Co- and Posttranslational Translocation Mechanisms Direct Cystic Fibrosis Transmembrane Conductance Regulator N Terminus Transmembrane Assembly*

(Received for publication, July 8, 1997, and in revised form, October 1, 1997)

Yun Lu‡, Ximing Xiong‡, Andrew Helm, Kabuya Kimani, Alvina Bragin, and William R. Skach§

From the Department of Molecular and Cellular Engineering and Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Transmembrane topology of most eukaryotic polypopic proteins is established cotranslationally at the endoplasmic reticulum membrane through the action of alternating signal and stop transfer sequences. Here we demonstrate that the cystic fibrosis transmembrane conductance regulator (CFTR) achieves its N terminus topology through a variation of this mechanism that involves both co- and posttranslational translocation events. Using a series of defined chimeric and truncated proteins expressed in a reticulocyte lysate system, we have identified two topogenic determinants encoded within the first (TM1) and second (TM2) membrane-spanning segments of CFTR. Each sequence independently (i) directed endoplasmic reticulum targeting, (ii) translocated appropriate flanking residues, and (iii) achieved its proper membrane-spanning orientation. Signal sequence activity of TM1, however, was inefficient due to the presence of two charged residues, Glu92 and Lys95, located within its hydrophobic core. As a result, TM1 was able to direct correct topology for less than half of nascent CFTR chains. In contrast to TM1, TM2 signal sequence activity was both efficient and specific. Even in the absence of a functional TM1 signal sequence, TM2 was able to direct CFTR N terminus topology through a ribosome-dependent posttranslational mechanism. Mutating charged residues Glu92 and Lys95 to alanine improved TM1 signal sequence activity as well as the ability of TM1 to independently direct CFTR N terminus topology. Thus, a single functional signal sequence in either the first or second TM segment was sufficient for directing proper CFTR topology. These results identify two distinct and redundant translocation pathways for CFTR N terminus transmembrane assembly and support a model in which TM2 functions to ensure correct topology of CFTR chains that fail to translocate via TM1. This novel arrangement of topogenic information provides an alternative to conventional cotranslational pathways of polypeptide biogenesis.

Eukaryotic polypeptide integral membrane proteins acquire proper transmembrane orientation through the action of discrete topogenic determinants (i.e. signal, stop transfer, and signal anchor sequences) encoded within transmembrane segments and their flanking residues (1–3). It has been proposed that, like secretory (4) and bitopic transmembrane proteins (5, 6), polytopic topology is established cotranslationally as these topogenic determinants emerge from ER-bound ribosomes and direct sequential rounds of translocation initiation, termination, and membrane integration (1, 7–9). Consistent with this view, any given polytopic orientation may be generated de novo from combinations of signal, stop transfer, and/or signal anchor sequences engineered into a single polypeptide (10–12). Similarly, independent topogenic determinants have been identified from eukaryotic proteins such as bovine rhodopsin (13, 14), acetylcholine receptor (15, 16), P-glycoproteins (17–19), aquaporins (20, 21), band 3 anion exchanger (22), and calcium and P-type ATPases (23, 24). Whereas the molecular details of polytopic protein assembly into the ER membrane remain largely unknown, this process appears to require cytosolic and membrane bound components of the ER translocation machinery including signal recognition particle (25), the Sec61 complex, and TRAM (18, 26, 27).

If polytopic proteins followed a strict mechanism of cotranslational translocation, then failure of a given topogenic determinant to target or translocate the nascent chain at the ER membrane would result in irreversible protein misfolding. Structural requirements for directing protein topology would therefore be predicted to significantly restrict sequence diversity of topogenic determinants (e.g. transmembrane (TM) segments and/or their flanking regions). For this reason, certain polytopic proteins appear to have developed variations on cotranslational assembly pathways. For example, during transmembrane assembly of the Escherichia coli lac permease protein, interactions between TM9 and TM10 were required to correctly position residue Arg302 within the bilayer and to ensure transmembrane topology of the TM9–10 peptide loop (28, 29). Similarly, membrane integration of human P-glycoprotein (MDR1) into the lipid bilayer required cooperativity between independent signal sequences encoded within TM1 and TM2 (18). More recently, it was shown that interactions between topogenic determinants within the yeast protein Sec61p were involved in directing topology of weakly hydrophobic internal transmembrane segments (30). Finally, it has been shown in vitro that some but not all polytopic proteins are capable of assembling posttranslationally into the ER membrane after polypeptide synthesis has been completed (8, 31, 32). Despite these studies however, the mechanisms by which

---

*This work was supported by National Institutes of Health Grants CA01614, DK51818, and GM53457 and the Cystic Fibrosis Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ These authors contributed equally to this work.

§ To whom correspondence should be addressed: Dept. of Molecular and Cellular Engineering, University of Pennsylvania, Philadelphia, PA 19104. Tel.: 215-898-0953; Fax: 215-573-8590.

1 The abbreviations used are: ER, endoplasmic reticulum; CFTR, cystic fibrosis transmembrane conductance regulator; PCR, polymerase chain reaction; WT, wild type; TM, transmembrane segment; RRL, rabbit reticulocyte lysate; PK, proteinase K.
multiple topogenic determinants cooperate to direct complex topology in a non-cotranslational manner remain poorly understood.

In the current study we examine the biogenesis of the cystic fibrosis transmembrane conductance regulator, CFTR, and show that both cotranslational and posttranslational mechanisms are involved in establishing correct topology of the first two N terminus transmembrane segments. This mechanism of assembly resulted from the presence of charged residues (Glu<sup>92</sup> and Lys<sup>95</sup>) located within the hydrophobic core of TM1, which markedly decreased TM1 signal sequence activity. Thus TM1 was capable of directing translocation for only a subset of CFTR chains, whereas signal sequence activity of TM2 was required to direct topology of remaining chains. CFTR N terminus biogenesis therefore does not rely on the sequential action of independent topogenic sequences to establish its multispanning topology. Rather CFTR utilizes two distinct and redundant translocation pathways: a cotranslational pathway directed by TM2 and a ribosome-dependent posttranslational pathway directed by TM1. An important consequence of this mechanism is that redundant topogenic information encoded by TM2 enables TM1 to acquire critical structural features that would otherwise interfere with conventional cotranslational assembly.

**MATERIALS AND METHODS**

dNA Construction—E92A and K95A mutations were engineered into CFTR by site-directed mutagenesis using a single stranded (M-13) (plasmid template) and oligonucleotides TATATTTAGGCGCCGTCAC-

**RESULTS**

**CFTR TM1 Functions as an Inefficient Signal Sequence to Direct CFTR Translocation**—To define the translocation activity of CFTR-TM1, CFTR cDNA encoding residues Met<sup>1</sup> through Thr<sup>100</sup> was added to RRL translation mixtures diluted in 0.5 ml of buffer A (1:1000 dilution) or to clarified oocyte homogenates. After a 10–30 min preincubation, 5 µl of antiprolactin antisera (ICN Immunologicals, Costa Mesa, CA) was added to RRL translation reaction mixture. Resultant polypeptides were resolved on SDS-PAGE, transferred to nitrocellulose, and visualized by the P reporter.

**Immunoprecipitation and Autoradiography—**Antiprolactin antisera (ICN Immunologicals, Costa Mesa, CA) was added to RRL translation mixtures diluted in 0.5 ml of buffer A (1:1000 dilution) or to clarified oocyte homogenates. After a 10–30 min preincubation, 5 µl of protein A Affi-Gel (Bio-Rad) was added, and the samples were mixed at 4 °C for 2–10 h prior to washing three times with buffer A and twice with 0.1 M NaCl, 0.1 M Tris, pH 8.0. Samples were analyzed by SDS-PAGE, ENHANCE (NEN Life Science Products) fluorography, and autoradiography. Autoradiograms were digitized with an AGFA Studio Scan II, and band intensities were quantitated on unmodified images using Adobe Photoshop software as described previously (17, 20). Several different film exposure times were compared to ensure that determinations remained within the linear range.

**RESULTS**

**CFTR N Terminus Transmembrane Assembly**

Made in the identical manner except that pSPCFTR(G85E) was used as the template for the initial 5′ PCR reactions. gtM2.P plasmids were generated by PCR amplification of WT and mutant CFTR using sense primer TACTGCAGCGGTGACGGCGCCTAA, digestion of the PCR fragment with BstI and ligated into HindIII/BstEI-digested S.LST<sub>gt</sub>P. This intermediate plasmid, pSPf/TM2.P, was then digested with NcoI and ligated to a synthetic oligonucleotide linker containing NcoI compatible ends (sense strand CATTAGACCGAGATCATC). The resultant plasmid encodes N terminus residues MNGSSVMV followed by CFTR residues Ala<sup>173</sup>–Asn<sup>186</sup> followed by the P reporter. mutant constructs were sequenced throughout PCR-amplified or -mutagenized cDNA to verify the presence of appropriately engineered mutations and the absence of inadvertent PCR errors.

**Transcription and Translation—**mRNA was transcribed in vitro with SP6 RNA polymerase (New England Biolabs) using 2–4 µg of plasmid DNA in a 10–µl volume at 40 °C for 1 h as described (11). Transcription reaction mixture was translated directly in a transcription-linked rabbit reticulocyte lysate (RRL) system as described previously (18). Where indicated, canine pancreas microsomes (final concentration, 8.0 µg/ml) and/or the tripeptide Ac-Asn-Tyr-Thr (final concentration, 0.2 mM) were added at the start of translation. Prior to proteolysis, chains synthesized on truncated mRNA (e.g. lacking a termination codon) were released from ribosomes by incubation in 1 m Tris, pH 7.5, at 24 °C for 10 min. For oocyte expression 50 µCi of [35S]Met (ICN Pharmaceuticals, Irvine, CA) (0.5 µl of a 10 × concentrated mixture) was added to 2 µl of transcription mixture and injected into mature Xenopus laevis oocytes (XO) (50 nl/oocytes). Oocytes were incubated at 18 °C for 3–4 h and homogenized on ice in 10 volumes of 0.25 M sucrose, 50 mM KAc, 5 mM MgAc<sub>2</sub>, 1.0 mM dithiothreitol, 50 mM Tris (pH 7.5). Prior to proteolysis, CaCl<sub>2</sub> was added to 10 mM final concentration.

**Proteolysis Digestion—**Protease K (PK) was added to RRL translation mixture or aliquots of oocyte homogenate (0.2 mg/ml final concentration) in the presence or absence of 1% Triton X-100 and incubated on ice for 1 h (18). Residual protease was inactivated by adding phenylmethylsulfonyl fluoride (10 mM) and rapidly mixing with 10 volumes of 1% SDS, 0.1 M Tris, pH 8.0 (preheated to 100 °C). Oocyte samples were subsequently diluted in 7× volumes of buffer A and combined with Triton X-100, 2× EDTA, 0.1 mM phenylmethylsulfonyl fluoride, and 0.1 M Tris, pH 8.0), incubated at 4 °C for 2–4 h, centrifuged at 14,000 × g for 15 min, and the supernatants were used for subsequent immunoprecipitation.
Fragments were also observed prior to protease digestion (19-kDa band, lanes 1 and 2, upward arrows; more evident in lanes 14 and 17). These latter fragments were dependent on the addition of ER membranes (data not shown) and were likely generated by signal peptidase cleavage of nascent chains at cryptic recognition sites unmasked by truncation of TM2 as has been observed previously for similar constructs (17, 18, 20, 36). In contrast to WT chains, no protease-protected fragments were observed for TM1.P(G85E) or TM1.P(G91R) chains, demonstrating that these inherited cystic fibrosis-related mutations abolished the ability of TM1 to direct translocation of the P reporter (lanes 4–6 and 7–9, respectively).

Quantitation of autoradiograms in Fig. 1 and correction for methionine content of fragments revealed that CFTR-TM1 was significantly less efficient at directing P reporter translocation than other polytopic protein-derived signal sequences that have been examined in this same context. The P reporter resided in the ER lumen in only 12% of newly synthesized WT TM1.P chains compared with 80–95% translocation efficiency previously reported by us and others (15–18, 21). Because all TM1-flanking residues were included in these constructs, translocation efficiency of TM1 could not be explained by removal of topogenic information. Similarly, the P reporter has been shown to faithfully follow topogenic information in a wide variety of contexts and would not be expected to negatively influence TM1 activity (11). Rather, it seemed plausible that TM1 either lacked translocation specificity (i.e. some chains spanned the membrane in the opposite orientation with their N terminus translocated) or that primary structural features within TM1, such as charged residues Glu92 and/or Lys95, limited its signal sequence activity. We therefore examined translocation efficiency of polypeptides generated from plasmids TM1.P(E92A), TM1.P(K95A), and TM1.P(E92A/K95A). As shown in Fig. 1A, lanes 10–18, the E92A and E92A/K95A mutations both improved TM1 signal sequence activity (43% and 79% of P translocated, respectively), whereas the K95A mutation by itself had little effect (10% of P translocated). This suggested that residues Glu92 and Lys95 within the hydrophobic core of TM1 directly reduced TM1 C terminus translocation activity and prevented TM1 from achieving a stable transmembrane orientation.

To address whether the weak TM1 signal sequence activity observed in vitro might be due to artifacts introduced by the reconstituted cell free RRL system, we also examined TM1 signal sequence activity in microinjected X. laevis oocytes (Fig. 1B). These in vitro studies confirmed results obtained in vitro and demonstrated the following: (i) signal sequence activity of WT TM1 was significantly impaired (only 28% of chains trans-
located); (ii) G85E and G91R mutations essentially abolished TM1 signal sequence activity (<5% of chains translocated); and (iii) E92A and E92A/K95A mutations improved TM1 signal sequence activity (36% and 70% of chains translocated, respectively).

**CFTR-TM2 Encodes Signal Sequence Activity Complementary to TM1**—If CFTR utilized a strict cotranslational mechanism of assembly, then the weak signal sequence activity of WT TM1 would result in failure of most CFTR chains to translocate and assemble properly in the ER membrane. This, however, was not the case because greater than 70% of WT as well as G85E and G91R chains achieved their correct N terminus topology in the ER membrane (33) and Fig. 4). These results suggested that topogenic information in addition to TM1 must function to ensure correct CFTR topology. We therefore examined the signal sequence activity of TM2 using the plasmid ggTM2.P, which encodes two (engineered) N-linked glycosylation consensus sites 12 and 18 residues from the membrane-spanning segment, followed by TM2 together with its flanking sequences (CFTR residues Ala107–Asn186), followed by the P reporter (diagrammed in Fig. 2). For initial experiments, plasmid ggTM2.P was truncated at codon Asn186 (Fig. 2). Translation of ggTM2.P (lanes 1–3) as shown in panel A. (C) PK digestion of chains generated from plasmid ggTM2.P. Total translation products (lanes 1–3) and translation products immunoprecipitated with antiprolactin antisera (lanes 4–6). Topology of TM2 is indicated beneath the autoradiograms.

For the experiments outlined above, microsomal membranes for ggTM2.P chains containing the P reporter (65% of chains glycosylated) (Fig. 2B). Both glycosylated and unglycosylated full-length ggTM2.P chains were accessible to protease (Fig. 2C, lanes 1–3), and immunoprecipitation with antiprolactin antisera confirmed that the P reporter was cytosolic (lanes 4–6). Thus TM2 directed an N-trans (or type I) transmembrane topology in which N terminus flanking residues resided in the ER lumen and C terminus residues faced the cytosol. This topology is consistent with the predicted orientation for TM2 in native CFTR (37, 38).

For the experiments outlined above, microsomal membranes were titrated to achieve 90–95% translocation efficiency based on protease protection of control secretory proteins (data not shown). However, estimating translocation efficiency by N-linked glycosylation is more difficult because utilization of consensus sites is context dependent. Lack of ggTM2.P glycosylation might therefore have resulted either from failure of the chain to translocate into the microsome lumen or failure of a particular translocated consensus site to be utilized by oligosaccharyltransferase. For readily accessible sites, glycosylation efficiency of our microsomal membranes is approximately 85–90% of translocated chains (18). Therefore, the 65% glycosylation efficiency observed for ggTM2.P chains indicates that TM2 is efficient in translocating N terminus flanking residues and directing a specific N-trans or type I transmembrane topology complimentary to that of TM1. It should be noted that the TM1.P constructs tested in Fig. 1 also encoded a potential N-linked glycosylation consensus site 33 residues N-terminal to the putative membrane-spanning segment. Whereas lack of glycosylation of TM1.P chains does not rule out the possibility that some chains were oriented with their N terminus in the ER lumen (for reasons stated above), the different glycosylation and proteolysis results observed for polypeptides TM1.P and ggTM2.P strongly indicate that TM1 and TM2 encode distinct and complimentary translocation specificities.

To better define the role of TM2 in directing CFTR topology, we attempted to decrease TM2 signal sequence activity by introducing three mutations previously identified in cystic fibrosis patients (G126D, ΔE115, and E116K) (39). These mutations were engineered alone or in combination into the plasmid ggTM2.P, and topology of the resulting chains was determined in RRL (Fig. 3). Mutations ΔE115 or E116K had only minor effects on TM2 signal sequence activity based on the fraction of glycosylated chains (Fig. 3A, lanes 1–6). However, the double mutations E115K/E116K, E116K/G126D and the triple mutation E115K/E116K/G126D all reduced N-linked glycosylation to approximately 20% of chains, a 65–70% reduction from WT levels (Fig. 3A, lanes 7–15). Because the location and context of glycosylation sites were unchanged in these mutants, it was unlikely that accessibility of consensus sites to oligosaccharyltransferase was altered. Rather, these results suggest a corresponding decrease in N terminus translocation efficiency by mutant TM2.

PK digestion of mutant ggTM2.P chains (ΔE115, E116K, and E115K/E116K) indicated that the introduction of basic residues flanking the N terminus of TM2, altered TM2 translocation specificity and enabled TM2 to translocate C terminus flanking sequences in a subset of chains. This was particularly evident for the E115K/E116K mutant (Fig. 3B, lanes 1–9, upward ar-
reported was observed (Fig. 4, downward arrows) which has been shown to contain TM1, TM2, and the first extracellular loop (33). The size of these protected fragments suggests that they contain 40–60% of the total methionine residues present in undigested chains (Met1, Met155, and possibly Met156 are digested, whereas Met82 and Met150 are likely protected). Quantitation of eleven different experiments for truncated WT chains revealed that 42% (range, 31–55%) of total 35S was protected from PK. Thus, CFTR N terminus assembly was greater than 70% efficient in RRL (depending on the precise site of PK cleavage).

Fig. 4B shows the translocation efficiency of chains that contain a single mutation in either TM1 or TM2 relative to WT chains. For comparison purposes, the translocation efficiency of WT chains was tested in each experiment and normalized to 1. TM2 was independently capable of translocating the CFTR N terminus in the absence of a functional TM1 signal sequence (Fig. 4, lanes 4–9). However, TM2 mutations E116K/G126D and E115K/E116K/G126D had a relatively minor but reproducible effect on CFTR N terminus topology, 82 and 79% of WT translocation activity, respectively (Fig. 4B, lanes 16–24 and Fig. 5, A and B). When the TM1 mutation G85E was introduced into chains containing E116K/G126D or E115K/E116K/G126D mutations, translocation efficiency was further reduced to 45% and 48% of WT levels, respectively (Fig. 5, A and B). This level of translocation correlated well with the residual TM2 signal sequence activity in these chains and indicated that WT TM1 was able to partially compensate for the loss of TM2 signal sequence activity. TM1, however, was unable to ensure that all CFTR chains were correctly oriented in the ER membrane in the absence of a fully functional TM2 signal sequence. This conclusion was further supported by the observation that improving TM1 signal sequence activity using the mutant E92A/K95A completely restored N terminus translocation in TM2 mutants (Fig. 5). Thus an efficient signal sequence in either the TM1 or the TM2 position was sufficient to ensure proper CFTR N terminus topology.

We also observed that the E115K/E116K mutation, which partially reversed TM2 translocation specificity, was more disruptive than other TM2 mutations. Only 55% of truncated E115K/E116K chains achieved correct topology (Fig. 5C), and the G85E mutation had little effect on these chains. Furthermore, even an efficient TM1 signal sequence (E92A/K95A) restored translocation efficiency to only 69% of WT levels. Thus when topogenic information encoded within TM2 directly conflicted with the information encoded within TM1, these determinants appeared to exhibit an antagonistic effect on transmembrane assembly.

CFTR N Terminus—Our results predict that as TM2 emerges from the ribosome it posttranslationally directs topology of TM1 and the TM1–2 peptide loop. However it was unclear whether TM2-mediated translocation might require earlier (e.g., cotranslational) membrane targeting by upstream residues prior to the initiation of translocation. We therefore tested whether TM2 was able to direct translocation of presynthesized cytosolic nascent chains by truncating CFTR cDNA at codon Asn186 and translating in RRL in the absence of microsomal membranes. Translation was then terminated by addition of cyclohexamide, and the RRL mixture was then incubated with microsomal membranes prior to PK digestion. As shown in Fig. 6A, CFTR chains containing WT TM1 and TM2 posttranslationally achieved correct transmembrane topology as evidenced by the presence of 17-kDa protease-protected fragments (upward arrows). Furthermore, the efficiency of posttranslational translo-
**DISCUSSION**

CFTR assembly and maturation requires coordinated folding events in the cytosol, ER lumen, and lipid bilayer. In aqueous environments, CFTR folding appears to be facilitated by cellular chaperones including calnexin (46) and hsp70 (47), which repetitively bind to incompletely or improperly folded chains. This ATP-dependent process is believed to (i) slow protein folding, (ii) correct off-pathway folding intermediates, and/or (iii) maintain a transient “folding competent” state until protein synthesis and assembly have been completed (reviewed in Refs. 48 and 49). Transmembrane domains of polytopic proteins must also fold properly, and one key aspect of this process is the acquisition of correct topology and association with membrane lipids, here referred to as topogenesis. As with aqueous polypeptide domains, topogenesis is facilitated by a complex set of cellular proteins. These proteins, collectively termed the transloco (50, 51), interact with topogenic determinants to direct protein translocation across the ER membrane and integration into the lipid bilayer (18, 26, 27, 52–54).

Because translation and translocation are usually coupled vectorially at the ER membrane, polytopic protein topogenesis has been proposed to occur through the sequential action of alternating signal (anchor) and stop transfer sequences (1). Consistent with this hypothesis, cotranslational topogenesis has been demonstrated for several eukaryotic polytopic proteins such as the erythrocyte anion transporter and the mercury-insensitive water channel where topology of N terminus TM segments is established even before the synthesis of C terminus TM segments has been completed (7, 21). However, cotranslational assembly requires each individual topogenic determinant to act independently, sequentially, and efficiently, because a mistake at any point in topogenesis will potentially misdirect a region of polypeptide into the wrong cellular compartment. In the current study we provide evidence both in vitro and in vivo that the initial topogenic determinant of CFTR, TM1, is quite inefficient at directing nascent chain translocation due to the presence of specific charged residues within its hydrophobic core. Thus, if CFTR solely utilized a cotranslational mechanism of topogenesis, chains failing to target and/or translocate via TM1 would fail to assemble into the membrane. This was not the case, however, because WT CFTR chains achieved their correct topology. Furthermore, even chains containing a completely defective TM1 signal sequence (G85E and G91R mutants) were properly oriented in the membrane but only if a functional TM2 signal sequence was present.

Our data indicate therefore that CFTR utilizes two independent and redundant topogenesis pathways at the ER membrane, each of which contributes to directing correct N terminus topology (Fig. 7). In one pathway, TM1 emerges from the ribosome, targets CFTR chains to the ER membrane, initiates translocation of C terminus flanking sequences, and spans the membrane in a C-trans (type II) orientation. In this subset of CFTR chains, translocation is subsequently terminated by TM2, and topology is thus established in a cotranslational manner. However, when TM1 fails to translocate the chain, topogenic information encoded within TM2 provides a second chance for chains to acquire their correct transmembrane orientation. In this case, we propose that TM2 initiates translocation of its own N terminus flanking residues in a manner similar to an N-trans (type I) signal anchor sequence (2, 3, 14, 55). The unusual feature of this process is that TM1 is posttranslationally directed into the transloco channel from a cytosolic orientation where it then terminates translocation and establishes its proper membrane-spanning topology. The consequence of such a mechanism is that the direction of nas-
cent chain movement through the translocation channel is
different in these two pathways. In cotranslational transloca-
tion (e.g. conventional pathway), the chain moves into the ER
lumen from N to C terminus as it leaves the ribosome, whereas
in the posttranslational pathway the chain enters the ER lu-
men from C to N terminus, presumably from a location in the
cytosol.

Because protein topogenesis is a multistep process involving
signal recognition particle binding (56), ER targeting (57), sig-
nal recognition particle release (58), translocon gating, trans-
location initiation (59), and membrane integration (52), several
details of CFTR topogenesis remain unknown. It is possible
that residues Glu92 and/or Lys95 within TM1 interfere with ER
membrane targeting. If this were the case, then TM1 would fail
to bind signal recognition particle and emerge from the ribo-
some directly into the cytosol. TM2 would therefore be required
to perform all targeting and translocation events. Alterna-
tively, if TM1 were able to dock ribosomes at the ER membrane
but failed to initiate translocation (i.e. failed to effectively gate
the translocon channel), then TM2 would be required only to
initiate translocation of chains already in the physical vicinity
of the translocon. A third possibility is that TM1 might initiate
translocation of CFTR chains but independently be unable to
close the translocon channel, allowing the P reporter to fall
back into the cytosol where it would appear to have never translocated (60). In this case, TM2 might act to stabilize an
otherwise transient transmembrane orientation of TM1. This
is an important consideration because CFTR-TM1, like MDR1-
TM1, is unable to independently integrate the nascent chain into
the lipid bilayer (18), implying that neither of these se-
quences completely disengages from the translocon and/or its
associated proteins (52). However, because MDR1-TM1 effi-
ciently directed nascent chain translocation despite its failure
to integrate, lack of integration per se does not necessarily
result in retrograde translocation. Furthermore, unlike MDR1
where TM1 and TM2 together are sufficient for integration
(18), stable membrane integration of CFTR requires synthesis
of at least four TM segments. Thus if TM2 acted primarily to
prevent TM1 from falling out of the translocon, it would likely
do so in the context of ER translocation machinery. Given the
limitations of the current study, we cannot rule out the possi-
bility that such a mechanism might contribute in some manner
to the cooperativity between TM1 and TM2 observed in CFTR
assembly.

It should also be noted that despite the widespread and
accepted use of RRL in defining mechanisms of translocation
and membrane integration, it is always possible that
in vitro systems may introduce potential artifacts that do not reflect
in vivo assembly pathways. This may be particularly important
for proteins such as CFTR with specialized requirements for
assembly. In this regard, our analyses of TM1 signal sequence
activity using an in vivo Xenopus oocyte expression system
confirmed all of the essential features observed in vitro, namely
the inefficient activity of TM1 and the effects charged residues
Glu85, Arg91, Glu92, and Lys95. These results, together with
recent evidence that up to 100% of newly synthesized CFTR
protein matures in Xenopus oocytes (33), argues that both in
vivo and in vitro the weak TM1 signal sequence activity must

2 William R. Skach, unpublished observations.
be compensated by topogenic information encoded within TM2 to ensure efficient and proper N terminus assembly.

Why might CFTR have evolved this unusual mechanism of topogenesis? One possibility is that such a mechanism arose from conflicting structural requirements within TM1 for protein assembly (at the ER membrane) and for protein function (elsewhere in the cell). In this regard, charged residues within TM segments are known to influence protein topogenesis, folding, and oligomerization (40, 42, 43, 61, 62), as well as structural properties important for ion conduction, selectivity, and gating (63–66). Consistent with this view, we observed that full-length CFTR encoding E92A or the double mutation, E92A/K95A, exhibited markedly reduced chloride channel activity when expressed in Xenopus oocytes. Adding scanning cysteine accessibility studies have revealed that Lys95 resides on a hydrophilic surface of TM1 and likely faces the CFTR chloride channel pore, whereas Glu92 appears to face 40° away from the pore surface, suggesting that it contributes to ionic interactions within the plane of the bilayer (67). Thus it seems likely that functional requirements for charged residues Glu92 and/or Lys95 within mature CFTR necessitated the presence of redundant topogenic information encoded by TM2 to ensure efficient protein topogenesis. It is interesting, however, that unlike other ATP binding cassette transporters, up to 80% of WT CFTR protein expressed in mammalian cells may be degraded prior to exit from the ER (68). This is proposed to result from inefficient folding of WT protein and subsequent degradation of CFTR through a ubiquitin proteasome-mediated pathway (69, 70). Why such a complex mechanism for ensuring fidelity of TM1–2 topogenesis should have evolved in light of the overall inefficient folding of full-length CFTR, and whether these two aspects of CFTR biogenesis are related, remain unanswered questions.

Finally, cooperativity between topogenic determinants has also been described for additional polytopic proteins such as the E. coli protein lac permease (29), the yeast protein Sec61p (30), and others. In the case of lac permease a charged residue critical for lactose transport, Arg302, markedly decreased TM9 signal sequence activity and led the authors to propose that a salt bridge between Arg302 and Glu325 (within TM10) allowed these helices to insert into the membrane as a pair. Whereas this observation resembles the behavior of TM1 and TM2 from CFTR, our data support a slightly different model in which CFTR topogenesis is directed by two distinct signal sequences, respectively.

ACKNOWLEDGMENTS—The authors thank Kenneth Moss for technical assistance, William Gresh, Drs. Michael Marks, and Carol Deutsch for helpful comments.

REFERENCES
1. Blobel, G. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 1496–1500
2. Wickner, W. T., and Lodish, H. F. (1985) Science 226, 400–407
3. Skach, W., and Lingappa, V. (1983) in Mechanisms of Intracellular Trafficking and Processing of Pro-proteins (Loh, Y. P., ed), pp. 19–77, CRC Press Inc., Boca Raton, FL
4. Lingappa, V. R., Devillers-Thiry, A., and Blobel, G. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 2432–2436
5. Rothman, J. E., and Lodish, H. F. (1977) Nature 270, 775–780
6. Lingappa, V. R., Lingappa, J. R., Prasad, R., Ehner, K. E., and Blobel, G. (1978) Proc. Natl. Acad. Sci. U. S. A. 75, 2338–2342
7. Brasel, W. A., and Lodish, H. F. (1982) Cell 25, 23–31
8. Brown, D., and Simon, R. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 1674–1678
9. Borel, A., and Simon, S. (1996) Biochemistry 35, 10587–10594
10. Lipp, J., Flint, N., Haeuptle, M. T., and Dobberstein, B. (1989) J. Cell Biol. 109, 2013–2022
11. Rothman, R. E., Andrews, D. W., Calayag, M. C., and Lingappa, V. R. (1988) J. Biol. Chem. 263, 16470–16480

3 K. Foskett and W. Skach, unpublished observations.
