocratic high-performance liquid chromatography.

Spectroscopic Investigation of Peptide Secondary Structures. Several spectroscopic techniques were employed to study the secondary structures of each of the model transmembrane peptides. The far-ultraviolet absorption spectrum contains several features that can be used to distinguish conformations of proteins (*41*). Specifically, a red shift of

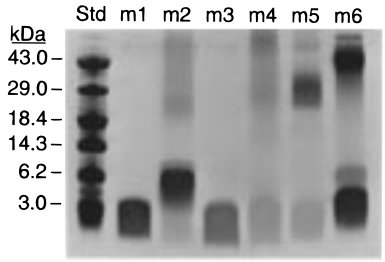

FIGURE 3: Tricine SDS−PAGE of synthetic peptides. The six synthetic peptides were further characterized by electrophoresis through an SDS−16.5% polyacrylamide gel utilizing a tricine buffer system. The peptides, dissolved in 0.5% SDS and 5 mM phosphate buffer (pH 7.2), were diluted with an equal volume of loading buffer and electrophoresed as described in Experimental Procedures. Each lane (m1−m6) represents 10 μg of the corresponding peptide except lane m4, where 30 μg was loaded.

the main observable peptide bond absorbance is associated with α to β secondary structure transitions. The ultraviolet absorption spectra of each of the TMD1 peptides dissolved in 0.5% SDS are shown in Figure 4. Each of the spectra contain features consistent with the presence of polypeptides in solution. It is intriguing to note that there is a red shift in the maximal intensity of the peptide bond absorbance of peptides m2 and m6, from 197 to 202 nm, relative to that of the other peptides. Peptide m3 exhibits a large unique absorbance shoulder at about 228 nm not seen in the other peptides.

The six TMD1 peptides were analyzed by CD spectroscopy in both SDS micelles and in 40% TFE/10% HFIP. CD spectra for all peptides (m1−m6) in both environments are presented in Figure 5. All of the spectra contain the classical characteristics of α-helical secondary structure with two minima, 209 and 222 nm, and a single maximum at about 195 nm. As summarized in Table 2, the helicity was calculated to be 50% or greater for all the peptides in both environments, with a range between 51% and 87%. In general, the peptides assume structures characterized by higher helical content in SDS micelles as opposed to 40% TFE/10% HFIP.

As seen in Figure 6, assessment of peptide secondary structures by ATR-FTIR in SDS micelles and 40% TFE/10% HFIP reveals a predominant secondary structural motif for each peptide consistent with that originally predicted by Riordan et al. (*1*) and in agreement with the CD data. Briefly, the dominant absorbance band in the amide I region of a protein FTIR spectrum associated with α-helix formation is found between 1654 and 1658 cm^{-1}. Absorbance bands at 1624, 1627, 1632, 1638, and 1642 cm^{-1} are associated with β-sheet and extended structures. An absorbance band at 1649−1651 cm^{-1} indicates unordered or random structures (*42*).

When integrated into detergent micelles, each of the peptides, with the exception of m2, appears to be predominantly α-helical. Peptide m2 has an α-helical component as well as a significant β-sheet contribution. In 40% TFE/10% HFIP, however, the peptides all assume conformations with significantly higher contributions of β-sheet secondary structure. These data correlate well with the CD % helicities summarized in Table 2. Peptides m2 and m6, having the highest % helicities by CD, appear to have the highest α-helical content by ATR-FTIR. Likewise, peptide m5 has the most β-sheet absorbance band contributions to its FTIR spectrum and also the lowest helicity by CD.

To assess the structures of m1−m6 in a lipid membrane, peptides were incorporated into PS/PC liposomes. To confirm the integrity of the liposomes, preparations were

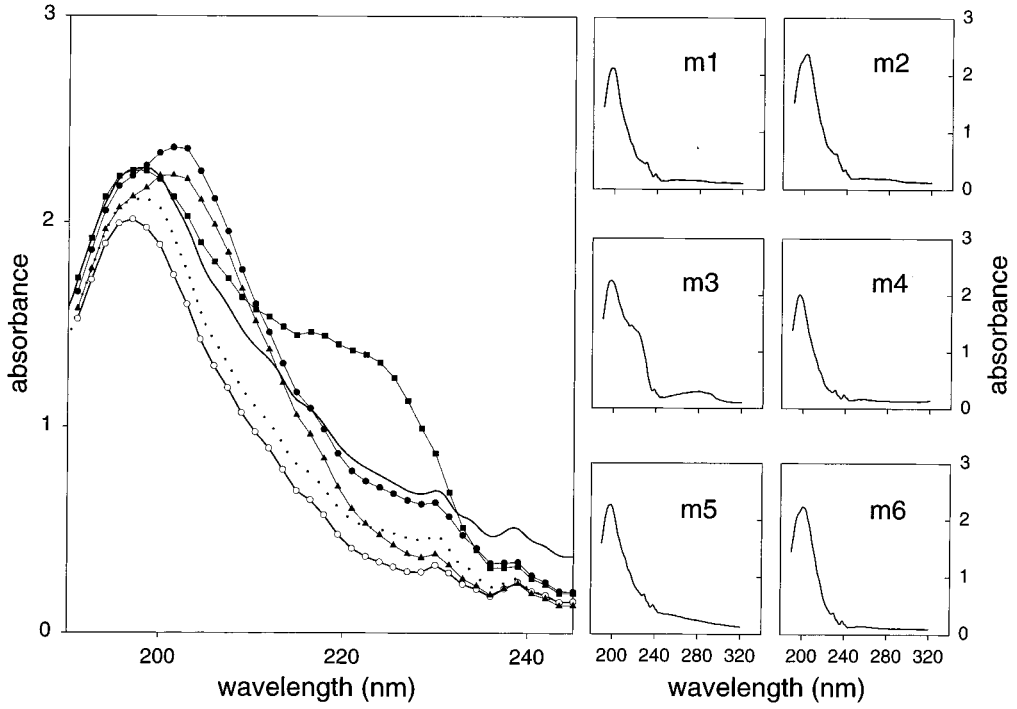

FIGURE 4: UV absorbance spectroscopy of synthetic peptides. Peptides m1−m6 were analyzed by UV spectroscopy from 190 to 320 nm. Peptides were dissolved in 0.5% SDS and 5 mM phosphate (pH 7.2) to 0.5 mg/mL and an absorbance spectrum was collected for an aliquot of each as described in Experimental Procedures. The entire UV spectrum collected of each peptide is shown in the small panels to the right as indicated. An enlarged portion of each spectrum from 190 to 240 nm is also presented in the large panel. (The symbols used to represent each peptide are as follows: m1, •; m2, ●; m3, ■; m4, ○; m5, —; m6, ▲).

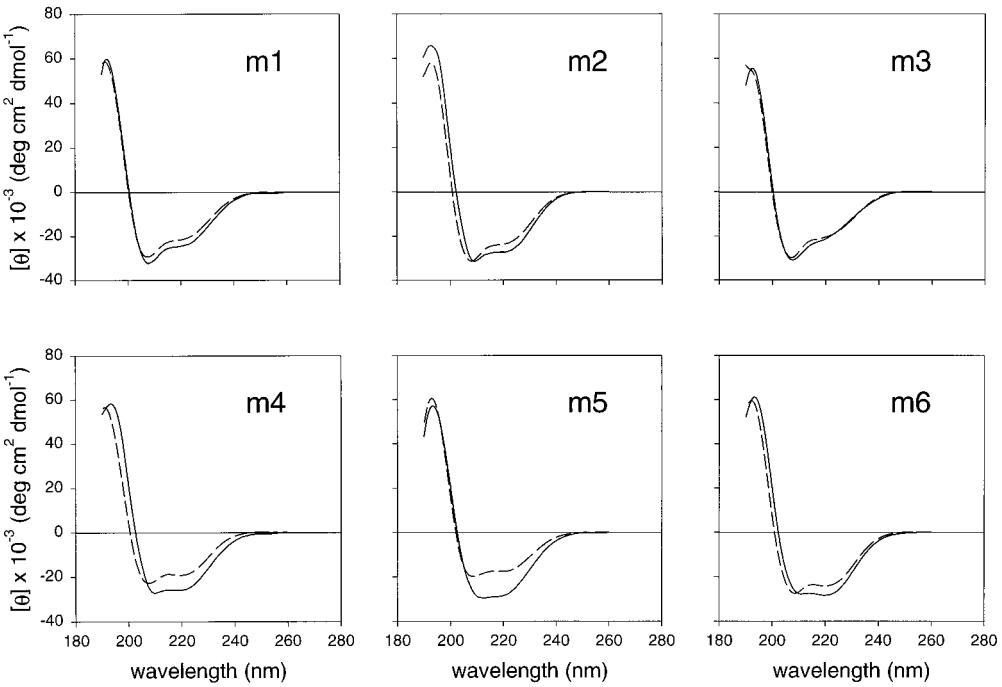

FIGURE 5: CD spectroscopy of synthetic peptides. The secondary structures of the six synthetic peptides (m1−m6) were analyzed by CD spectroscopy as described in Experimental Procedures. In each of the six panels, the solid line represents the indicated peptide dissolved in 0.5% SDS and 5 mM phosphate (pH 7.2) and the dashed line represents the peptide dissolved in 10% HFIP and 40% TFE. Each spectrum represent the smoothed average of three scans.

Table 2

	CD		ATR-FTIR		
peptide	0.5% SDS	40%TFE/10% HFIP	liposomes	0.5% SDS	40%TFE/10% HFIP
m1	73	63	α	α	α-*β*
m2	83	71	α	α-*β*	α-*β*
m3	60	59	α	α	α-*β*
m4	78	59	α	α	α-*β*
m5	85	51	α	α	α-*ββ*
m6	87	73	α	α	α-*β*

[a] Peptide percent helicities were calculated from the mean molar per residue ellipticities at 222 nm as described under Experimental Procedures. A purely qualitative assessment of the predominant peptide secondary structural contributions is presented for the ATR-FTIR data. (The symbolism is as follows: α denotes predominantly α-helix; α-*β* denotes α-helix with significant *β*-sheet and other contributions; α-*ββ* denotes α-helix with strong *β*-sheet and other contributions.)

negatively stained with 1.0% uranyl acetate and visualized by electron microscopy as described under Experimental Procedures. Uniformly shaped liposomes were observed in both peptide-reconstituted and peptide-minus controls. Notably, peptide solubilization required the presence of liposomes, consistent with peptide incorporation. The sizes of the resulting liposomes ranged from between about 75 and 100 nm in diameter (data not shown).

We were unable to obtain UV or CD spectra for the peptides reconstituted into liposomes due to excessive light scattering at the required concentrations. However, as light scattering is not a concern with ATR-FTIR, we employed this method to obtain additional information about the lipid-reconstituted peptides. All six peptides assume predominantly α-helical structures when reconstituted into liposomes (Figure 6).

CD and ATR-FTIR of m6 in Methanol. Peptide m6 is soluble in 100% methanol and in dilutions as low as 20% methanol in water. The CD spectrum of m6 in 20% methanol (Figure 7A) contains features generally indicative of *β*-sheet protein secondary structure: a single minimum at 218 nm and a maximum at about 195 nm, both of which are of a reduced magnitude relative to α-helices. Therefore, m6 assumes a *β*-sheet conformation in methanol. This is in contrast to its conformation in the other three environments examined.

This conformation was confirmed by ATR-FTIR of m6 in 20% methanol (Figure 7B). Intensity of the peak at 1624 cm^{-1}, indicative of *β* secondary structure, is significantly intensified relative to the data in polyfluorinated alcohol solvents and detergent or liposome solutions. In addition, two strong absorbances representative of turns (*42*) are present at 1680 and 1690 cm^{-1}. Finally, the very low absorbance observed at 1650 cm^{-1} indicates that the unordered or random structural contribution to the IR spectrum is minimal. The amplified signals at 1624, 1680, and 1690 cm^{-1} are consistent with the presence of a highly ordered array of *β*-strands that form a large sheet of m6 peptides.

DISCUSSION

Several lines of evidence suggest that important structural information, currently obscure to more traditional approaches, regarding transmembrane regions may be obtained by studying the structures and interactions of model peptides. Synthetic peptides modeling transmembrane segments of the integral membrane protein bacteriorhodopsin fold into native conformations (*24*). In addition, membrane-spanning peptide sequences from glycophorin A form stable dimers in both detergent micelles and lipid bilayers (*26*, *27*). In another example, several segments from the hydrophobic core of the Shaker K^+ channel were modeled by synthetic peptides (*43*) that interact specifically with one another but not with a homologous segment of another channel. These observations

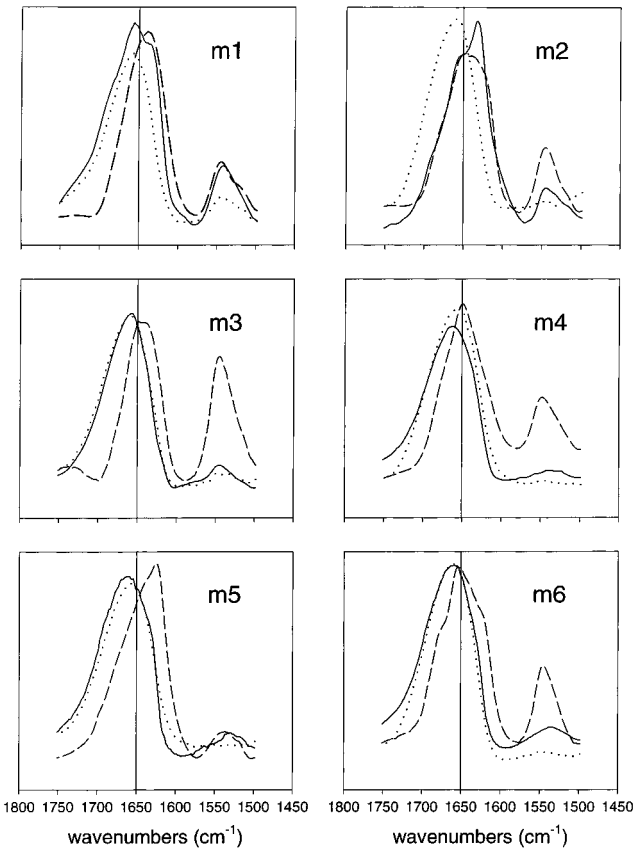

FIGURE 6: ATR-FTIR spectroscopy of synthetic peptides. The secondary structures of the six synthetic peptides (m1−m6) were analyzed by ATR-FTIR spectroscopy as described in Experimental Procedures. In each of the six panels, the solid line represents the indicated peptide dissolved in 0.5% SDS and 5 mM phosphate (pH 7.2), the dashed line represents the peptide dissolved in 10% HFIP and 40% TFE, and the dotted line represents the peptide incorporated into synthetic phospholipid vesicles. Each spectrum is the result of averaging 250 individual interferograms. The vertical line at 1650 cm^{-1} dividing the α-helical (1654−1658 cm^{-1}) and extended/β-sheet absorbance regions (1624−1642 cm^{-1}) is included as a visual aid to the reader.

demonstrate that at least some of the information necessary for the formation of native conformations as well as higher-order structures in membranes is contained in the model peptide sequences. Synthetic peptides are therefore useful to investigate not only structure, but also assembly and folding of organized membrane regions of proteins.

The problem of studying membrane protein folding and structure can, at its most basic level, be distilled into two stages: determination of the secondary structure of spanning sequences and subsequent description of their two-dimensional arrangement in the plane of the bilayer. The studies described herein support this two-stage model of membrane protein folding (*44*). In the first stage, individual transmembrane helices are formed followed by insertion into the bilayer. Significantly, each of the model peptides forms a stable structure in three different environments in the absence of the remainder of the protein. Only as additional helices enter the bilayer does the second stage of helix association occur. Evidence indicates that, once in the bilayer, helical associations are stabilized by tight packing of the interface resulting in favorable van der Waals contacts (*44*, *28*). Using the system established here, further studies will focus on elucidating these possible helix−helix interactions within the membrane.

Of course in the cell the situation is more complex. Recently, evidence has been reported indicating that the translocon may play a critical role in presenting these spanning segments to the membrane (*45*, *46*). In addition, the ribosome may participate in membrane integration through a transmembrane domain recognition and subsequent translocon signaling pathway (*47*, *48*).

Synthetic model peptides have also been used to address functional questions about multipass transmembrane regions. In addition to their ability to assume native conformations, fragments corresponding to the transmembrane segments of bacteriorhodopsin were shown to associate and form functional proton pumps (*49*). The relative orientations of these peptides were not constrained by the extramembrane connecting loops, thus supporting the idea of transmembrane sequence determination of conformation and interactions within the bilayer. In CFTR, peptides representing the second and sixth membrane-spanning segments of TMD1 are able to form anion-specific channels, and apparently interact, after spontaneous insertion into membranes (*30*).

We therefore predicted that peptides corresponding to putative transmembrane segments of CFTR would be a suitable model system of TMD folding and structure in the membrane. An additional advantage of this system is the large number of mutations that have been introduced into TMD1 by nature, allowing the investigation of a variety of structure and function relationships associated with the CF pathology. We decided to focus on the first transmembrane domain of CFTR, containing peptides m1−m6, for the studies described herein. It was assumed that this region contained the putative CFTR channel pore-forming segments because expression of the amino-terminal half of CFTR alone, truncated between the R domain and TMD2, is sufficient to form a functional, ATP-dependent chloride channel (*21*).

Several of the peptides apparently form higher molecular weight oligomers when solubilized in SDS. The tricine−SDS−polyacrylamide gel presented in Figure 3 shows peptides m1 and m3 present only in forms consistent with monomers. Experiments involving mixing and gel electrophoresis of all possible combinatorial pairs of micelle associated peptides had no discernible effect on the apparent molecular weights of the species observed (data not shown). This observation indicates that, under the conditions described here, there is no evidence for interactions at quite high peptide concentrations (200 μM) in detergent micelles. The apparently oligomeric species formed under these conditions by peptides m2, m5, and m6 are of well-defined molecular masses and differ significantly from the nonspecific aggregation that is present at higher molecular weights as a smear at the top of lanes m2 and m4. In addition, this apparent multimerization phenomenon appears to persist at even lower concentrations. Dilution by as much as 10-fold has no apparent effect on the multimer formation (data not shown).

Equal amounts of protein were loaded into each lane in Figure 3 with the exception of m4, which was loaded at 3-fold higher concentration due to poor staining. Staining with Coomassie blue is not necessarily proportional to protein concentration; a similar staining differential was also ob-

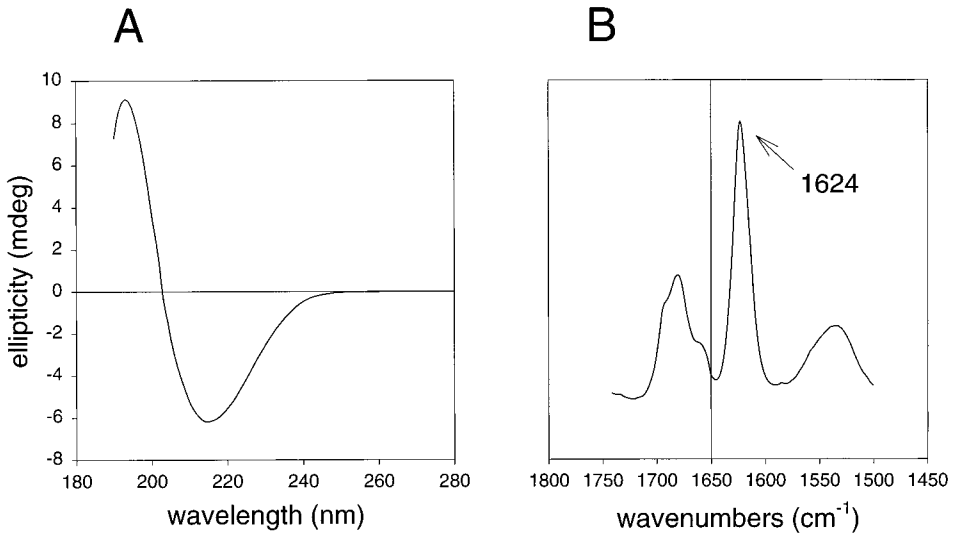

FIGURE 7: CD and ATR-FTIR of peptide m6 in 20% methanol. Synthetic peptide m6, dissolved in 20% methanol to 0.5 mg/mL, was analyzed by (A) CD spectroscopy and (B) ATR-FTIR spectroscopy as described in Experimental Procedures. The maximum spectral absorbance frequency at 1624 cm^{-1} is indicated by an arrow. The vertical line is included as in the legend to Figure 6.

served between the α and β subunits of the F_1 ATPase, yielding stain intensity ratios as high as 4 to 1 even after normalization for differences in molecular weight (*50*).

Two significant spectral features are apparent in the UV absorption spectra of the peptides in SDS micelles (Figure 4). First, the red shift of 5 nm in the peptide bond absorbance maxima of m2 and m6 indicates the presence of an increased proportion of β-structure in these peptides (*41*), perhaps related to the tendency of these peptides to self-associate in solution. Recall that m2 and m6 are the two peptides that form anion-specific channels in planar lipid bilayers which has been taken to indicate that these regions may form part of the CFTR channel (*30*). Moreover, cysteine scanning suggests a β-strand-like labeling pattern in the terminal end of m6 (*6*). Second, the absorbance shoulder at about 228 nm seen in the spectrum of m3 is characteristic of ring-charge interactions (*51*), perhaps involving one of the two tryptophan residues present in this 24 amino acid peptide. Thus, m3 may not form a simple helix. It is interesting to note that m3 contains a central proline residue, which would introduce a bend or kink into its transmembrane α-helix (*52*). Mutational analysis of proline residues in the lutropin/choriogonadotropin receptor transmembrane helices demonstrated the importance of these intramembrane proteins for both hormone binding and cell surface expression (*53*). The bend at proline 205 may bring one or both m3 Trp residues proximal to a charged residue. The critical importance of this intramembrane residue to CFTR structure is highlighted by the CF-causing mutation of proline 205 to serine (*54*), which prevents proper folding and processing of CFTR (*52*) and causes a form of cystic fibrosis similar to other misprocessing mutants such as A455E and P574H (*55*).

One question we hoped to address was the validity and applicability of several secondary structure prediction methods toward structures in various environments. It is interesting to compare these predictions with experimental secondary structural evaluations. In Figure 1, schematics representing predictions of secondary structure made using several widely available prediction routines are presented. Notably, there appears to be very little consensus among the five prediction methods employed here except with regard to m3 (predicted to be α-helical in four of five with partial α-character in the fifth), m6 (predicted to be a β-strand in four of five with partial β-character in the fifth), and to some extent m5 (predicted to be primarily β in conformation).

One of the most widely used prediction methods, that of Chou and Fasman (*40*), which predicts secondary structure based upon statistically calculated propensities of residues in soluble proteins to form a particular structure, predicts all of the peptides to adopt β secondary structure with the lone exception being m3, predicted to be α/β. This is not a surprise, as it has been clearly demonstrated that the hydrophobicity of the individual residue defines the probability of adopting α-helical conformations in hydrophobic environments such as the membrane-mimetic SDS micelle (*24*, *56*). The current findings further support these observations. In direct comparisons with experimental results presented in the current study, the predictions of PHDsec (*38*, *39*)(Figure 1) correlate most closely to our experimentally determined TMD1 secondary structures (Table 2).

As predicted, all six peptides are primarily α-helical in polyfluorinated organic solvents, membrane-mimetic detergent micelles, and phospholipid membranes as demonstrated by both CD and ATR-FTIR (Figures 5 and 6, Table 2). These experimental data agree well with the predictions of transmembrane helices made in the working model proposed by Riordan et al. (*1*) and during the design stage of this project, a summary of which is presented in Table 1. However, the similarities between secondary structures of the peptides in membrane-mimetic detergent micelles and in liposomes when contrasted with the increase in β-sheet contributions observed in the polyfluorinated alcohol solvent system underscore the importance and influence of the environment on the structure of transmembrane sequences. These results are consistent with the notion that the structures in polyfluorinated alcohols reflect the propensity of a sequence to assume a secondary structure and the structure, in membranes or membranelike environments reflects the relevant native secondary structure (*57*, *58*).

This has implications for the effects of environment on the secondary structure of proteins. For example, the phospholipid phosphatidylethanolamine has recently been

shown to be a required chaperone in the assembly of lactose permease (*59*). Lipid-mediated effects on structure may also have implications for regulation of function, perhaps through the formation of homogeneous microdomains or highly concentrated lipid "rafts". Interactions between protein and phospholipids in these "rafts" have been implicated in the activation of ARF1 by a guanine nucleotide exchange factor (*60*). In addition, the formation of dioleoylglycerol-rich membrane domains has been suggested to activate protein kinase C (*61*).

In this regard, the ability of peptide m6 to assume multiple stable environmentally dependent secondary structures may have relevance and is not without precedent. One example involves the conformational shift of the α-helical normal cellular prion protein (PrP^{c}) into the infectious scrapie prion (PrP^{Sc}), which is highly β-sheet in secondary structure and is responsible for prion propagation (*62*). In another example, studies suggest the peptide β/A4, derived from the inappropriate proteolysis of a membrane-spanning region of the Alzheimer's precursor protein, adopts an antiparallel β-sheet secondary structure (*63*−*65*). The observation that m6 is primarily β-sheet in 20% methanol (Figure 7) rather than α-helical as observed in the membrane-mimetic detergent and polyfluorinated alcohol organic environments may have many important implications for the folding of CFTR.

Evidence has been reported supporting the hypothesis that m6 forms a portion of the channel lining. For example, mutational data demonstrated that the substitution of glutamic acid for lysine 335 alters channel permeability of CFTR (*4*), and mutation of arginine 347 to histidine gives rise to pH-dependent ion selectivity of the channel (*66*). Thus, the residue R347 may act as an anion selectivity filter. Substituted cysteine accessibility studies conducted on the sixth transmembrane-spanning region identified several solvent-exposed residues, including R347, implicating them as forming a portion of the lining of the CFTR pore (*6*). The pattern of residues modified indicated both α-helical and extended structure in the m6 transmembrane region. Therefore, a change of the conformation in this crucial region of CFTR, such as an α-helix to β-sheet, could play a direct role in anion selectivity or even act as a switch for gating of the channel. In any case, it is important to consider these and other possibilities for this highly environmentally sensitive membrane-spanning segment.

ACKNOWLEDGMENT

We thank the members of the Thomas laboratory for advice and helpful discussions, Steve Madden for excellent technical assistance, Bikash Pramanik for mass spectral assistance, and Imma Fernandez for critical review of the manuscript.

REFERENCES

1. Riordan, J. R., Rommens, J. M., Kerem, B.-S., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.-L., Drumm, M. L., Iannuzzi, M. C., Collins, F. S., & Tsui, L.-C. (1989) *Science 245*, 1066−1073.
2. Zeng, W., Lee, M. G., Yan, M., Diaz, J., Benjamin, I., Marino, C. R., Kopito, R., Freedman, S., Cotton, C., Muallem, S., & Thomas, P. (1997) *Am. J. Physiol. 273*, C442−C455.
3. Kartner, N., Augustinas, O., Jensn, T. J., Naismith, A. L., & Riordan, J. R. (1992) *Nat. Genet. 1*, 321−327.
4. Anderson, M. P., Gregory, R. J., Thompson, S., Souza, D. W., Paul, S., Mulligan, R. C., Smith, A. E., & Welsh, M. J. (1991) *Science 253*, 202−205.
5. Linsdell, P., & Hanrahan, J. W. (1996) *J. Physiol. 496*, 687−693.
6. Cheung, M., & Akabas, M. H. (1996) *Biophys. J. 70*, 2688−2695.
7. Schwiebert, E. M., Egan, M. E., Hwang, T.-H., Fulmer, S. B., Allen, S. S., Cutting, G. R., & Guggino, W. B. (1995) *Cell 81*, 1063−1073.
8. Stutts, M. J., Rossier, B. C., & Boucher, R. C. (1997) *J. Biol. Chem. 272*, 14037−14040.
9. McNicholas, C. M., Guggino, W. B., Schwiebert, E. M., Hebert, S. C., Giebisch, G., & Egan, M. E. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 8083−8088.
10. Egan, M. E., Schwiebert, E. M., & Guggino, W. B. (1995) *Am. J. Physiol. 268*, C243−C251.
11. Higgins, C. F. (1992) *Annu. Rev. Cell Biol. 8*, 67−113.
12. Cheng, S. H., Rich, D. P., Marshall, J., Gregory, R. J., Welsh, M. J., & Smith, A. E. (1991) *Cell 66*, 1027−1036.
13. Anderson, M. P., Berger, H. A., Rich, D. P., Gregory, R. J., Smith, A. E., & Welsh, M. J. (1991) *Cell 67*, 775−784.
14. Ko, Y. H., & Pedersen, P. L. (1995) *J. Biol. Chem. 270*, 22093−22096.
15. Gunderson, K. L., & Kopito, R. R. (1994) *J. Biol. Chem. 269*, 19349−19353.
16. Gunderson, K. L., & Kopito, R. R. (1995) *Cell 82*, 231−239.
17. Baukrowitz, T., Hwang, T. C., Nairn, A. C., & Gadsby, D. C. (1994) *Neuron 12* (3), 473−482.
18. Li, C., Ramjeesingh, M., Wang, W., Garami, E., Hewryk, M., Lee, D., Rommens, J. M., Galley, K., & Bear, C. E. (1996) *J. Biol. Chem. 271*, 28463−28468.
19. Chang, X.-B., Hou, Y.-X., Jensen, T. J., & Riordan, J. R. (1994) *J. Biol. Chem. 269*, 18572−18575.
20. McDonough, S., Davidson, N., Lester, H. A., & McCarty, N. A. (1994) *Neuron 13*, 623−634.
21. Sheppard, D. N., Ostedgaard, L. S., Rich, D. P., & Welsh, M. J. (1994) *Cell 76*, 1091−1098.
22. Akabas, M. H., Kaufmann, C., Cook, T. A., & Archdeacon, P. (1994) *J. Biol. Chem. 269*, 14865−14868.
23. Shai, Y. (1995) *Trends Biochem. Sci. 20*, 460−464.
24. Hunt, J. F., Earnest, T. N., Bousché, O., Kalghatgi, K., Reilly, K., Horvath, C., Rothschild, K. J., & Engelman, D. M. (1997) *Biochemistry 36*, 15156−15176.
25. Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckman, E., & Downing, K. H. (1990) *J. Mol. Biol. 213*, 899−929.
26. Engelman, D. M., Steitz, T. A., & Goldman, A. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 321−353.
27. Adair, B. D., & Engelman, D. M. (1994) *Biochemistry 33*, 5539−5544.
28. MacKenzie, K. R., Prestegard, J. H., & Engelman, D. M. (1997) *Science 276*, 131−133.
29. Doak, D. G., Mulvey, D., Kawaguchi, K., Willalain, J., & Campbell, I. D. (1996) *J. Mol. Biol. 258* (4) 672−687.
30. Oblatt-Montal, M., Reddy, G. L., Iwamoto, T., Tomich, J. M., & Montal, M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1495−1499.
31. Schägger, H., & von Jagow, G. (1987) *Anal. Biochem. 166*, 368−379.
32. Jackson, M. L., & Litman, B. J. (1985) *Biochim. Biophys. Acta 812*, 369−376.
33. Jones, J. D., Bourell, K. W., Norgard, M. V., & Radolf, J. D. (1995) *Infect. Immun. 63*, 2424−2434.
34. Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
35. Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 8*, 4108−4116.
36. Ben-Efraim, I., & Shai, Y. (1997) *Biophys. J. 72*, 85−96.
37. Eisenberg, D., Schwarz, E., Komaramy, R., & Wall, R. (1984) *J. Mol. Biol. 179*, 125−142.
38. Rost, B., & Sander, C. (1993) *J. Mol. Biol. 232*, 584−599.

39. Rost, B., & Sander, C. (1994) *Proteins: Struct., Funct., Genet. 19*, 55−72.
40. Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45−148.
41. Rosenheck, K., & Doty, P. (1961) *Proc. Natl. Acad. Sci. U.S.A. 47*, 1775−1785.
42. Dong, A., Huang, P., & Caughey, W. S. (1990) *Biochemistry 29*, 3303−3308.
43. Peled-Zehavi, H., Arkin, I. T., Engelman, D. M., & Shai, Y. (1996) *Biochemistry 35*, 6828−6838.
44. Lemmon, M. A., & Engelman, D. M. (1994) *Q. Rev. Biophys. 27*, 157−218.
45. Mothes, W., Heinrich, S. U., Graf, R., Nilsson, I., von Heiijne, G., Brunner, J., & Rappaport, T. A. (1997) *Cell 89*, 523−533.
46. Hamman, B. D., Chen, J.-C., Johnson, E. E., & Johnson, A. E. (1997) *Cell 89*, 535−544.
47. Siegel, V. (1997) *Cell 90*, 5−8.
48. Liao, S., Lin, J., Do, H., & Johnson, A. E. (1997) *Cell 90*, 31−41.
49. Kahn, T. W., & Engelman, D. M. (1992) *Biochemistry 31*, 6144−6151.
50. Catterall, W. A., & Pedersen, P. L. (1971) *J. Biol. Chem. 246*, 4987−4994.
51. Clark, P. L., Liu, Z.-P., Zhang, J., & Gierasch, L. M. (1996) *Protein Sci. 5*, 1108−1117.
52. Sheppard, D. N., Travis, S. M., Ishihara, H., & Welsh, M. J. (1996) *J. Biol. Chem. 25*, 14995−15001.
53. Hong, S., Ryu, K.-S., Oh, M.-S., Ji, I., & Ji, T. H. (1997) *J. Biol. Chem. 272*, 4166−4171.
54. Chillón, M., Casals, T., Nunes, V., Giménez, J., Ruiz, E. P., & Estivill, X. (1993) *Hum. Mol. Genet. 2*, 1741−1742.
55. Sheppard, D. N., Ostedgaard, L. S., Winter, M. C., & Welsh, M. J. (1995) *EMBO J.* 14, 876−883.
56. Li, S.-C., & Deber, C. M. (1994) *Nat. Struct. Biol. 1*, 368−373.
57. Bruch, M. D., & Gierasch, L. M. (1990) *J. Biol. Chem. 265*, 3851−3858.
58. Rizo, J., Blanco, F. J., Kobe, B., Bruch, M. D., & Gierasch, L. M. (1993) *Biochemistry 32*, 4881−4894.
59. Bogdanov, M., Sun, J., Kaback, H. R., & Dowhan, W. (1996) *J. Biol. Chem. 271*, 11615−11618.
60. Paris, S., Berand-Dufour, S., Robineau, S., Bigay, J., Antonny, B., Chabre, M., & Chardin, P. (1997) *J. Biol. Chem. 272*, 22221−22226.
61. Dibble, A. R. G., Hinderliter, A. K., Sando, J. J., & Biltonen, R. L. (1996) *Biophys. J. 71*, 1877−1890.
62. Pan, K.-M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R. J., Cohen, F. E., & Prusiner, S. B. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10962−10966.
63. Kang, J., Lemaire, H.-G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K.-H., Multhaup, G., Beyreuther, K., & Muller-Hill, B. (1987) *Nature 325*, 733−736.
64. Hilbich, C., Kisters, W. B., Teed, J., Masters, C. L., & Beyreuther, K. (1991) *J. Mol. Biol. 218*, 149−163.
65. Wood, S. J., Wetzel, R., Martin, J. D., & Hurle, M. R. (1995) *Biochemistry 34*, 724−730.
66. Tabcharani, J. A., Rommens, J. M., Hou, Y.-X., Chang, X.-B., Tsui, L.-C., Riordan, J. R., & Hanrahan, J. W. (1993) *Nature 366*, 79−82.
67. Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol. 120(1)*, 97−120.
68. Solovyev, V. V., & Salamov, A. A. (1994) *Comput. Appl. Biosci. 10* (6), 661−669.
69. Kneller, D. G., Cohen, F. E., & Langridge, R. (1990) *J. Mol. Biol. 214*, 171−182.
70. McClelland, J. L., & Rumelhart, D. E. (1988) *Explorations in Parallel Distributed Processing*, Vol 3; pp 318−362, MIT Press, Cambridge, MA.

BI972293N