Topogenesis of Cystic Fibrosis Transmembrane Conductance Regulator (CFTR): Regulation by the Amino Terminal Transmembrane Sequences†

Mingang Chen‡ and Jian-Ting Zhang*,§

Department of Pharmacology and Toxicology, IU Cancer Center and Walther Oncology Center, Indiana University School of Medicine, Indianapolis, Indiana 46202

Received September 8, 1998; Revised Manuscript Received March 2, 1999

ABSTRACT: Cystic fibrosis transmembrane conductance regulator (CFTR) is a member of the ATP-binding cassette (ABC) transport superfamily. CFTR folding and assembly appear to involve several events occurred in the cytosol and ER. Misfolding of CFTR causes cystic fibrosis, and thus, understanding the folding mechanism of CFTR is extremely important. Recently, detailed study of membrane insertion process suggests that the first two transmembrane (TM) segments of CFTR have two distinct but independent mechanisms to ensure the correct membrane folding of its amino terminal end [Lu, Y., Xiong, X., Helm, A., Kimani, K., Bragin, A., Skach, W. R. (1998) *J. Biol. Chem. 273*, 568−576]. To understand how other TM segments are ensured to insert into membranes correctly, we investigated the topogenesis of TM3 and TM4 of CFTR in a cell-free expression system. We found that the correct membrane insertion of TM3 and TM4 of CFTR was ensured by their flanking amino acid sequences and controlled by the correct membrane insertion of their preceding TM1 and TM2. Thus, correct membrane insertion and folding of TM1 and TM2 play an essential role in the membrane insertion and folding of the subsequent TM segments of CFTR.

Integral membrane proteins that have one or multiple transmembrane (TM)[1] segments are known as bitopic and polytopic membrane proteins, respectively. They acquire their membrane topology during or immediately after synthesis on the rough endoplasmic reticulum (ER). For bitopic membrane proteins, "positive inside rule" has been used to explain the formation of topology. It was predicted that these membrane proteins orient themselves with the most positively charged end in the cytoplasm (*1*). For polytopic membrane proteins, it was suggested that only the charged amino acids flanking the first TM segment are important in determining the overall membrane topology of the protein. The remaining TM segments simply follow the lead of the first one (*2*, *3*).

ATP-binding cassette (ABC) membrane transporters are polytopic membrane proteins that transport a wide variety of structurally unrelated substrates. Members of this superfamily of transporters include cystic fibrosis transmembrane conductance regulator (CFTR) and P-glycoprotein (Pgp) (*4*, *5*). The topological folding of both CFTR and Pgp has been investigated. More than one topological folding was found with Pgp (*6*−*9*) but not with CFTR (*10*, *11*). One of the variations found in Pgp topology is located in the region including TM3 and TM4 (*12*). Positively charged residues in the loop between TM3 and TM4 have been shown to regulate the membrane folding of TM3 and TM4 in Pgp (*13*). Furthermore, dissection of topological signal in Pgp suggested that Pgp TM4 had high de novo signal-anchor activity to initiate membrane insertion which also contributed to the variation in Pgp topology (*14*).

	Charge N	C	C-N				
PGP	+2	0	-2	TM3	TRGWKLT	TM4	K I L S S FTDKE LQAYA
CFTR	-1	+4	+5	TM3	ELLQ	TM4	RMMMKYRDQRAGKIS
E217R Q220K	+2	+4	+2	TM3	R--K	TM4	----------
R242E	-1	+2	+3	TM3	----	TM4	E----------
K246N R248E	-1	+1	+2	TM3	----	TM4	----N-E-------
R251E	-1	+2	+3	TM3	----	TM4	---------E-----

FIGURE 1: Sequence comparison between CFTR and Pgp. Amino acid sequences flanking TM4 of Pgp and CFTR are shown in single-letter code. The net charge difference (C−N) across TM4 was calculated by subtracting the charges at the amino terminal side (N) from that of the carboxyl terminal side (C) of TM4. The dashed line indicates amino acid residues that are the same as in the wild-type CFTR.

Although CFTR and Pgp have similar hydropathy-predicted topology, the folding mechanism for their topology may be different. Sequence comparison showed that distribution of the charged residues in the domains flanking TM4 was different between CFTR and Pgp (Figure 1). This difference may be responsible for the topogenesis difference between CFTR and Pgp. Recently, Lu et al. (15) showed

† This work was supported by grants (CA64539 and GM59475) from the National Institutes of Health.

* To whom correspondence should be addressed: IU Cancer Research Institute, 1044 W. Walnut St., R4−Room 166, Indiana University School of Medicine, Indianapolis, IN 46202. Phone: (317) 278-4503. Fax: (317) 274-7592.

‡ Present address: Department of Medical Biochemistry and Genetics, 116 Reynold Medical Bldg., Texas A&M University, College Station, TX 77843.

§ Recipient of a Career Investigator Award from the American Lung Association.

[1] Abbreviations: CFTR, cystic fibrosis transmembrane conductance regulator; TM, transmembrane; ABC, ATP-binding cassette; Pgp, P-glycoprotein; PCR, polymerase chain reaction; RRL, rabbit reticulocyte lysate.

10.1021/bi982153t CCC: $18.00
Published on Web 04/06/1999

that CFTR has two independent but redundant membrane translocation pathways to ensure correct topological folding for TM1 and TM2. In this study, we analyzed the biogenesis of TM3 and TM4 of CFTR and tested whether the charged residues flanking TM segments are responsible for the topogenesis difference between CFTR and Pgp. We found that the topogenesis of TM3 and TM4 of CFTR is determined by their flanking amino acid sequences and is controlled by the correct folding of their preceding TM1 and TM2.

MATERIALS AND METHODS

Materials. pGEM-4z, Taq DNA polymerase, SP6 and T7 RNA polymerases, RNasin, ribonucleotides, deoxyribonucleotides, RQ1 DNase, rabbit reticulocyte lysate (RRL), and dog pancreatic microsomal membranes were obtained from Promega. [^{35}S]Methionine and Amplify were purchased from DuPont/NEN and Amersham, respectively. m^{7}G(5′)ppp(5′)G cap analogue was obtained from Pharmacia Biotech. Peptide N-glycosidase F was purchased from Boehringer Mannheim and restriction endonucleases were obtained from either GIBCO-BRL or New England Biolab. All other chemicals were obtained from Sigma or Fisher Scientific.

Mutant Construct Engineering. Site-directed mutagenesis in CFTR-N4R construct (*11*) was performed using two polymerase chain reactions (PCR). In the first PCR, primers carrying mutations and a universal primer (T7 or SP6) and CFTR-N4R template were used. PCR products were isolated and used as a primer, together with another universal primer (SP6 or T7), to amplify the remaining sequence in CFTR-N4R. The second PCR product was digested with *Xba*I and *Bgl*II and the *Xba*I−*Bgl*II cassettes carrying mutations were then used to replace the *Xba*I−*Bgl*II fragment in wild-type CFTR-N4R DNA. Primers carrying various mutations are 5′-AATCTGGAGGTTGTTAAAGGCGTC-3′ (E217R/Q220K), 5′-CATCATTTCCCCTAGCCC-3′ (R242E), 5′-CTGATCTTCGTAATTCATCATCAT-3′ (K246N/R248E), and 5′-TCCCAGCTTCCTGATCT-3′ (R251E). To make the CFTR-N4R(−8) and CFTR-N4R(−5) mutants, two PCR reactions were performed as described above using CFTR-N4R(E217R/Q220K) or wild-type CFTR-N4R as templates and the mutant primer 5′-AGGGGAAATGATGATGGAGTACGAAGATCAGGAAGCTGGGGATATCAG-3′. The resulting constructs were named CFTR-N4R(E217R/Q220K), CFTR-N4R(R242E), CFTR-N4R(K246N/R248E), CFTR-N4R(R251E), CFTR-N4R(−8), and CFTR-N4R(−5).

To remove TM1 and TM2 from CFTR-N4R, CFTR-N4R(E217R/Q220K), CFTR-N4R(−8), and CFTR-N4R(−5), PCR was performed using a primer 5′-GAGACCATGCAGATGAGAATAG-3′ containing Kozak translation initiation codon and a primer 5′-CACTTTTGCCAACCAG-3′ in the glycosylation reporter sequence on templates CFTR-N4R, CFTR-N4R(E217R/Q220K), CFTR-N4R(−8), and CFTR-N4R(−5), respectively. The PCR products were cloned and propagated using a TA cloning kit (Invitrogen). The cDNA fragments were then released by *Eco*RI and *Bgl*II digestion and ligated into a CFTR-N4R vector containing a reporter cDNA digested with *Eco*RI and *Bgl*II. The final DNA clones were named CF-TM3,4R, CF-TM3,4R(E217R/Q220K), CF-TM3,4R(−8), and CF-TM3,4R(−5), respectively. To engineer R242E and K246N/R248E mutations into CF-TM3,4R, an *Eco*RI-*Eco*NI fragment encoding TM3 and part of TM4 was released from CF-TM3,4R and used to replace the amino terminal-encoding sequence in CFTR-N4R(R242E) and CFTR-N4R(K246N/R248E). The resulting constructs were named CF-TM3,4R(R242E) and CF-TM3,4R(K246N/R248E), respectively.

To replace TM1 and TM2 of CFTR with that of Pgp, CFTR-N4R(−8) was linearized with *Xba*I, treated with Klenow DNA polymerase supplemented with dNTP in the absence of dGTP to avoid filling at the G position, and then digested with *Hin*dIII. The *Xba*I−*Hin*dIII fragment from CFTR-N4R(−8) and an *Eco*RI-*Nci*I fragment encoding TM1 and TM2 of Pgp from pGPGP-N3 (12) were ligated into a pGEM-4z vector digested with *Eco*RI and *Hin*dIII. The resulting construct was named P1,2CF3,4R(−8).

To replace the loop between TM3 and TM4 of CFTR with that of Pgp, we did two PCR reactions. In the first reaction, the universal T7 primer and the specific primer 5′-CCACAGAAGGCAGACGCGGTTAGCTTCCAGCCTCGAGTCCAGATTAGCCCC-3′ (the underlined sequence encodes the Pgp loop linking TM3 and TM4) were used to amplify CF-TM3,4R and CF-TM3,4R(−8). The PCR products were then used as primers, together with the SP6 primer to amplify the remaining sequence of CF-TM3,4R and CF-TM3,4R(−8). These new constructs were named CF3PL4R and CF3PL4R(−8), respectively. All constructs were sequenced to confirm the designed mutations and to eliminate possible errors generated during DNA manipulation.

In Vitro Transcription and Translation. In vitro transcription and in vitro translation using rabbit reticulocyte lysate was performed as previously described (*6*). Posttranslational treatments of membrane-associated nascent proteins using endoglycosidase PNGase F and proteinase K were also performed as previously described (*16*). Translation in the presence of glycosylation acceptor NLT (Bz-Asn-Leu-Thr, final concentration of 0.7 mM) was performed as described by Lu et al. (*15*). Translation products were separated on a SDS−PAGE. The gel was then fixed with methanol and acetic acid for 30 min and treated with Amplify for 30 min. The gel was finally dried and exposed to an X-ray film at −80 °C. The images were digitized using a HP ScanJet 6100C and Adobe Photoshop 4.0.

RESULTS

Previously, it has been shown that the positive charges between TM3 and TM4 of Pgp is important for regulating the topogenesis of Pgp and responsible for generating more than one topology (*13*). Sequence comparison shows that CFTR does not have positive charges in the loop between TM3 and TM4 (Figure 1) which may be responsible for generating only one topology for CFTR. To test this hypothesis, two amino acids between TM3 and TM4 were mutated to mimic the charge distribution in Pgp (Figure 1). Effects of this mutation on the topogenesis of TM4 in CF-TM3,4R (*14*) were examined in a cell-free expression system. Charged amino acids at the C-terminal side (within 15 amino acid range) of TM4 were also mutated to test their effect on TM4 biogenesis (Figure 1).

The mutant molecules used were CF-TM3,4R(E217R/Q220K), CF-TM3,4R(R242E), and CF-TM3,4R(K246N/R248E). All mutant molecules have a glycosylation reporter linked to the C-terminal side of TM4 (14) (see also Figure

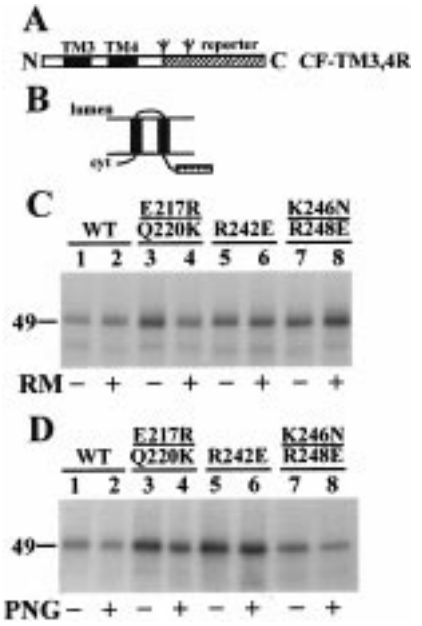

FIGURE 2: Translation and membrane orientation of wild-type and mutant CF-TM3,4R. (A and B) Schematic presentation of the linear structure (A) and topology (B) of wild-type and mutant CF-TM3,-4R. Solid and open bars represent TM segments and their linking loops, respectively. The hatched bar represents the glycosylation reporter sequence. The branched symbol indicates the potential glycosylation sites in the reporter. (C and D) In vitro translation and endoglycosidase treatment of wild-type and mutant CF-TM3,-4R. Wild-type and mutant CF-TM3,4R were translated in the absence (panel C, lanes 1, 3, 5, and 7) or presence (panel C, lanes 2, 4, 6, and 8) of RM. Proteins translated in the presence of RM were then treated with endoglycosidase PNGase F (PNG) (panel D, lanes 2, 4, 6, and 8). Lanes 1, 3, 5, and 7 in panel D are control samples treated in the absence of PNGase F.

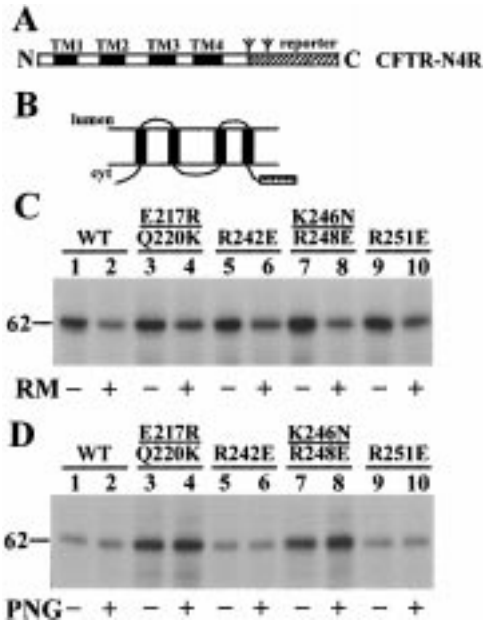

FIGURE 3: Translation and membrane orientation of wild-type and mutant CFTR-N4R. (A and B) Schematic presentation of the linear structure (A) and topology (B) of wild-type and mutant CFTR-N4R. The symbols are the same as described in Figure 2. (C and D) In vitro translation and endoglycosidase treatment of wild-type and mutant CFTR-N4R. Wild-type and mutant CFTR-N4R were translated in the absence (panel C, lanes 1, 3, 5, 7, and 9) or presence (panel C, lanes 2, 4, 6, 8, and 10) of RM. Proteins translated in the presence of RM were then treated with endoglycosidase PNGase F (PNG) (panel D, lanes 2, 4, 6, 8, and 10). Lanes 1, 3, 5, 7, and 9 in panel D are control samples treated in the absence of PNGase F.

2A). Glycosylation status of the reporter therefore serves as an indicator for the sidedness of the reporter and thus the membrane orientation of TM4. In vitro translation of wild-type and the mutant CF-TM3,4R transcripts was performed in RRL in the absence or presence of RM. As shown in Figure 2C, protein products of identical size were generated for all CF-TM3,4R constructs in the absence and presence of RM. This observation suggests that no glycosylation occurred to the nascent proteins. This is confirmed by endoglycosidase PNGase F treatment of the proteins generated in the presence of RM. The sizes of all protein products were not changed by PNGase F treatment (Figure 2D). Therefore, the reporters of all mutant CF-TM3,4R proteins were not glycosylated and, thus, located in the cytoplasm. This conclusion is confirmed by proteolysis analysis which did not reveal any protected reporter from the mutant CF-TM3,4R (data not shown). Similar results were observed with the mutant CFTR-N4R (*11*) proteins that contain all four TM segments (Figure 3). Thus, mutation of the positively charged amino acids surrounding TM4 of CFTR does not affect its membrane orientation.

The above studies suggest that the biogenesis properties of CFTR TM4 are different from that of Pgp. However, examination of the overall charge flanking TM4 indicates that the net charge difference across TM4 of Pgp is −2 whereas that of wild-type and mutant CFTR are all positive (Figure 1). To determine whether the net charge difference plays any role in CFTR TM4 topogenesis, we engineered a dramatic charge mutation into CF-TM3,4R and the new protein now has a net charge difference of −8 across TM4 of CFTR (Figure 4A). Translation of CF-TM3,4R(−8) generated an interesting observation. As shown in Figure 4C, in comparison with a single band of 49 kDa generated in the absence of RM (lane 1), a doublet of proteins around 49 kDa were generated in the presence of RM (lane 2). The lower band (indicated by an arrowhead) of the doublet was reduced to 44 kDa by endoglycosidase treatment (compare lanes 2 and 3). These results suggest that about 50% of nascent CF-TM3,4R(−8) proteins inserted into membranes in an inverted orientation with the reporter located in the RM lumen and glycosylated (Figure 4B, models II and III). It is possible that inversion of TM4 orientation revealed a site for cleavage by an unknown mechanism in the RM lumen (see Discussion). Thus, the reporter in the RM lumen was modified by glycosylation and was also cleaved off from TM4 (Figure 4B, models II and III). It should be noted that both models II and III are possible and cannot be differentiated based on the above results.

Proteinase K digestion of the membrane-associated nascent CF-TM3,4R(−8) indicates that the glycosylated proteins are also protected from protease digestion by the membranes (Figure 4D). This observation further confirms the inversion of TM4 orientation and cleavage. The remaining 50% molecules (upper band of the doublet) apparently have the model I orientation (Figure 4B) since they are not glycosylated and can be removed by proteinase K digestion. To further differentiate the molecule with model I topology and the one with cleaved and glycosylated reporter, we performed a translation of CF-TM3,4R(−8) in the presence or absence of a glycosylation acceptor peptide (NLT). As shown in

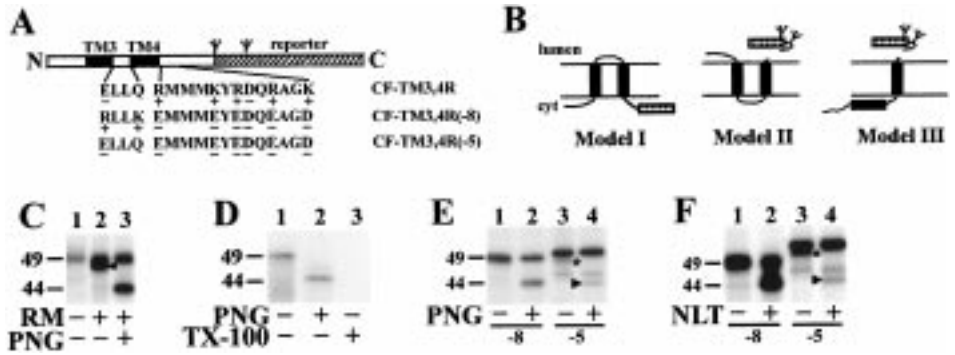

FIGURE 4: Translation and membrane orientation of CF-TM3,4R(−8) and CF-TM3,4R(−5). (A and B) Schematic presentation of the linear structure (A) and topology (B) of CF-TM3,4R(−8) and CF-TM3,4R(−5). The symbols are the same as described in Figure 2. The amino acid sequence flanking TM4 is shown with charged amino acids indicated. Note that the reporters in model II and III (panel B) are located in the RM lumen and were modified by addition of two-oligosaccharide chains and by signal peptidase cleavage. (C) In vitro translation and endoglycosidase treatment of CF-TM3,4R(−8). CF-TM3,4R(−8) was translated in the absence (lane 1) or presence (lane 2) of RM. Endoglycosidase PNGase F (PNG) treatment was performed with the proteins translated in the presence of RM (lane 3). (D) Proteinase K digestion of CF-TM3,4R(−8). Membrane-associated CF-TM3,4R(−8) was treated by proteinase K in the absence (lanes 1 and 2) or presence of Triton X-100 (TX-100) (lane 3). Samples in lane 2 were further treated with endoglycosidase PNGase F (PNG). (E) In vitro translation and endoglycosidase treatment of CF-TM3,4R(−5). CF-TM3,4R(−8) (lanes 1 and 2) and CF-TM3,4R(−5) (lanes 3 and 4) were translated in the presence of RM and treated without (lanes 1 and 3) or with endoglycosidase PNGase F (PNG) (lanes 2 and 4). The asterisk and arrowhead indicate glycosylated and deglycosylated proteins, respectively. (F) Inhibition of glycosylation by glycosylation acceptor peptide. In vitro translation of CF-TM3,4R(−8) (lanes 1 and 2) and CF-TM3,4R(−5) (lanes 3 and 4) was performed in the absence (lanes 1 and 3) or presence (lanes 2 and 4) of glycosylation acceptor peptide NLT and membrane associated proteins were separated by SDS−PAGE. The asterisk and arrowhead indicate glycosylated and deglycosylated proteins, respectively. Note that lanes 1−2 and 3−4 are from different gels.

Figure 4F, in the presence of NLT, the 44 kDa protein was translated in addition to the 49-kDa protein (lane 2) whereas only the 49 kDa protein was produced in the absence of NLT (lane 1). Thus, we conclude that the 44 kDa protein represents the cleaved reporter that, when modified by addition of two oligosaccharide chains, has the same mobility as the 49 kDa protein.

To determine whether a less net negative charge difference across TM4 will have similar effect on TM4 orientation, we made a new construct that has a net charge difference of −5 [CF-TM3,4R(−5), Figure 4A]. As shown in Figure 4E, endoglycosidase treatment of membrane-associated nascent CF-TM3,4R(−5) generated only a very small fraction of deglycosylated proteins in comparison with that of CF-TM3,-4R(−8) (compare lanes 1−2 with lanes 3−4). This is confirmed by using the glycosylation acceptor peptide NLT (compare lanes 3 with 4, Figure 4F). Thus, a net charge difference of −5 cannot significantly convert the orientation of TM4 of CFTR and only an extreme charge difference across TM4 can reverse the orientation of TM3 and TM4 to a maximum of ∼50% of the population. This is different from TM4 of Pgp of which a net charge difference of −2 forced TM4 to adopt the model III orientation (Figure 4B) (*13*).

We next examined whether the loop between TM3 and TM4 is important for the correct folding of TM4 in addition to the net charge difference. The loop linking TM3 and TM4 of Pgp was used to replace that in both wild-type and mutant CF-TM3,4R (Figure 2A) to generate CF3PL4R and CF3PL4R-(−8), respectively (see Figure 5A). As shown in Figure 5C, no glycosylation of the reporter occurred to the nascent CF3PL4R, suggesting that CF3PL4R has the orientation of model I (Figure 5B). However, a cleaved and glycosylated reporter was observed with the nascent CF3PL4R(−8) (Figure 5D). Note that there are two protein bands at the 49 kDa position (lane 2, Figure 5D), and the top one (indicated by an arrowhead) disappeared after endoglycosidase treatment. No other protein precursors were produced in the absence of RM (lane 1, Figure 5D). Thus, the top band produced in the presence of RM (lane 2, Figure 5D) represents the glycosylated reporter cleaved off from the same 49 kDa precursor. The glycosylated CF3PL4R(−8) peptide was also protected from proteinase K digestion (Figure 5E). These results suggest that the reporter of

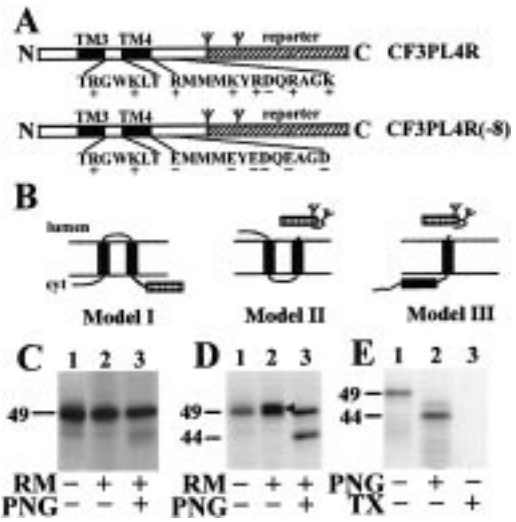

FIGURE 5: Translation and membrane orientation of CF3PL4R and CF3PL4R(−8). (A and B) Schematic presentation of the linear structure (A) and topology (B) of CF3PL4R and CF3PL4R(−8). The symbols are the same as described in Figure 2. (C and D) In vitro translation and endoglycosidase treatment of CF3PL4R and CF3PL4R(−8). CF3PL4R (C) and CF3PL4R(−8) (D) were translated in the absence (lane 1) or presence (lane 2) of RM. Endoglycosidase PNGase F (PNG) treatment was performed with the proteins translated in the presence of RM (lane 3). (E) Proteinase K digestion of CF3PL4R(−8). Membrane-associated CF3PL4R-(−8) was treated with proteinase K in the absence (lanes 1 and 2) or presence of Triton X-100 (TX-100) (lane 3). Samples in lane 2 were further treated with endoglycosidase PNGase F (PNG). The minor band in lane 2 represents the incompletely deglycosylated fragment that still contains one oligosaccharide chain.

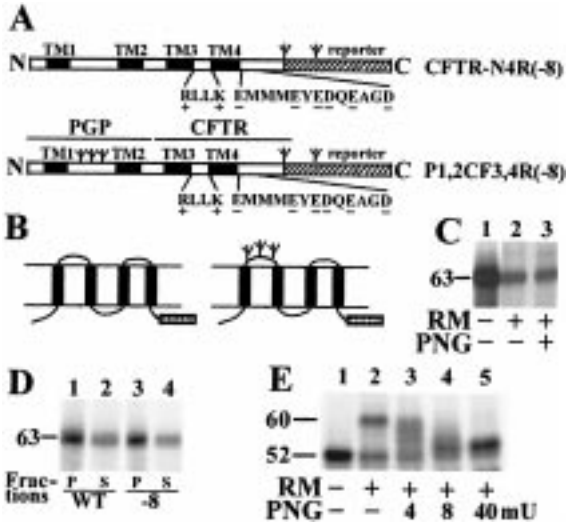

FIGURE 6: Translation and membrane orientation of CFTR−N4R-(−8) and P1,2CF3,4R(−8). **(**A and B) Schematic presentation of the linear structure (A) and topology (B) of CFTR-N4R(−8) and P1,2CF3,4R(−8). The symbols are the same as described in Figure 2. (C) In vitro translation and endoglycosidase treatment of CFTR-N4R(−8). CFTR-N4R(−8) was translated in the absence (lane 1) or presence (lane 2) of RM. Endoglycosidase PNGase F (PNG) treatment was performed with proteins translated in the presence of RM (lane 3). (D) Separation of membrane-associated wild type (WT) and mutant (−8) CFTR-N4R products. Translation was performed in the presence of RM and centrifuged to separate membrane pellet (P) from soluble fractions (S). (E) In vitro translation and endoglycosidase treatment of P1,2CF3,4R(−8). P1,-2CF3,4R(−8) was translated in the absence (lane 1) or presence (lane 2) of RM. Endoglycosidase PNGase F (PNG) treatment was performed with proteins translated in the presence of RM (lanes 3−5).

CF3PL4R(−8) was not only translocated into the RM lumen and glycosylated, but also cleaved off from the TM domain. On the basis of these observations, we conclude that the net charge difference across TM4, not the loop linking TM3 and TM4, is important for the folding of TM4.

To determine whether the amino terminal sequences including TM1 and TM2 influence the folding of TM3 and TM4, we engineered CFTR-N4R(−8) (Figure 6A). As shown in Figure 6C, translation of CFTR-N4R(−8) generated a protein of 63 kDa both in the absence and presence of RM (compare lanes 1 and 2). Endoglycosidase PNGase F treatment did not change the size of the protein (compare lanes 2 and 3, Figure 6C). These results suggest that the reporter in CFTR-N4R(−8) is not glycosylated and thus likely located in the cytoplasm (Figure 6B). This was confirmed by using proteinase K digestion that did not generate any membrane-protected reporter fragments (data not shown). To ensure that CFTR-N4R(−8) is membrane associated, a translation was centrifuged and membrane-associated proteins were separated from soluble fractions. As shown in Figure 6D, more than half of nascent CFTR-N4R(−8) was found in the membrane pellet (compare lanes 3 and 4). This is similar to that of wild-type CFTR-N4R (lanes 1 and 2, Figure 6D). Thus, mutation of amino acids surrounding the TM4 did not affect the membrane insertion properties of CFTR-N4R. On the basis of the above observations, we conclude that the ability of TM4 to translocate its C-terminal reporter into the RM lumen is inhibited by the amino terminal sequences including TM1 and TM2 (either by the correct membrane insertion of TM1 and TM2 or by simply the existence of these sequences; (see Discussion).

To determine whether the inhibition of TM4 to translocate its C-terminal reporter into the RM lumen by TM1 and TM2 is sequence specific, we swapped the amino terminal sequence including TM1 and TM2 of CFTR-N4R(−8) with that of Pgp (Figure 6A). The newly created molecule is named P1,2CF3,4R(−8). As shown in Figure 6E, translation of P1,2CF3,4R(−8) in the presence of RM generated a protein of higher molecular weight than that produced in the absence of RM (compare lanes 1 and 2). This suggests that proteins generated in the presence of RM are glycosylated. However, there are three potential glycosylation sites between TM1 and TM2 of Pgp and addition of ∼8 kDa oligosaccharide means that only the sites between TM1 and TM2 of Pgp were used (assuming that each high mannose oligosaccharide chain weighs ∼2.5 kDa). Partial deglycosylation analysis (lanes 3−5, Figure 6E) confirmed that only three oligosaccharide chains were added to the nascent P1,2CF3,-4R(−8). These results indicate that the reporter of P1,2CF3,-4R(−8) is located in the cytoplasm. Therefore, we conclude that the membrane folding of TM3 and TM4 of CFTR is controlled by their preceding membrane-anchorage amino acid sequences (TM1 and TM2) and this process is not sequence specific.

DISCUSSION

CFTR folding and assembly appear to involve several events occurring in the cytosol and ER. Misfolding of CFTR causes malfunction of CFTR and thus cystic fibrosis. Therefore, it is extremely important to understand the folding mechanism of CFTR. It has been suggested that the overall folding of CFTR depends on ER chaperons such as calnexin (*17*). Recently, detailed analysis of membrane insertion process suggests that the first two TM segments of CFTR have two distinct but independent mechanisms to ensure the correct membrane folding of the amino terminal end of CFTR (*15*). TM2 has a function to direct posttranslational membrane insertion to warrant correct topology of CFTR chains that fail to translocate via TM1. To understand how other TM segments were ensured to insert into membranes correctly, we investigated the topogenesis of TM3 and TM4 of CFTR. We found that the correct membrane insertion of TM3 and TM4 of CFTR is determined by their surrounding amino acid sequences and by the correct membrane insertion of TM1 and TM2. Thus, correct membrane insertion and folding of TM1 and TM2 play an essential role in correct membrane insertion and folding of other TM segments in CFTR.

Membrane insertion and folding of polytopic proteins are complicated processes and are not well understood. It was initially proposed that these proteins insert into membranes in a sequential process (*18*). That is, each TM segment of a polytopic membrane protein functions as a signal-anchor or stop-transfer sequence and these segments insert into membranes in a sequential event. However, exceptions to this rule have been found with several polytopic membrane proteins (*14*, *19*−*22*). For example, it was found that the membrane insertion and folding of Pgp TM3 and TM4 did not follow the prevailing sequential insertion model (*12*). The alternative insertion resulted in the generation of a

different folding for Pgp. Examination of individual segments revealed that the alternative membrane insertion and folding of Pgp were due to the charged amino acids flanking TM4 (*13*) and the high de novo $N_{in}-C_{out}$ membrane insertion activity of TM4 (*14*).

In contrast, distribution of charged amino acids flanking TM3 and TM4 of CFTR favors the generation of the predicted topology for CFTR and that the membrane orientation of TM3 and TM4 of CFTR is determined by the membrane insertion of the preceding TM1 and TM2. Previously, it has also been shown that the de novo $N_{in}-C_{out}$ membrane insertion activity of CFTR TM4 is not stronger than that of CFTR TM3 (*14*). Thus, we conclude that the membrane insertion of TM3 and TM4 of CFTR is different from that of Pgp. Although the membrane orientation of CFTR TM4 was altered by making the net charge difference of −8 across TM4, this alteration was overcome by the existence of TM1 and TM2 of CFTR. It is unknown, however, whether the effect of TM1 and TM2 on the membrane orientation of TM4 of CFTR is due to correct membrane insertion or simply by the existence of these sequences. Although we do not know whether TM1 and TM2 in CFTR-N4R(−8) inserted into membranes correctly, full glycosylation of the first loop in P1,2CF3,4R(−8) suggests that TM1 and TM2 were likely inserted into membranes as predicted. It is also possible that TM3 and TM4 in CFTR-N4R(−8) and P1,2CF3,4R(−8) were not inserted into membranes. However, this possibility is unlikely because both TM3 and TM4 have the ability to initiate efficiently the membrane insertion (*14*; see also Figure 4 and discussion below).

Previously, it has been thought that positive charges flanking a TM segment played an important role in determining the orientation of the TM segment according to the "positive inside" rule (*1*). Positively charged amino acids have been shown to be important for the membrane insertion and folding of both eukaryotic and prokaryotic proteins (*23*−*26*). Interestingly, the membrane orientation of TM3 and TM4 of CFTR was only converted partially with the changes of all charged amino acids flanking TM4 [CF-TM3,4R(−8)]. That is, only about 50% of CF-TM3,4R(−8) adopted an orientation to have the more positively charged domain linking TM3 and TM4 located cytoplasmically. There is almost no conversion of the orientation with CF-TM3,4R-(−5). This suggests that other unknown factors in addition to charged amino acids may also be important in determining the membrane orientation of CFTR TM segments.

It is interesting that when the orientation of TM4 in CF-TM3,4R(−8) and CF3PL4R(−8) was altered by mutation of charged amino acids surrounding TM4, a cleavage occurred to the nascent glycosylated protein. The mechanism for this cleavage is unknown, but possibly by enzymes located in the RM lumen. One possibility is that inversion of TM4 revealed a cryptic cleavage site for signal peptidase located in RM lumen. Examination of the sequence C-terminal to TM4 shows that Gly^{241} at the C-terminal end of TM4 is a potential cleavage site. Gly^{241} (−1 position) is a small residue and the Gly^{239} at −3 position is also small. This is consistent with the consensus signal-sequence cleavage site reported previously (*27*). However, we cannot rule out the possibility that the truncation may be due to proteolysis by enzymes in the RM lumen other than the signal peptidase. Further studies are clearly needed to clarify these issues. In addition, the truncation may also occur at the N-terminal end and not at the C-terminal end of TM4 of CF-TM3,4R(−8) with models II or III (Figure 4B), thus leaving the TM4 with reporter in the membrane. However, this possibility is unlikely because alkaline extraction studies showed that the cleaved reporter fragment is not integrated in the membrane (unpublished observation) and it is soluble in the RM lumen.

It has also been proposed that the membrane orientation of the first TM segment in a polytopic protein determines the folding of its subsequent TM sequences (*2*). While it seems that our observation in this study supports this hypothesis, "frustrated" topologies have been found in many other cases. Formation of these "frustrated" topologies was found due to the change in positive charges surrounding internal TM segments of polytopic proteins (*22*, *28*). In these "frustrated" topologies, internal TM segments adopt a "leave one out" strategy in order for other TM segments to insert into membranes properly (*28*). Interestingly, in this study, we found that the "frustrated" topology was not formed when TM1 and TM2 are present even though the mutation would have caused the "frustration" for TM3 and TM4. The effect of TM1 and TM2 on TM3 and TM4 is not sequence specific since replacing TM1 and TM2 with that of Pgp also overcome the mutation-induced change on CFTR TM4 orientation. This suggests that the membrane insertion of CFTR TM3 and TM4 is not independent from TM1 and TM2. Likely, TM3 and TM4 interact with TM1 and TM2 inside a translocon before or during their membrane insertion process.

In another study on the topogenesis of human MRP (multidrug resistance-associated protein), we found that the correct membrane insertion and orientation of TM1 were controlled by TM2.[2] These observations indicate that in a polytopic membrane protein the membrane insertion of each TM segment may not be necessarily independent of each other. Correct membrane insertion of one TM segment may rely on the correct folding of the other in a polytopic protein. The detailed mechanism for this type of regulation of membrane insertion for polytopic proteins remains to be determined.

Together with the previous observations, the results from this study strongly suggest that alternative folding mechanism for "frustrated" internal TM segments may be involved in the folding of at least some polytopic membrane proteins, such as Pgp and bacterial leader peptidase. In the case of CFTR, the correct membrane insertion and folding of TM1 and TM2 determine the correct membrane folding of the subsequent TM sequences. "Frustrating" the internal TM segments of CFTR by mutations will not alter their membrane insertion and folding assuming that TM1 and TM2 fold correctly. Thus, correct folding of TM1 and TM2 is essential for the biogenesis of CFTR. The results from this study imply that mutations of charged amino acids flanking internal TM sequences (e.g., TM3 and TM4) may not be responsible for the misfolding of CFTR and thus may not contribute to the cause of cystic fibrosis.

[2] Zhang, J. T. manuscript submitted for publication.

ACKNOWLEDGMENT

This work was initiated while the authors were still at the Department of Physiology and Biophysics, University of Texas Medical Branch, Galveston. We would like to thank Mr. Robert Broussard for technical assistance and our colleagues at both University of Texas Medical Branch and Indiana University School of Medicine for their helpful suggestions. Critical comments on the earlier version of this manuscript from Drs. Qun Chen, Ariel Castro, and Ms. Roxana Pincheira are greatly appreciated.

REFERENCES

1. von Heijne, G. (1986) *EMBO J.* 5, 3021−3027.
2. Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5786−5790.
3. Andrews, D. W., and Johnson, A. E. (1996) *Trends Biochem. Sci. 21*, 365−369.
4. Higgins, C. F. (1992) *Annu. Rev. Cell Biol. 8*, 67−113.
5. Doige, C. A., and Ames, G. F. (1993) *Annu. Rev. Microbiol. 47*, 291−319.
6. Zhang, J. T., and Ling, V. (1991) *J. Biol. Chem. 266*, 18224−18232.
7. Skach, W. R., Calayag, M. C., and Lingappa, V. R. (1993) *J. Biol. Chem. 268*, 6903−6908.
8. Bibi, E., and Beja, O. (1994) *J. Biol. Chem. 269*, 19910−19915.
9. Zhang, M., Wang, G., Shapiro, A., and Zhang, J. T. (1996) *Biochemistry 35*, 9728−9736.
10. Chang, X. B., Hou, Y. X., Jensen, T. J., and Riordan, J. R. (1994) *J. Biol. Chem. 269*, 18572−18575.
11. Chen, M., and Zhang, J. T. (1996) *Mol. Membr. Biol. 13*, 33−40.
12. Zhang, J. T., Duthie, M., and Ling, V. (1993) *J. Biol. Chem. 268*, 15101−15110.
13. Zhang, J. T., Lee, C. H., Duthie, M., and Ling, V. (1995) *J. Biol. Chem. 270*, 1742−1746.
14. Zhang, J. T., Chen, M., Han, E., and Wang, C. (1998) *Mol. Biol. Cell 9*, 853−863.
15. Lu, Y., Xiong, X., Helm, A., Kimani, K., Bragin, A., and Skach, W. R. (1998) *J. Biol. Chem. 273*, 568−576.
16. Zhang, J. T., and Ling, V. (1993) *Biochim. Biophys. Acta 1153*, 191−202.
17. Pind, S., Riordan, J. R., and Williams, D. B. (1994) *J. Biol. Chem. 269*, 12784−12788.
18. Blobel, G. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1496−1500.
19. Lipp, J., Flint, N., Haeuptle, M. T., and Dobberstein, B. (1989) *J. Cell Biol. 109*, 2013−2022.
20. Shi, L. B., Skach, W. R., Ma, T., and Verkman, A. S. (1995) *Biochemistry 34*, 8250−8256.
21. Zhang, J. T. (1996) *Mol. Biol. Cell 7*, 1709−1721.
22. Gafvelin, G., Sakaguchi, M., Andersson, H., and von Heijne, G. (1997) *J. Biol. Chem. 272*, 6119−6127.
23. Boyd, D., and Beckwith, J. (1990) *Cell 62*, 1031−1033.
24. Nilsson, I., and von Heijne, G. (1990) *Cell 62*, 1135−1141.
25. Beltzer, J. P., Fiedler, K., Fuhrer, C., Geffen, I., Handschin, C., Wessels, H. P., and Spiess, M. (1991) *J. Biol. Chem. 266*, 973−978.
26. Parks, G. D., and Lamb, R. A. (1991) *Cell 64*, 777−787.
27. Skach, W. R., and Lingappa, V. R. (1993) in *Mechanisms of intracellular trafficking and processing of proproteins* (Loh, Y. P., Ed.) CRC Press, Boca Raton, FL.
28. Gafvelin, G., and von Heijne, G. (1994) *Cell 77*, 401−412.

BI982153T