Role of transmembrane spanning domain 1 in cystic fibrosis transmembrane conductance regulator folding

Anna E. Patrick^{1*,#a}, Linda Millen¹, Philip J. Thomas¹

¹ Department of Physiology, University of Texas Southwestern Medical Center, Dallas, Texas, United States of America

^{#a} Current Address: Department of Pediatrics, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America

* Anna E. Patrick

E-mail: anna.e.patrick@vumc.org (AEP)

Keywords: CFTR; cystic fibrosis; protein folding; multidomain membrane protein; domain interactions

Abstract

Cystic fibrosis (CF) is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) protein that disrupt its folding pathway. The most common mutation causing CF is a deletion of phenylalanine at position 508 ($\Delta F508$). CFTR contains five domains that each form cotranslational structures that interact with other domains as they are produced and folded. CFTR is comprised of two transmembrane spanning domains (TMDs), two nucleotide binding domains (NBDs) and a unique regulatory region (R). The first domain translated, TMD1, forms interdomain interactions with the other domains in CFTR. In TMD1, long intracellular loops extend into the cytoplasm and interact with both NBDs via coupling helices and with TMD2 via transmembrane spans (TMs). We examined mutations in TMD1 to determine the impact on individual domain and multidomain constructs. We found that mutations in a TM span or in the cytosolic ICLs interfere with specific steps in the hierarchical folding of CFTR. TM1 CF-causing mutants, G85E and G91R, directly affect TMD1, whereas most ICL1 and ICL2 mutant effects become apparent in the presence of TMD2. A single mutant in ICL2 worsened CFTR trafficking in the presence of NBD2, supporting its role in the ICL2-NBD2 interface. Mutation of hydrophobic residues in ICL coupling helices tended to increased levels of pre-TMD2 biogenic intermediates but caused ER accumulation in the presence of TMD2. This suggests a tradeoff between transient stability during translation and final structure. NBD2 increased the efficiency of mutant trafficking from the ER, consistent with stabilization of the full-length constructs. While the G85E and G91R mutants in TM1 have immediately detectable effects, most of the studied mutant effects and the $\Delta F508$ mutant are apparent after production of TMD2, supporting this intermediate as a major point of recognition by protein quality control.

Introduction

Cystic fibrosis (CF) is a common, lethal monogenetic disease caused by mutations in the cystic fibrosis transmembrane conductance regulator protein (CFTR) [1-3]. CFTR acts as a chloride channel at the surface of secretory epithelia, and its loss of function results in a reduced liquid layer and thick mucus secretions [4]. CF-causing mutations are identified throughout the protein, however 90% of CF patients have at least one allele with a deletion of phenylalanine at position 508 ($\Delta F508$) [2]. The $\Delta F508$ and other mutations result in misfolding of the CFTR protein, ER accumulation, and eventually degradation. Thus, the majority of CF is caused by mutations that disrupt CFTR folding.

CFTR is a member of the ABC transporter superfamily that uses ATP binding and hydrolysis to move substrates across membranes [5]. CFTR contains two transmembrane spanning domains (TMDs) and two nucleotide binding domains (NBDs) like other ABC transporters. CFTR also has a unique and less structured regulatory region (R) [6]. ABC transporter structures have varying resolutions and include Sav1866 [7, 8], MsbA [9], P-glycoprotein [10], MRP1 [11], zebrafish CFTR [12], and human CFTR [13]. Several homology models of CFTR have been constructed [14-17]. In CFTR, two TMDs wrap around each other in a domain-swapped fashion with each having two long intracellular loops (ICLs) that extend into the cytoplasm and interact with both NBDs. The NBDs interact in a head-to-tail fashion sandwiching the ATP binding sites [18-20]. Conformational signals from NBD ATP binding and hydrolysis are transmitted by the ICLs to the TMDs, resulting in chloride channel opening and closing [21, 22]. A coupling helix at the end of each ICL interacts with the NBDs. The coupling helices of ICL1 and ICL3 interact with both NBDs, with primary interactions occurring between ICL1 and NBD1 or ICL3 and NBD2. The coupling helices of ICL2 and ICL4 interact with NBD2 and NBD1 respectively [14].

As CFTR is translated, each domain folds and can then interact with previously translated domains to form multidomain intermediates [23-29]. The domains are translated in the order: TMD1-NBD1-R-TMD2-NBD2. A minimal construct of TMD1-NBD1-R-TMD2 is able to traffic from the ER [24, 27, 29, 30]. Finally, NBD2 undergoes a posttranslational incorporation into the CFTR structure [24]. A mutation in one domain can destabilize other domains in CFTR, highlighting the importance of interdomain interactions in this process [29, 31]. In this study, we use specific mutations in TMD1 to test this process.

The first domain, TMD1, contains extracellular loops, six transmembrane spans (TMs), and cytosolic regions including an N-terminal region, two ICLs, and a linker that connects to NBD1. The surface formed between the ICL1 coupling helix and NBDs is a mix of hydrophobic and hydrophilic interactions, whereas the surface formed by ICL2 is largely hydrophobic [14]. During TMD1 translation, regions that later form interdomain interactions are present for minutes until interaction partners are produced. Whether ICL structure forms in the TMD1 biogenic intermediate is not known.

CF-associated mutations are within the TMs and ICLs of TMD1 (www.genet.sickkids.on.ca). Many of these mutations introduce charge into the hydrophobic TM spans, and are thereby predicted to disrupt TMD1 structure [28]. In TM1, the CF-causing G85E and G91R mutations result in accumulation of CFTR in the ER [32]. These mutants maintain normal CFTR TM1 topology and have multidomain constructs with reduced stability [33]. The G85E mutant destabilizes the TM1 span in the ER membrane, producing an early defect for cellular recognition [32]. It is not clear whether G85E and G91R cause cellular recognition of TMD1 alone or if this requires multidomain intermediates.

CF-associated mutations in ICL1 and ICL2 are mostly in the helical bundle with few near the coupling helices (www.genet.sickkids.on.ca) [34]. For comparison, ICL4 interacts near the F508 position and contains far more CF-mutations [35, 36]. Predicted ICL1 and ICL2 amino acid residues include 141 to 194 and 242 to 307 respectively [14]. These sequences contain

four disease-causing mutations that inhibit maturation of CFTR (H139R, G149R, D192G, and R258G) and two that reduce CFTR function (G178R and E193K) [34]. Specific residues are predicted as critical for ICL1 and ICL2 interactions with NBDs, including Y275 and W277 with NBD2 [17, 37] and D173, S169, and R170 with nucleotide and NBD1 [17]. These mutations have not been experimentally tested in multidomain folding intermediates.

The $\Delta F508$ mutation alters multiple steps during CFTR folding. The $\Delta F508$ -NBD1 domain has tends to aggregate, is destabilized, and folds less efficiently, indicating a domain folding problem [26, 38-41]. NBD1 structure forms during translation [25] and $\Delta F508$ effects are apparent during translation [24, 41]. In full-length $\Delta F508$ -CFTR crosslinks between the TMDs and NBDs are altered [15, 42] and multidomain folding intermediates are destabilized [24, 27, 29, 31, 43]. The $\Delta F508$ mutant effects can be rescued independently by suppressor mutations within NBD1 [41, 44-46] and by a suppressor mutation in ICL4 [47]. Both these steps have additive effects on rescue of the $\Delta F508$ mutant protein [48]. The additive effects on rescuing and potentiating $\Delta F508$ mutant protein is essential for therapeutics used to treat CF patients [49]. Importantly, current clinical therapeutics stabilize TMD1 [50, 51]. Understanding the role of TMD1 in CFTR folding is important for ongoing therapeutic development.

This study examines biogenic intermediates and full-length CFTR containing CF-causing mutations in TM1, basic residue mutations in the ICL1 and ICL2 coupling helices, and the $\Delta F508$ mutation to test their roles in CFTR domain and multidomain folding.

Materials and methods

Plasmids, DNA techniques

An expression plasmid of full-length, wild type CFTR (pCMV-CFTR- pBQ4.7) was a gift from J. Rommens (The Hospital for Sick Children, Toronto) and was mutagenized using standard protocols for site-directed mutagenesis [52]. Site-directed mutagenesis was

performed by PCR techniques using PfuUltra high-fidelity DNA Polymerase (Stratagene). All mutations were confirmed by DNA sequencing. In these constructs, two stop codons were introduced after amino acid residues 414, 670, or 1174. For TMD1 C-terminal truncations, stop codons were introduced after positions 414, 388, and 354 in wild type or G85E mutant CFTR generated from the pBQ4.7 CFTR cloned into a pBI bidirectional vector. An ECL1 glycosylation site in CFTR is previously described [32]. Stop codons were introduced into this construct to make truncation mutants.

Multiple sequence alignment

ABC transporters of the exporter type include Sav1866 (pdb 2ONJ and 2HYD), MsbA from *E. coli* (pdb 3B5W), MsbA from *V. cholerae* (3B5X), MsbA from *S. typhimurium* (3B60), P-glycoprotein from mouse (pdb 3G5U, 3G60, 3G61). A multisequence alignment was by entering the UniProt accession numbers for these proteins into ClustalW and selected regions examined. UniProt accession numbers are human CFTR (P13569), Sav1866 (Q99T13), MsbA from *E. coli* (P60752), *V. cholerae* (Q9KQW9), and *S. typhimurium* (P63359), and P-glycoprotein from mouse (P21447). The coupling helices for CFTR were selected based on the identified helices in the highest resolution structure, Sav1866 [7].

Mammalian cell protein expression

HEK293 cells (American Type Culture Collection, ATCC) were maintained in Dulbecco's Modified Eagle Medium (DMEM, Invitrogen) supplemented with 10% Fetal Calf Serum (Gemini Bio-Products), 50µg mL⁻¹ penicillin, and 50 units mL⁻¹ streptomycin using standard culture techniques. CFTR constructs were transfected at 20,000 cells per mL in suspension using polyethylenimine, (PEI Polysciences Inc.) transfection reagent, plated in 24 well plates, and expressed for 48 hours. Cell lysis was performed in RIPA buffer (20mM tris pH7.6, 150mM

NaCl, 0.1% SDS, 1% IGEPAL, 0.5% deoxycholic acid, 1mg EDTA-free protease inhibitor tablet (Roche)) at 4°C and centrifuged at 13000g's to generate cleared lysate. Sample buffer (60mM tris pH6.8, 5% glycerol, 2% SDS, bromophenol blue, 280mM β -mercaptoethanol) was added to cleared lysate and incubated at 37°C for 20 minutes.

Western blot analysis

SDS-PAGE analysis of full-length CFTR was performed on equal volumes of lysate. Samples were transferred to PVDF Immobilon membranes (Millipore) and Western blotting performed using CFTR antibodies or actin antibody (Millipore). The CFTR antibody utilized for full-length protein was 596 against amino acids 1204-1211 (UNC School of Medicine), for TMD2x was 570 against amino acids 731-742 (UNC School of Medicine), and for TMD1x, NBD1x, and Rx was MM13-4 (AbCAM). The secondary antibody for all primary antibodies was peroxidase-AffiniPure goat anti-mouse IgG (Jackson ImmunoResearch Laboratories, Inc.). All Western blots were developed using Amersham ELCPlus Western blotting detection reagent (GE Healthcare).

Quantitation

Signal from all membranes was collected by film and using a Typhoon 9410 Variable Mode Imager (GE Healthsciences). The TMD1x, NBD1x, and Rx samples were quantified from film and the TMD2x and full-length samples were quantified from Typhoon. Quantification was performed using Image J Software. Each membrane contained all samples, and individual sample intensity was normalized to total membrane sample intensity. This was then compared across membranes.

Glycosylation analysis

Glycosylation analysis of the ECL1 site was performed on HEK293 cells transiently transfected with the ECL1 site constructs, expressed for 48 hours, and lysed in RIPA buffer. 500U of PNGaseF (New England Biolabs) or Endoglycosidase H (New England Biolabs) was added to 40 μ L of cleared lysate and incubated at 37° for 2 hours. Sample buffer was added to the reaction and incubated at 37° for 20 minutes before SDS-PAGE and Western blot analysis using antibody MM13-4 (AbCAM). For each transfected construct, mock treated lysate was loaded immediately next to EndoH and PNGaseF treated lysate to detect gel shifts that occur upon glycosylation removal.

Results

Mutations in TMD1 alter CFTR cellular trafficking

TMD1 is an integral membrane domain and has amino acid residues in the cytosol, ER membrane, and ER lumen during CFTR translation. Positions in TM1, ICL1, and ICL2 known to disrupt CFTR folding or with predicted locations near the ICL coupling helices were perturbed by introduction of charge. The positions selected are mapped onto the homologous Sav1866 structure (Fig 1A). The CF-causing mutations, G85E and G91R, are in TM1. The G178K, M265R, and W277R CF-associated mutations are in ICL1 and ICL2. These mutations are predicted to introduce a basic residue close to the coupling helices (www.genet.sickkids.on.ca). The coupling helices for CFTR were modeled by aligning the sequences of homologous exporter structures and selecting residues in CFTR utilizing the Sav1866 structure (Fig 1B). The mutated positions within CFTR and respective positions in other transporters are highlighted (Fig 1B). Positions were mutated to the basic amino acid residues lysine or arginine.

The TM1, ICL1, and ICL2 mutations were introduced into full-length CFTR, transiently transfected in HEK293 cell culture, and CFTR maturation monitored by Western blot analysis (Fig 1C). The progression of CFTR as it traffics in the cell is reflected by the glycosylation status of two natural N-linked glycosylation sites in TMD2 that become core glycosylated in the ER (Band B) and are then modified by complex glycosylation in the Golgi (Band C). The $\Delta F508$ mutation accumulates in the ER as Band B. The CF-causing TM1 mutants G85E and G91R accumulate in the ER, indicated by the sole presence of Band B. Mutations at the test positions are integrated into the ER and mature at a variety of efficiencies as compared to WT. G85E, G91R, $\Delta F508$, S168K, S169K, V171K, D173K, M265R, I266K, V272K, and A274K were chosen for further study based on inefficient maturation or predicted location in the ICL-NBD interfaces.

Generation of CFTR biogenic intermediates

We made CFTR multidomain biogenic intermediates to examine TMD1 interactions with other domains. To generate biogenic intermediates, stop codons, designated by x, were introduced after different domains based on the structural model, resulting in constructs producing residues 1-414 (TMD1x), 1-670 (NBD1x), 1-836 (Rx), and 1-1174 (TMD2x) (Fig 2A). Consistent with other studies [27, 29], the full-length and TMD2x constructs traffic from the ER and are complex glycosylated in the Golgi (Fig 2B). The TMD1x, NBD1x, and Rx constructs do not contain the natural N-linked glycosylation sites. Membrane integration and trafficking of these constructs was monitored using a reporter glycosylation site introduced in extracellular loop 1 [32]. The TMD1x, NBD1x, and Rx constructs are core glycosylated, but not complex glycosylated, indicating integration into the ER membrane but inefficient trafficking to the Golgi (Figs 2B and 2C).

TMD1 forms cotranslational structure that then can interact with other domains or proteins. The C-terminal boundary of TMD1 was examined by making constructs containing different C-terminal residues. Stop codons were introduced after residues 354 (354x), 388 (388x), and 414 (414x) to include all TM6, an additional linker region, and the N-terminal residues of NBD1 respectively (Fig 3A). Steady state levels of transiently transfected constructs in HEK293 cells were examined by Western blot analysis for wild type and G85E mutant protein (Fig 3B). The 354x construct expresses at low levels for both wild type and G85E, suggesting this construct is less stable (Fig 3B). The wild type 388x and 414x constructs have higher steady-state levels compared to the G85E 388x and G85E 414x constructs (Fig 3B). This suggests that G85E effects are identifiable at steady state levels in TMD1 constructs containing more than 388 residues.

TMD1 mutant impact on TMD1x

Selected TMD1 mutations were examined in the TMD1x biogenic intermediate containing residues 1-414 (Fig 4A). Steady state levels of transiently transfected constructs in HEK293 cells were examined by Western blot analysis, quantified, and normalized to total image CFTR signal (Fig 4B). The CF-causing mutants G85E and G91R have significantly reduced steady-state levels with respect to wild type. The mutant D173K replaces an acidic with a basic residue and has a significantly decreased level with respect to wild type. Mutations that replace a polar residue with a basic residue, S168K and S169K, had no effect on steady-state TMD1x. Many mutants that replace a hydrophobic residue with a positively charged residue (V171K, I266K, V272K, and A274K) have an increased steady-state level of TMD1x. Importantly, the mutant effects in TMD1x are different than in full-length, indicating other domains are required for their full impact.

TMD1 mutant effects are less apparent in NBD1x and Rx

The TMD1 mutants were studied in the NBD1x construct. The NBD1x construct contains residues 1-670 of CFTR (Fig 5A). In this construct, mutations can disrupt an individual domain or interdomain interactions. Steady-state levels of the constructs were monitored after transient transfection in HEK293 cells by Western blot analysis (Fig 5B). The CF-mutant G85E significantly reduces steady state levels as compared to wild type, indicating NBD1 does not rescue nor amplify the defect. G91R and D173K trend to a lower expression, however unlike the TMD1x construct this is not statistically significant. Again, multiple mutants that replace a hydrophobic residue with a positively charged residue tend to increase steady-state levels of NBD1x. Most mutants have NBD1x levels the same as wild type, suggesting that the cell maintains steady state levels of NBD1x for wild type and most mutants in a similar fashion.

The R domain has potential interaction surfaces for a biogenic intermediate. Other studies found that CFTR constructs truncated after the R domain had a longer half-life than shorter constructs both in cell culture and *in vitro*, suggesting an increased stability [33, 43]. We examined the selected TMD1 mutants in the Rx construct containing residues 1-836 (Fig 6A). Steady state levels of no mutants were statistically different from wild type, though a decreased trend is seen for G85E, G91R, and D173K (Fig 6B). Again, the mutants that replace a hydrophobic residue with a positively charged residue show a tendency to increase steady-state levels of Rx. Mutants have Rx levels more like wild type steady-state cellular levels than either the TMD1x or NBD1x biogenic intermediates.

TMD1 mutant effects are dramatic in TMD2x

The TMD2x construct contains the first four domains, TMD1-NBD1-R-TMD2, and forms a trafficking competent structure that can be monitored by glycosylation of sites in TMD2 (Fig

2B). The TM1, ICL1, and ICL2 mutations were examined in TMD2x containing residues 1 to 1174 of CFTR (Fig 7A). The CF-causing mutants G85E and G91R accumulate in the ER in TMD2x indicated by presence of solely Band B* (Fig 7B). Most ICL mutants have a reduction in steady-state levels of trafficked TMD2x, with significant reduction in band C* (Figs 7B). This is a striking difference from the effect of TMD1x, NBD1x, and Rx constructs. This suggests that a stable TMD2x structure is disrupted by the TMD1 mutants, and this abnormal structure is recognized by the cell resulting in reduced steady-state levels. An exception is that D173K is like wild type, suggesting that the presence of TMD2 improved the significant defect in the TMD1x and NBD1x constructs. The A274K mutant remains like wild type. Interestingly, the V272K mutant shows a small but significant deviation from wild type that is less pronounced than the nearby M265R and I266K mutants. None of the mutants show a difference in the ER steady state levels of mutants as indicated by Band B* (Fig 7B).

NBD2 contributes to stabilization of CFTR

Full-length, functional CFTR contains extensive interactions between all five domains. Its cellular trafficking is monitored by glycosylation of sites in TMD2. The TMD1 mutants were examined in a full-length construct containing residues 1-1480 of CFTR (Fig 8A). The CF-causing mutants G85E and G91R accumulate in the ER without trafficking, indicated by a lack of Band C (Figs 8B). The D173K and A274K mutants have similar levels of Band C compared to wild type, which is like their effect in the TMD2x construct. In full-length, all other ICL mutants have a significantly reduced steady state level of trafficked construct, indicated by Band C. Importantly, the efficiency of ICL mutant trafficking is higher in the full-length protein than in TMD2x. This is consistent with NBD2 conferring stabilization and improved trafficking of CFTR. A notable exception is the V272K mutant, which exhibits a worse efficiency of trafficking in the full-length protein compared to TMD2x. The V272 position is centrally located within the

coupling helix of ICL2, which forms major interactions with NBD2 that might only be disrupted in the full-length construct. In full-length constructs, a decrease in mutant trafficking from the ER is detected by a significant increase in Band B for mutants G85E, G91R, S168K, M265R, and I266K. Both mutants with wild type like trafficking, D173K and A274K, also have wild type like levels of Band B.

Δ F508 mutant effects are largely revealed in TMD2x

The biogenic intermediates containing the Δ F508 mutation were examined to determine the intermediate most significantly deviating from CFTR steady state levels. Multiple NBD1x constructs (n=9) and Rx constructs (n=7) were generated to minimize a plasmid specific effect. Constructs were transiently expressed in HEK293 cell culture in 6 biologic replicates, with multiple constructs in some experiments (Fig 9A). The steady-state levels of NBD1x Δ F508 is significantly reduced from wild type. The steady-state level of Rx Δ F508 is like wild type (Fig 9A). The Δ F508 mutant was examined in the TMD2x and full-length constructs. Constructs were transiently expressed in HEK293 cell culture and Bands B* and C* or Bands B and C quantified (Fig 9B). Δ F508 has dramatically lower steady-state levels of Band C* and Band C in TMD2x and full-length constructs respectively (Fig 9B). While early effects of Δ F508 are identifiable in NBD2x, the cell makes a major distinction between wild type and Δ F508 in TMD2x (Fig 9B). Similar to our other studies mutants, in TMD2x Band B* for Δ F508 is no different from wild type. In the full-length construct, Δ F508 has a significant accumulation of Band B in the ER (Fig 9B). A model depicts potential interactions between the domains of Δ F508-CFTR during translation with red arrows representing possible interdomain disruptions in the TMD2x and full-length constructs (Fig 9C).

Discussion

This study examines the role of TMD1 in formation of CFTR domain structure and interdomain intermediates. The structural units are disrupted by specific mutations in TMD1, including CF-causing mutations. CF-associated mutants and polymorphisms (www.genet.sickkids.on.ca) have been identified throughout TMD1. Furthermore, TMDs are stabilized by CF therapeutics. The corrector VX-809 (lumacaftor) stabilizes TMD1 [50] and also stabilizes TMD1 mutant CFTR and $\Delta F508$ -CFTR [51]. Understanding TMD1 and its role within the formation of CFTR structure is increasingly important since this domain is promising for future therapeutics.

Many disease-associated mutations introduce charge into TMD1 TM spans [28] or ICLs [34]. More CF-causing mutants have been identified in ICL1 than in ICL2. Of these, few are near the coupling helices suggesting that mutations in these regions do not result in CF or are not represented in the database. We found that introduction of basic residues in many positions in ILC1 and ICL2 did not completely block CFTR trafficking (Fig 1C). Thus, few individual residues in ICL1 and ICL2 have a major impact on the formation of the ICL-NBD interface. A similar finding occurs for growth hormone binding to the human growth hormone receptor, whereby few residue mutations in the protein interface alter binding affinity [53]. This results from the interface ability to alter binding energetics [54] and remodel [55] to maintain similar binding, which could relate to the described observations for CFTR ICL-NBD interactions. We identified selected mutations and studied them in biogenic intermediates. A model depicts the earliest biogenic intermediate wherein a persistent TMD1 mutant impact was found (Fig 10).

CFTR is monitored by protein folding and quality control proteins during translation [56]. Quality control involves a protein complex containing the E3 RMA1, E2 Ubc6e, and Derlin-1 interacting with N-terminal regions, and then the Hcs70/CHIP E3 complex interacting during and after more C-terminal regions are produced [43, 56-58]. TMD1 mutants could be recognized by

either or both complexes. G85E mutant CFTR is recognized by the Derlin-1 associated cellular quality control machinery early during biogenesis [56, 59, 60]. Derlin-1 mediates the retro translocation and ER-associated degradation of misfolded proteins [61, 62]. Our study shows that G85E has effects within TMD1 (Fig 4). Our prior work showed that G85E destabilizes TM1 [63]. TMD1 likely contains structural signals for the recognition and degradation of the G85E mutant by quality control machinery. The G91R mutant is associated with the RMA1 E3 ubiquitin ligase that interacts with Derlin-1 [56]. G91R has effects within TMD1 (Fig 4). Yet, *in vitro* work suggests G85E and G91R are recognized in multidomain constructs [33]. For both mutants, it is likely that multiple levels of recognition are involved in cellular mistrafficking and reduced steady state levels.

TMD1 makes a proteolytically protected structure during translation [25]. Integration of the TM spans is required for its formation. If a TM span is marginally hydrophobic, the efficiency of integration becomes more efficient with flanking regions and other TM spans present [64]. In TMD1, TM6 has several basic residues that are important for pore formation [65]. TM6 *in vitro* integration is inefficient, and truncation at 364 results in TM6 incorrect integration [66]. In our study, we found that moving the end of TMD1 more C-terminal from 354x to 388x to 414x increased the steady state level of the TMD1x construct (Fig 3). In the presence of G85E, these construct levels were only reduced for 388x and 414x. This suggests that the 354x construct is not stable with reduced steady state levels and that G85E is not able to make this detectably lower. Multiple CF-associated mutants and polymorphisms are found in the region between 388-414 (www.genet.sickkids.on.ca). Our study supports an essential role for these C-terminal residues of TMD1 for stable domain formation and thereby CFTR function, which could be perturbed by mutations in this region.

ICL1 and ICL2 form extensive interdomain interactions in the final CFTR structure, yet their role during folding is not clear. In this study, we focused on the ICL coupling helices, which are cytosolic and approximately 25Å away from the membrane in homologous structures [7]. In

CFTR the predicted coupling helix and NBD interfaces are both hydrophilic and hydrophobic for ICL1, and largely hydrophobic for ICL2 [14]. Surprisingly, the TMD1x, NBD1x, and Rx constructs trend towards increased steady-state levels when a hydrophobic residue is mutated to a basic residue. Thereby, the coupling helix hydrophobic residues may decrease level of biogenic intermediates. Hsc70 recognizes short hydrophobic sequences [67] and could form an interaction with the coupling helices. Furthermore, the Hsc70/CHIP machinery is found in association with TMD1 [43] and targets immature CFTR for degradation [43, 57, 58]. Our work supports a hypothesis that the coupling helices are monitored during CFTR folding through interactions with Hsc70, and persistent hydrophobic residue exposure can lead to targeting of misfolded CFTR for degradation.

The impact of most positively charged ICL mutants in TMD1 is detected in TMD2x (Fig 10). TMD2x contains domains wherein TMD rearrangements, ICL helical bundle structure, and interactions with NBD1 can occur. The TMD1 mutants have reduced trafficking efficiency from the ER, suggesting an intermediate that is recognized in the presence of TMD2 or failure of appropriate trafficking signals. A di-acidic motif within NBD1 is important for COPII-dependent trafficking of CFTR from the ER [68]. The exposure of this motif could depend on a structural rearrangement that occurs after TMD2 is produced. If TMD1 mutants do not have the native multidomain rearrangements, then this site may not be revealed resulting in mistrafficking.

In the full-length protein, most lysine mutants trafficked more efficiently than in the TMD2x construct. The addition of NBD2 confers a greater folding efficiency and trafficking from the ER. Thus, NBD2 is not strictly required for CFTR trafficking [27, 29, 30, 47] however its posttranslational association into the CFTR structure [24] imparts stability to the other domains. The exceptions to this are for the mutations D173K, A274K, and V272K. The D173K mutation in ICL1 reduces biogenic intermediates prior to TMD2 and behaves more like wild type in the presence of TMD2. One explanation is that this position is involved in formation of a nonnative contact that is lost in the constructs containing TMD2. The A274K mutation is like wild type in

all constructs, suggesting this position is tolerant to mutation. The V272K mutant lowered trafficking efficiency in the full-length construct more dramatically than in TMD2x, which lacks NBD2. In CFTR homology model, position V272 points between the two helical stems of ICL2 to support ICL structure, thereby stabilizing the ICL2-NBD2 interface indirectly. An equivalent position in ICL4, L1065, has a CF-causing mutation of L1065P that accumulates in the ER [35, 36]. We propose that stabilization of the ICL and coupling helix structures rather than direct stabilization of the NBD/ICL interfaces should be considered a potential therapeutic target for improving mutant CFTR folding and function.

The major recognition of many CFTR mutants requires TMD2, which is also true for $\Delta F508$ (Fig 9). Some reports indicate that $\Delta F508$ minimally reduces stability of TMD1-NBD1 constructs, but significantly reduces stability of constructs containing the R domain [31, 69]. We found that the steady-state levels of NBD1x- $\Delta F508$ is reduced from wild type and the Rx- $\Delta F508$ construct is not altered from wild type. These data are inconsistent with previous reports. In our experimental data, we could observe this difference within the range of expressed individual constructs. We have tried to minimize this possibility by using multiple independently generated constructs in our studies between wild type and $\Delta F508$. In TMD2x- $\Delta F508$, the protein may fail to achieve a complex, multidomain fold that could allow the cell to distinguish misfolded from appropriately folded CFTR. The lower levels of NBD1x- $\Delta F508$ may reflect misfolded domain, and the dramatic decrease in TMD2x- $\Delta F508$ reflects disruption of interdomain interactions and multidomain structure formation. This is consistent with compounds or suppressor mutations for $\Delta F508$ within NBD1 [44, 45, 70] or distant in the CFTR protein [71-73] improving trafficking and function. The importance of targeting multiple steps for $\Delta F508$ -CFTR is highlighted by clinical therapeutics used to treat CF patients [49].

This study identifies roles for TMD1 in CFTR folding and details for CF-causing mutant misfolding. This data is important for understanding CFTR misfolding and for identifying potential therapeutic targets in that process.

Acknowledgments

We thank Juan Mendoza for help with sequence alignment and selection of the intracellular loop residues. We thank J. Rommens for providing the CFTR construct. This research was supported by NIH grants R37 DK049835 (P.J.T.) and K12 HD087023 (Research Scholar, A.E.P.).

References

1. Rommens JM, Iannuzzi MC, Kerem B, Drumm ML, Melmer G, Dean M, et al. Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science*. 1989;245(4922):1059-65. Epub 1989/09/08. PubMed PMID: 2772657.
2. Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science*. 1989;245(4922):1066-73. Epub 1989/09/08. PubMed PMID: 2475911.
3. Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, et al. Identification of the cystic fibrosis gene: genetic analysis. *Science*. 1989;245(4922):1073-80. Epub 1989/09/08. PubMed PMID: 2570460.
4. Rowe SM, Miller S, Sorscher EJ. Cystic fibrosis. *N Engl J Med*. 2005;352(19):1992-2001. Epub 2005/05/13. doi: 352/19/1992 [pii] 10.1056/NEJMra043184. PubMed PMID: 15888700.
5. Holland IB. ABC proteins : from bacteria to man. 1st ed. Amsterdam ; Boston: Academic Press; 2003. xxiii, 647 p. p.
6. Baker JM, Hudson RP, Kanelis V, Choy WY, Thibodeau PH, Thomas PJ, et al. CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. *Nat Struct Mol Biol*. 2007;14(8):738-45. Epub 2007/07/31. doi: nsmb1278 [pii] 10.1038/nsmb1278. PubMed PMID: 17660831.
7. Dawson RJ, Locher KP. Structure of a bacterial multidrug ABC transporter. *Nature*. 2006;443(7108):180-5. Epub 2006/09/01. doi: nature05155 [pii] 10.1038/nature05155. PubMed PMID: 16943773.
8. Dawson RJ, Locher KP. Structure of the multidrug ABC transporter Sav1866 from *Staphylococcus aureus* in complex with AMP-PNP. *FEBS Lett*. 2007;581(5):935-8. Epub 2007/02/17. doi: 10.1016/j.febslet.2007.01.073. PubMed PMID: 17303126.
9. Ward A, Reyes CL, Yu J, Roth CB, Chang G. Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc Natl Acad Sci U S A*. 2007;104(48):19005-10. Epub 2007/11/21. doi: 10.1073/pnas.0709388104. PubMed PMID: 18024585; PubMed Central PMCID: PMC2141898.
10. Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, et al. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science*.

- 2009;323(5922):1718-22. Epub 2009/03/28. doi: 10.1126/science.1168750. PubMed PMID: 19325113; PubMed Central PMCID: PMC2720052.
11. Johnson ZL, Chen J. Structural Basis of Substrate Recognition by the Multidrug Resistance Protein MRP1. *Cell*. 2017;168(6):1075-85 e9. Epub 2017/02/28. doi: 10.1016/j.cell.2017.01.041. PubMed PMID: 28238471.
12. Zhang Z, Chen J. Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator. *Cell*. 2016;167(6):1586-97 e9. Epub 2016/12/03. doi: 10.1016/j.cell.2016.11.014. PubMed PMID: 27912062.
13. Liu F, Zhang Z, Csanady L, Gadsby DC, Chen J. Molecular Structure of the Human CFTR Ion Channel. *Cell*. 2017;169(1):85-95 e8. Epub 2017/03/25. doi: 10.1016/j.cell.2017.02.024. PubMed PMID: 28340353.
14. Mendoza JL, Thomas PJ. Building an understanding of cystic fibrosis on the foundation of ABC transporter structures. *J Bioenerg Biomembr*. 2007;39(5-6):499-505. Epub 2007/12/15. doi: 10.1007/s10863-007-9117-7. PubMed PMID: 18080175.
15. Serohijos AW, Hegedus T, Aleksandrov AA, He L, Cui L, Dokholyan NV, et al. Phenylalanine-508 mediates a cytoplasmic-membrane domain contact in the CFTR 3D structure crucial to assembly and channel function. *Proc Natl Acad Sci U S A*. 2008;105(9):3256-61. Epub 2008/02/29. doi: 0800254105 [pii] 10.1073/pnas.0800254105. PubMed PMID: 18305154.
16. Mornon JP, Lehn P, Callebaut I. Molecular models of the open and closed states of the whole human CFTR protein. *Cell Mol Life Sci*. 2009;66(21):3469-86. Epub 2009/08/27. doi: 10.1007/s00018-009-0133-0. PubMed PMID: 19707853.
17. Mornon JP, Lehn P, Callebaut I. Atomic model of human cystic fibrosis transmembrane conductance regulator: membrane-spanning domains and coupling interfaces. *Cell Mol Life Sci*. 2008;65(16):2594-612. Epub 2008/07/04. doi: 10.1007/s00018-008-8249-1. PubMed PMID: 18597042.
18. Moody JE, Millen L, Binns D, Hunt JF, Thomas PJ. Cooperative, ATP-dependent association of the nucleotide binding cassettes during the catalytic cycle of ATP-binding cassette transporters. *J Biol Chem*. 2002;277(24):21111-4. Epub 2002/04/20. doi: 10.1074/jbc.C200228200 C200228200 [pii]. PubMed PMID: 11964392.
19. Smith PC, Karpowich N, Millen L, Moody JE, Rosen J, Thomas PJ, et al. ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Mol Cell*. 2002;10(1):139-49. Epub 2002/08/02. doi: S1097276502005762 [pii]. PubMed PMID: 12150914.
20. Mense M, Vergani P, White DM, Altberg G, Nairn AC, Gadsby DC. In vivo phosphorylation of CFTR promotes formation of a nucleotide-binding domain heterodimer. *EMBO J*. 2006;25(20):4728-39. Epub 2006/10/13. doi: 7601373 [pii] 10.1038/sj.emboj.7601373. PubMed PMID: 17036051.
21. Riordan JR. CFTR function and prospects for therapy. *Annu Rev Biochem*. 2008;77:701-26. Epub 2008/02/29. doi: 10.1146/annurev.biochem.75.103004.142532. PubMed PMID: 18304008.
22. Gadsby DC, Vergani P, Csanady L. The ABC protein turned chloride channel whose failure causes cystic fibrosis. *Nature*. 2006;440(7083):477-83. Epub 2006/03/24. doi: nature04712 [pii] 10.1038/nature04712. PubMed PMID: 16554808; PubMed Central PMCID: PMC2720541.

23. Lukacs GL, Mohamed A, Kartner N, Chang XB, Riordan JR, Grinstein S. Conformational maturation of CFTR but not its mutant counterpart (delta F508) occurs in the endoplasmic reticulum and requires ATP. *EMBO J.* 1994;13(24):6076-86. Epub 1994/12/15. PubMed PMID: 7529176.
24. Du K, Sharma M, Lukacs GL. The DeltaF508 cystic fibrosis mutation impairs domain-domain interactions and arrests post-translational folding of CFTR. *Nat Struct Mol Biol.* 2005;12(1):17-25. Epub 2004/12/28. doi: nsmb882 [pii] 10.1038/nsmb882. PubMed PMID: 15619635.
25. Kleizen B, van Vlijmen T, de Jonge HR, Braakman I. Folding of CFTR is predominantly cotranslational. *Mol Cell.* 2005;20(2):277-87. Epub 2005/10/26. doi: S1097-2765(05)01607-2 [pii] 10.1016/j.molcel.2005.09.007. PubMed PMID: 16246729.
26. Thibodeau PH, Brautigam CA, Machius M, Thomas PJ. Side chain and backbone contributions of Phe508 to CFTR folding. *Nat Struct Mol Biol.* 2005;12(1):10-6. Epub 2004/12/28. doi: nsmb881 [pii] 10.1038/nsmb881. PubMed PMID: 15619636.
27. Cui L, Aleksandrov L, Chang XB, Hou YX, He L, Hegedus T, et al. Domain interdependence in the biosynthetic assembly of CFTR. *J Mol Biol.* 2007;365(4):981-94. Epub 2006/11/23. doi: S0022-2836(06)01511-7 [pii] 10.1016/j.jmb.2006.10.086. PubMed PMID: 17113596.
28. Cheung JC, Deber CM. Misfolding of the cystic fibrosis transmembrane conductance regulator and disease. *Biochemistry.* 2008;47(6):1465-73. Epub 2008/01/16. doi: 10.1021/bi702209s. PubMed PMID: 18193900.
29. Du K, Lukacs GL. Cooperative assembly and misfolding of CFTR domains in vivo. *Mol Biol Cell.* 2009;20(7):1903-15. Epub 2009/01/30. doi: E08-09-0950 [pii] 10.1091/mbc.E08-09-0950. PubMed PMID: 19176754; PubMed Central PMCID: PMC2663924.
30. Pollet JF, Van Geffel J, Van Stevens E, Van Geffel R, Beauwens R, Bollen A, et al. Expression and intracellular processing of chimeric and mutant CFTR molecules. *Biochim Biophys Acta.* 2000;1500(1):59-69. Epub 1999/11/24. doi: S0925-4439(99)00088-5 [pii]. PubMed PMID: 10564718.
31. Rosser MF, Grove DE, Chen L, Cyr DM. Assembly and misassembly of cystic fibrosis transmembrane conductance regulator: folding defects caused by deletion of F508 occur before and after the calnexin-dependent association of membrane spanning domain (MSD) 1 and MSD2. *Mol Biol Cell.* 2008;19(11):4570-9. Epub 2008/08/22. doi: E08-04-0357 [pii] 10.1091/mbc.E08-04-0357. PubMed PMID: 18716059.
32. Patrick AE, Karamyshev AL, Millen L, Thomas PJ. Alteration of CFTR transmembrane span integration by disease-causing mutations. *Molecular biology of the cell.* 2011;22(23):4461-71. Epub 2011/10/15. doi: 10.1091/mbc.E11-05-0396. PubMed PMID: 21998193; PubMed Central PMCID: PMC3226467.
33. Xiong X, Bragin A, Widdicombe JH, Cohn J, Skach WR. Structural cues involved in endoplasmic reticulum degradation of G85E and G91R mutant cystic fibrosis transmembrane conductance regulator. *J Clin Invest.* 1997;100(5):1079-88. Epub 1997/09/01. doi: 10.1172/JCI119618. PubMed PMID: 9276724; PubMed Central PMCID: PMC508282.
34. Seibert FS, Jia Y, Mathews CJ, Hanrahan JW, Riordan JR, Loo TW, et al. Disease-associated mutations in cytoplasmic loops 1 and 2 of cystic fibrosis transmembrane conductance

- regulator impede processing or opening of the channel. *Biochemistry*. 1997;36(39):11966-74. Epub 1997/10/08. doi: 10.1021/bi9712652. PubMed PMID: 9305991.
35. Seibert FS, Linsdell P, Loo TW, Hanrahan JW, Clarke DM, Riordan JR. Disease-associated mutations in the fourth cytoplasmic loop of cystic fibrosis transmembrane conductance regulator compromise biosynthetic processing and chloride channel activity. *The Journal of biological chemistry*. 1996;271(25):15139-45. Epub 1996/06/21. PubMed PMID: 8662892.
36. Cotten JF, Ostedgaard LS, Carson MR, Welsh MJ. Effect of cystic fibrosis-associated mutations in the fourth intracellular loop of cystic fibrosis transmembrane conductance regulator. *The Journal of biological chemistry*. 1996;271(35):21279-84. Epub 1996/08/30. PubMed PMID: 8702904.
37. He L, Aleksandrov AA, Serohijos AW, Hegedus T, Aleksandrov LA, Cui L, et al. Multiple membrane-cytoplasmic domain contacts in the cystic fibrosis transmembrane conductance regulator (CFTR) mediate regulation of channel gating. *J Biol Chem*. 2008;283(39):26383-90. Epub 2008/07/29. doi: M803894200 [pii] 10.1074/jbc.M803894200. PubMed PMID: 18658148.
38. Lewis HA, Zhao X, Wang C, Sauder JM, Rooney I, Noland BW, et al. Impact of the deltaF508 mutation in first nucleotide-binding domain of human cystic fibrosis transmembrane conductance regulator on domain folding and structure. *J Biol Chem*. 2005;280(2):1346-53. Epub 2004/11/06. doi: M410968200 [pii] 10.1074/jbc.M410968200. PubMed PMID: 15528182.
39. Qu BH, Thomas PJ. Alteration of the cystic fibrosis transmembrane conductance regulator folding pathway. *The Journal of biological chemistry*. 1996;271(13):7261-4. Epub 1996/03/29. PubMed PMID: 8631737.
40. Thomas PJ, Shenbagamurthi P, Sondek J, Hullihen JM, Pedersen PL. The cystic fibrosis transmembrane conductance regulator. Effects of the most common cystic fibrosis-causing mutation on the secondary structure and stability of a synthetic peptide. *The Journal of biological chemistry*. 1992;267(9):5727-30. Epub 1992/03/25. PubMed PMID: 1372891.
41. Hoelen H, Kleizen B, Schmidt A, Richardson J, Charitou P, Thomas PJ, et al. The primary folding defect and rescue of DeltaF508 CFTR emerge during translation of the mutant domain. *PLoS One*. 2010;5(11):e15458. Epub 2010/12/15. doi: 10.1371/journal.pone.0015458. PubMed PMID: 21152102; PubMed Central PMCID: PMC2994901.
42. Chen EY, Bartlett MC, Loo TW, Clarke DM. The DeltaF508 mutation disrupts packing of the transmembrane segments of the cystic fibrosis transmembrane conductance regulator. *J Biol Chem*. 2004;279(38):39620-7. Epub 2004/07/24. doi: 10.1074/jbc.M407887200 M407887200 [pii]. PubMed PMID: 15272010.
43. Meacham GC, Lu Z, King S, Sorscher E, Tousson A, Cyr DM. The Hdj-2/Hsc70 chaperone pair facilitates early steps in CFTR biogenesis. *EMBO J*. 1999;18(6):1492-505. Epub 1999/03/17. doi: 10.1093/emboj/18.6.1492. PubMed PMID: 10075921.
44. Teem JL, Berger HA, Ostedgaard LS, Rich DP, Tsui LC, Welsh MJ. Identification of revertants for the cystic fibrosis delta F508 mutation using STE6-CFTR chimeras in yeast. *Cell*. 1993;73(2):335-46. Epub 1993/04/23. PubMed PMID: 7682896.
45. Qu BH, Strickland EH, Thomas PJ. Localization and suppression of a kinetic defect in cystic fibrosis transmembrane conductance regulator folding. *The Journal of biological chemistry*. 1997;272(25):15739-44. Epub 1997/06/20. PubMed PMID: 9188468.

46. DeCarvalho AC, Gansheroff LJ, Teem JL. Mutations in the nucleotide binding domain 1 signature motif region rescue processing and functional defects of cystic fibrosis transmembrane conductance regulator delta f508. *The Journal of biological chemistry*. 2002;277(39):35896-905. Epub 2002/07/12. doi: 10.1074/jbc.M205644200. PubMed PMID: 12110684.
47. Thibodeau PH, Richardson JM, 3rd, Wang W, Millen L, Watson J, Mendoza JL, et al. The cystic fibrosis-causing mutation deltaF508 affects multiple steps in cystic fibrosis transmembrane conductance regulator biogenesis. *The Journal of biological chemistry*. 2010;285(46):35825-35. Epub 2010/07/30. doi: 10.1074/jbc.M110.131623. PubMed PMID: 20667826; PubMed Central PMCID: PMC2975206.
48. Mendoza JL, Schmidt A, Li Q, Nuvaga E, Barrett T, Bridges RJ, et al. Requirements for efficient correction of DeltaF508 CFTR revealed by analyses of evolved sequences. *Cell*. 2012;148(1-2):164-74. Epub 2012/01/24. doi: 10.1016/j.cell.2011.11.023. PubMed PMID: 22265409; PubMed Central PMCID: PMC3266553.
49. Southern KW, Patel S, Sinha IP, Nevitt SJ. Correctors (specific therapies for class II CFTR mutations) for cystic fibrosis. *Cochrane Database Syst Rev*. 2018;8:CD010966. Epub 2018/08/03. doi: 10.1002/14651858.CD010966.pub2. PubMed PMID: 30070364; PubMed Central PMCID: PMCPMC6513216.
50. Loo TW, Bartlett MC, Clarke DM. Corrector VX-809 stabilizes the first transmembrane domain of CFTR. *Biochem Pharmacol*. 2013;86(5):612-9. Epub 2013/07/10. doi: 10.1016/j.bcp.2013.06.028. PubMed PMID: 23835419.
51. Ren HY, Grove DE, De La Rosa O, Houck SA, Sopha P, Van Goor F, et al. VX-809 corrects folding defects in cystic fibrosis transmembrane conductance regulator protein through action on membrane-spanning domain 1. *Molecular biology of the cell*. 2013;24(19):3016-24. Epub 2013/08/09. doi: 10.1091/mbc.E13-05-0240. PubMed PMID: 23924900; PubMed Central PMCID: PMC3784376.
52. Sambrook J, Fritsch, E.F., and Maniatis, T. *Molecular Cloning: A laboratory Manual*. 1989 ed. New York: Cold Spring Harbor Laboratory Press; 1989.
53. Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor interface. *Science*. 1995;267(5196):383-6. Epub 1995/01/20. PubMed PMID: 7529940.
54. Pearce KH, Jr., Ultsch MH, Kelley RF, de Vos AM, Wells JA. Structural and mutational analysis of affinity-inert contact residues at the growth hormone-receptor interface. *Biochemistry*. 1996;35(32):10300-7. Epub 1996/08/13. doi: 10.1021/bi960513b. PubMed PMID: 8756685.
55. Atwell S, Ultsch M, De Vos AM, Wells JA. Structural plasticity in a remodeled protein-protein interface. *Science*. 1997;278(5340):1125-8. Epub 1997/11/14. PubMed PMID: 9353194.
56. Younger JM, Chen L, Ren HY, Rosser MF, Turnbull EL, Fan CY, et al. Sequential quality-control checkpoints triage misfolded cystic fibrosis transmembrane conductance regulator. *Cell*. 2006;126(3):571-82. Epub 2006/08/12. doi: S0092-8674(06)00908-1 [pii] 10.1016/j.cell.2006.06.041. PubMed PMID: 16901789.
57. Meacham GC, Patterson C, Zhang W, Younger JM, Cyr DM. The Hsc70 co-chaperone CHIP targets immature CFTR for proteasomal degradation. *Nat Cell Biol*. 2001;3(1):100-5. Epub 2001/01/09. doi: 10.1038/35050509. PubMed PMID: 11146634.
58. Younger JM, Ren HY, Chen L, Fan CY, Fields A, Patterson C, et al. A foldable CFTR{Delta}F508 biogenic intermediate accumulates upon inhibition of the Hsc70-CHIP E3

- ubiquitin ligase. *J Cell Biol.* 2004;167(6):1075-85. Epub 2004/12/22. doi: 10.1083/jcb.200410065. PubMed PMID: 15611333; PubMed Central PMCID: PMC2172621.
59. Sun F, Zhang R, Gong X, Geng X, Drain PF, Frizzell RA. Derlin-1 promotes the efficient degradation of the cystic fibrosis transmembrane conductance regulator (CFTR) and CFTR folding mutants. *J Biol Chem.* 2006;281(48):36856-63. Epub 2006/09/07. doi: M607085200 [pii] 10.1074/jbc.M607085200. PubMed PMID: 16954204.
60. Wang B, Heath-Engel H, Zhang D, Nguyen N, Thomas DY, Hanrahan JW, et al. BAP31 interacts with Sec61 translocons and promotes retrotranslocation of CFTRDeltaF508 via the derlin-1 complex. *Cell.* 2008;133(6):1080-92. Epub 2008/06/17. doi: S0092-8674(08)00616-8 [pii] 10.1016/j.cell.2008.04.042. PubMed PMID: 18555783.
61. Lilley BN, Ploegh HL. A membrane protein required for dislocation of misfolded proteins from the ER. *Nature.* 2004;429(6994):834-40. Epub 2004/06/25. doi: 10.1038/nature02592 nature02592 [pii]. PubMed PMID: 15215855.
62. Ye Y, Shibata Y, Yun C, Ron D, Rapoport TA. A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. *Nature.* 2004;429(6994):841-7. Epub 2004/06/25. doi: 10.1038/nature02656 nature02656 [pii]. PubMed PMID: 15215856.
63. Patrick AE, Thomas PJ. Development of CFTR Structure. *Front Pharmacol.* 2012;3:162. Epub 2012/09/14. doi: 10.3389/fphar.2012.00162. PubMed PMID: 22973227; PubMed Central PMCID: PMC3434365.
64. Hedin LE, Ojemalm K, Bernsel A, Hennerdal A, Illergard K, Enquist K, et al. Membrane insertion of marginally hydrophobic transmembrane helices depends on sequence context. *Journal of molecular biology.* 2010;396(1):221-9. Epub 2009/11/26. doi: 10.1016/j.jmb.2009.11.036. PubMed PMID: 19931281.
65. Cheung M, Akabas MH. Identification of cystic fibrosis transmembrane conductance regulator channel-lining residues in and flanking the M6 membrane-spanning segment. *Biophys J.* 1996;70(6):2688-95. Epub 1996/06/01. doi: 10.1016/S0006-3495(96)79838-7. PubMed PMID: 8744306; PubMed Central PMCID: PMC1225248.
66. Tector M, Hartl FU. An unstable transmembrane segment in the cystic fibrosis transmembrane conductance regulator. *The EMBO journal.* 1999;18(22):6290-8. Epub 1999/11/24. doi: 10.1093/emboj/18.22.6290. PubMed PMID: 10562541; PubMed Central PMCID: PMC1171692.
67. Rudiger S, Buchberger A, Bukau B. Interaction of Hsp70 chaperones with substrates. *Nat Struct Biol.* 1997;4(5):342-9. Epub 1997/05/01. PubMed PMID: 9145101.
68. Wang X, Matteson J, An Y, Moyer B, Yoo JS, Bannykh S, et al. COPII-dependent export of cystic fibrosis transmembrane conductance regulator from the ER uses a di-acidic exit code. *J Cell Biol.* 2004;167(1):65-74. Epub 2004/10/14. doi: 10.1083/jcb.200401035. PubMed PMID: 15479737; PubMed Central PMCID: PMC2172508.
69. Grove DE, Rosser MF, Ren HY, Naren AP, Cyr DM. Mechanisms for rescue of correctable folding defects in CFTRDelta F508. *Mol Biol Cell.* 2009;20(18):4059-69. Epub 2009/07/25. doi: E08-09-0929 [pii] 10.1091/mbc.E08-09-0929. PubMed PMID: 19625452; PubMed Central PMCID: PMC2743624.
70. Sampson HM, Robert R, Liao J, Matthes E, Carlile GW, Hanrahan JW, et al. Identification of a NBD1-binding pharmacological chaperone that corrects the trafficking defect

- of F508del-CFTR. Chem Biol. 2011;18(2):231-42. Epub 2011/02/23. doi: 10.1016/j.chembiol.2010.11.016. PubMed PMID: 21338920.
71. Loo TW, Bartlett MC, Clarke DM. Benzbromarone Stabilizes DeltaF508 CFTR at the Cell Surface. Biochemistry. 2011. Epub 2011/04/28. doi: 10.1021/bi2004813. PubMed PMID: 21520952.
72. Cebotaru L, Vij N, Ciobanu I, Wright J, Flotte T, Guggino WB. Cystic fibrosis transmembrane regulator missing the first four transmembrane segments increases wild type and DeltaF508 processing. The Journal of biological chemistry. 2008;283(32):21926-33. Epub 2008/05/30. doi: 10.1074/jbc.M709156200. PubMed PMID: 18508776; PubMed Central PMCID: PMC2494918.
73. Cormet-Boyaka E, Jablonsky M, Naren AP, Jackson PL, Muccio DD, Kirk KL. Rescuing cystic fibrosis transmembrane conductance regulator (CFTR)-processing mutants by transcomplementation. Proc Natl Acad Sci U S A. 2004;101(21):8221-6. Epub 2004/05/14. doi: 10.1073/pnas.0400459101. PubMed PMID: 15141088; PubMed Central PMCID: PMC419584.

Legends

Fig 1. TMD1 mutant effects on full-length CFTR cellular trafficking.

(A) Mutations mapped onto the Sav1866 structure (pdb 2HYD). Sav1866 is a homodimer, and each protein unit is colored with light gray representing CFTR TMD1 and NBD1 and dark grey representing CFTR TMD2 and NBD2. Respective CFTR mutation positions are shown as spheres for TM1 and Δ F508 (green), ICL1 (light blue), and ICL2 (light red). Mutants that were further studied are a darker blue or red. Views down ICL1 and ICL2 coupling helices axes with the protein backbone as a ribbon and mutated residues as sticks. (B) Alignments of human CFTR with ABC exporters using ClustalW software and sequences for *H. sapiens* CFTR, *S. aureus* Sav1866, *M. musculus* P-glycoprotein, and *S. typhimurium* MsbA (UniProt accession numbers: P13569, Q99T13, P21447, and P63359). Screened CFTR mutant positions and respective positions in transporters are highlighted blue for ICL1 and red for ICL2. The coupling helix for Sav1866 is underlined. (C) TMD1 mutants examined in full-length CFTR. Expression and trafficking in HEK293 cell culture with assessment by Western blot analysis using antibody 596. Band B is core glycosylated in the ER and Band C has been complex glycosylation in the

Golgi. Mutants further studied are bolded. The $\Delta F508$ mutation accumulates in the ER and is a control for Band B. Actin is a loading control.

Fig 2. CFTR biogenic intermediates.

(A) The CFTR biogenic intermediates, residues, and multidomain construct produced by introducing stop codons. (B) Full-length CFTR (1-1480) and biogenic intermediates were transiently expressed in HEK293 cells and assessed by Western blot analysis with mixed CFTR antibodies 570 and MM13-4. TMD2 contains the two natural glycosylation sites for monitoring cellular trafficking of CFTR. Full-length Bands B and C and 1174x core glycosylated Band B* and complex glycosylated Band C* are indicated. (C) A reporter core glycosylation site introduced into TMD1 assessed in TMD1x, NBD1x, and Rx. Unmodified biogenic intermediates or biogenic intermediates with the TMD1 glycosylation site were expressed in HEK293 cells and examined by treatment with glycosidase followed by Western blot analysis with MM13-4 antibody. Glycosidase treatment used EndoH to remove core glycosylation and PNGaseF to remove both core and complex glycosylation. The positions of core glycosylated (b) and non-glycosylated (a) constructs are noted. Unrelated background in the TMD1 glycosylation site Rx sample is designated by a triangle.

Fig 3. TMD1 C-terminal boundary mapping.

(A) A schematic of the TMD1 C-terminal stop codon placement after TM6 (354x), after the linker region (388x), or including N-terminal NBD1 residues (414x) is depicted with the position of G85E indicated (green circle). (B) Wild type (WT) and G85E constructs were transiently transfected into HEK293 cells and monitored by Western blot analysis with the antibody MM13-4. Actin is a loading control.

Fig 4. TMD1 mutants in the TMD1x biogenic intermediate.

(A) Schematic of TMD1x (residues 1-414) with six transmembrane spans, ER luminal, and cytosolic regions. Approximate locations of mutants in TM1 (green circles), ICL1 (blue circle), and ICL2 (red circle) are marked. (B) TMD1x containing a mutation was transiently transfected in HEK293 cells and monitored by Western blot analysis with the antibody MM13-4. The steady-state levels were quantified and normalized to total image TMD1x signal with the average and standard error of the mean plotted (N=4). TM1 mutants are green, ICL1 mutants are blue, and ICL2 mutants are red. Actin is a loading control. Analysis by ANOVA against WT, * $p < 0.05$, ** $p < 0.01$.

Fig 5. TMD1 mutants in the NBD1x biogenic intermediate.

(A) Schematic of NBD1x construct (residues 1-670) with TMD1 and NBD1. Approximate locations of mutants in TM1 (green circles), ICL1 (blue circles), and ICL2 (red circles) are marked. (B) NBD1x containing a mutation was transiently transfected into HEK293 cells and monitored by Western blot analysis with the antibody MM13-4. The steady-state levels were quantified and normalized to total image NBD1x signal with the average and the standard error of the mean plotted (N=4). TM1 mutants are green, ICL1 mutants are blue, and ICL2 mutants are red. Actin is a loading control. Analysis by ANOVA against WT, * $p < 0.05$.

Fig 6. TMD1 mutants in the Rx biogenic intermediate.

(A) Schematic of Rx construct (residues 1-836) with TMD1, NBD1, and R. Approximate locations of mutants in TM1 (green circles), ICL1 (blue circles), and ICL2 (red circles) are marked. (B) Rx containing a mutation was transiently transfected into HEK293 cells and monitored by Western blot analysis with the antibody MM13-4. The steady-state levels were quantified and normalized to total image Rx signal with the average and the standard error of the mean plotted (N=4). TM1 mutants are green, ICL1 mutants are blue, and ICL2 mutants are red. Actin is a loading control. Analysis by ANOVA against WT, * $p < 0.05$.

Fig 7. TMD1 mutants in the TMD2x biogenic intermediate.

(A) Schematic of TMD2x (residues 1-1174x) with TMD1, NBD1, R, and TMD2. Approximate locations of mutants in TM1 (green circles), ICL1 (blue circles), and ICL2 (red circles) are marked. (B) TMD2x containing a mutation was transiently transfected into HEK293 cells and monitored by Western blot analysis with the antibody 570. CFTR steady-state levels of Band B* and Band C* were quantified and normalized to total image TMD2x signal. The average is plotted with the standard error of the mean (N=4). TM1 mutants are green, ICL1 mutants are blue, and ICL2 mutants are red. Actin is a loading control. Analysis by ANOVA against WT, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Fig 8. TMD1 mutants in full-length CFTR.

(A) Schematic of full-length CFTR (residues 1-1480) depicts TMD1, NBD1, R, TMD2, and NBD2. Approximate locations of mutants in TM1 (green circles), ICL1 (blue circles), and ICL2 (red circles) are marked. (B) A full-length construct containing a mutant was transiently transfected into HEK293 cells and monitored by Western blot analysis with the antibody 596. CFTR steady-state levels of Band B and Band C were quantified and normalized to total image full-length signal. The average is plotted with the standard error of the mean (N=3). TM1 mutants are green, ICL1 mutants are blue, and ICL2 mutants are red. Actin was monitored as a loading control. Analysis by ANOVA against WT, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Fig 9. $\Delta F508$ dramatically alters TMD2x and full-length CFTR.

(A) Wild type (WT) and $\Delta F508$ NBD1x and Rx constructs were transiently transfected into HEK293 cells and monitored by Western blot analysis. The average and standard error of the mean for NBD1x, WT (N=8) and $\Delta F508$ (N=21), and Rx, WT (N=8) and $\Delta F508$ (N=17), are from

6 biologic replicates. (B) WT and $\Delta F508$ TMD2x and full-length constructs were transiently transfected into HEK293 cells and monitored by Western blot analysis. Steady-state levels of Bands B*, C*, B, and C were quantified and normalized to total image TMD2x and full-length respectively. The average is plotted with the standard error of the mean (N=3). (C) Depiction of a folding model with $\Delta F508$ in the presence of different multidomain intermediates. Red arrows represent possible interdomain disruptions in the TMD2x and full-length constructs.

Fig 10. Model of TMD1 impact on biogenic intermediates.

Schematic depicting the CFTR biogenic intermediates as they would occur during translation. Below each biogenic intermediate, an X denotes the initial intermediate in which the mutant has a significant and large impact on steady-state levels.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Figure 8

Figure 9

Figure 10