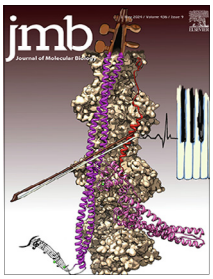
Research Article



The Folding Pathway of ABC Transporter CFTR: Effective and Robust

Peter van der Sluijs [1]*, Hanneke Hoelen [1,3], Andre Schmidt [1,2] and Ineke Braakman [1]

*1 - **Cellular Protein Chemistry,** Bijvoet Centre for Biomolecular Research, Utrecht University, 3584 CH Utrecht, the Netherlands*
*2 - **3D-Pharmxchange,** Tilburg, the Netherlands*
*3 - **Present address: GenDx,** Utrecht, The Netherlands*

Correspondence to Peter van der Sluijs:*Cellular Protein Chemistry, Bijvoet Centre for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, the Netherlands. *p.vandersluijs1@uu.nl (P. van der Sluijs)*
https://doi.org/10.1016/j.jmb.2024.168591
Edited by J. Buchner

Abstract

De novo protein folding into a native three-dimensional structure is indispensable for biological function, is instructed by its amino acid sequence, and occurs along a vectorial trajectory. The human proteome contains thousands of membrane-spanning proteins, whose biosynthesis begins on endoplasmic reticulum-associated ribosomes. Nearly half of all membrane proteins traverse the membrane more than once, including therapeutically important protein families such as solute carriers, G-protein-coupled receptors, and ABC transporters. These mediate a variety of functions like signal transduction and solute transport and are often of vital importance for cell function and tissue homeostasis. Missense mutations in multispan membrane proteins can lead to misfolding and cause disease; an example is the ABC transporter Cystic Fibrosis Transmembrane Conductance Regulator (CFTR). Even though our understanding of multispan membrane-protein folding still is rather rudimental, the cumulative knowledge of 20 years of basic research on CFTR folding has led to development of drugs that modulate the misfolded protein. This has provided the prospect of a life without CF to the vast majority of patients. In this review we describe our understanding of the folding pathway of CFTR in cells, which is modular and tolerates many defects, making it effective and robust. We address how modulator drugs affect folding and function of CFTR, and distinguish protein stability from its folding process. Since the domain architecture of (mammalian) ABC transporters are highly conserved, we anticipate that the insights we discuss here for folding of CFTR may lay the groundwork for understanding the general rules of ABC-transporter folding.

Introduction

CFTR is 1480-amino-acid anion channel expressed on the apical surface of epithelial cells and some other specialized cells.[1–3] CFTR is a member of the ABC-transporter superfamily, and consists of two TMDs (transmembrane domains) and two NBDs (nucleotide-binding domains).[4] In CFTR these domains are part of a single polypeptide chain, with a TMD1-NBD1-TMD2-NBD2 topology, and the first and second TMD-NBD pairs linked by the unstructured cytoplasmic regulatory (R) region (Figure 1). The TMDs associate to form the substrate-recognition site and provide the conduit for substrate transport. Limited sequence conservation in the TMDs likely reflects the diversity of substrates that can be transported by the different ABC transporters. In contrast, NBDs are the most conserved domains, containing the hallmarks of ABC transporters: the ATP-binding site at their

Please cite tts article as: P. van der Sluijs, H. Hoelen, A. Schmidt, et al., The Folding Pathway of ABC Transporter CFTR: Effective and RobustJournal of Molecular Biology, https://doi.org/10.1016/j.jmb.2024.168591

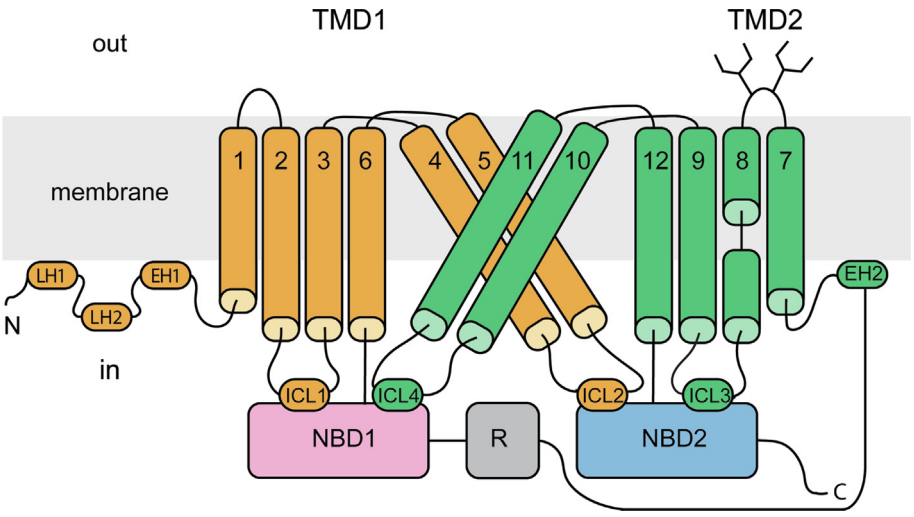


Figure 1. Domain architecture of CFTR. Characteristic tandem organization of CFTR where the two halves each consisting of a transmembrane domain (TMD) with 6 transmembrane helices (TMs) and a nuclear binding domain (NBD). The two halves are joined via the long unstructured regulatory region (R) connecting NBD1 with TMD2. Intracellular loops (ICLs) between TMs tether the TMDs to the NBDs. The cytoplasmic region preceding the first TM of TMD1 contains two lasso helices (LH) and is important for CFTR biogenesis. Both TMD1 and TMD2 are preceded by a short amphipathic elbow helix (EH1 and EH2). TMD1 in orange/wheat, NBD1 in pink, R in grey, TMD2 in green, and NBD2 in blue.

interface, allowing NBD dimerization in the presence of ATP. The TMDs of CFTR each consist of 6 transmembrane helices (TMs). Consecutive TMs are connected via short ectodomain sequences and long cytoplasmic loops. ATP binding and hydrolysis induce conformational changes in the NBDs, which are relayed to the TMDs to drive substrate transport. In contrast to most other ABC transporters, as CFTR is a channel rather than a transporter, these conformational changes open and close the channel pore in the TMDs.

Cystic Fibrosis

Autosomal recessive mutations in the CFTR gene are the root cause of Cystic Fibrosis (CF), a disease characterized by insufficient amounts of functional CFTR at the surface of epithelial cells of airways, intestine, and other organs. The reduced Cl^- permeation and coupled reduction in Na^+ transport leads to a dehydrated mucus. This mucus cannot be cleared by the epithelium's ciliary movements in for example the lungs, ultimately leading to obstruction of the respiratory tract and colonization by microorganisms.[5,6] The highest incidence of CF is seen in Northern European countries with 1/3,500–1/5,000 live births.[6] In the United States, the disease occurs in roughly 1 in 3,000 Caucasian Americans. CF is less common in other ethnic groups, affecting about 1 in 15,000 African Americans and 1 in 35,000 Asian Americans.[7] More than 4,000 genetic variants of CFTR have been described (http://www.genet.sickkids.on.ca/Home.html),[8] with over 700 demonstrated to cause CF disease (https://cftr2.org).[9] The most common CF-disease-causing variant is F508del, accounting for 90% of alleles sequenced in the largest worldwide patient database (https://cftr2.org).

Classification of CFTR Mutations

In pioneering work, Brown and Goldstein developed a classification system for Low Density Lipoprotein Receptor mutants[10] based on phenotypic outcome for the receptor, which later was adopted to classify CFTR patient mutations in groups with similar molecular defects.[1,11] Class-1 mutations encompass variants of CFTR with severely reduced or even absent protein production, such as premature stop codons. Class-2 mutations are characterized by impaired protein biogenesis at the endoplasmic reticulum (ER), with misfolding and degradation, severely reducing the cell-surface population of CFTR molecules. Class-3 mutations affect gating of the channel pore leading to reduced open probability. Class-4 mutations affect the ion conductance path through the pore, leading to reduced conductance. Class-5 mutations, typically in noncoding regions, decrease CFTR abundance, and class-6 mutations destabilize CFTR in post-ER compartments by reducing stability or introducing endocytic internalization signals, thus reducing CFTR levels on the cell surface. Binning CFTR mutations in classes with similar phenotypes might help to identify which mutations respond to specific CFTR modulators (theratyping).[12] A complicating factor for this classification arises from the vectorial principle of protein biosynthesis, which means that many variants have multiple defects: a defect in an early event in CFTR's life

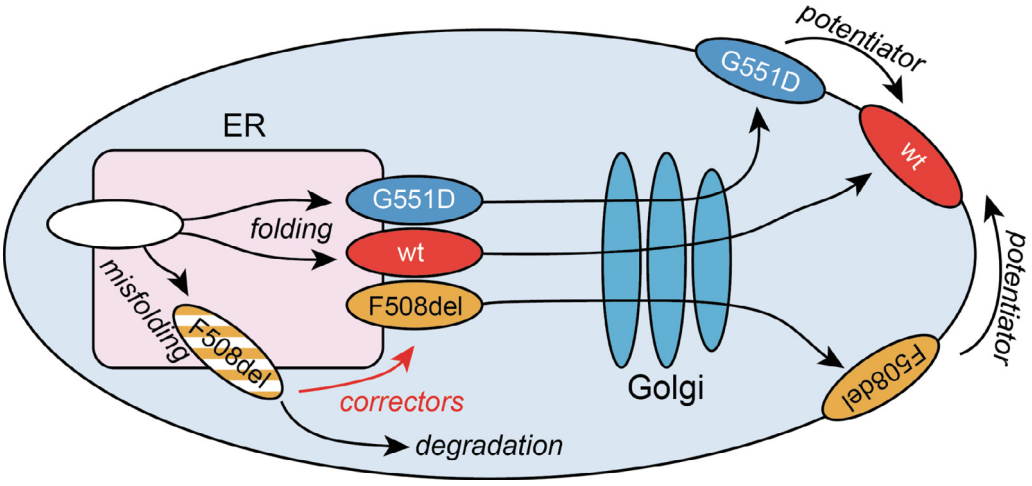


Figure 2. Strategies to rescue CFTR missense variants with impaired function. CFTR is made on ribosomes attached to ER, and translocated into the ER where it starts to fold its domains co-translationally. Folding is completed post-translationally during traversal of the biosynthetic pathway and deposition on the plasma membrane. Many missense mutations in CFTR can cause CF. Class 3 mutants such as G551D are impaired in channel gating but are not recognized by ER quality control; they reach the plasma membrane like wild-type CFTR. These mutants can be rescued functionally by the potentiator VX-770, which led to the concept of a 2-staged therapeutic approach for class-2 missense mutants like the most abundant F508del mutation. Since these are recognized by ER quality-control mechanisms and targeted for degradation, the concept emerged to first improve the class-2-mutant protein with a corrector drug allowing it to escape the ER for the cell surface. If the resulting, rescued conformation is insufficiently functional, this then would be improved with the potentiator. The result is ETI: the current combination therapy with two correctors VX-661 and VX-445 and potentiator VX-770.

in the biogenesis pathway often instigates a downstream molecular defect. An example of this is the F508del mutation, which principally causes misfolding and therefore qualifies as a class-2 mutant. Yet, the mutation also impairs downstream processes: even when a cohort of mutant molecules reaches the cell surface, F508del causes reduced channel conductance (class 4) and increased cell-surface turnover (class 6).[13]

Small-molecule Modulators

The discovery that mutations in CFTR are the basis for CF paved the way for molecular studies aimed to treat and perhaps at one time cure the disease.[14] Well-characterized cellular expression models with CFTR mutants provided deep insights in CFTR proteostasis and the rules for delivery of (mutant) CFTR to and retrieval from the cell surface.[13,15] Results of many studies collectively facilitated development of high-throughput phenotypic in-cellulo drug discovery screens. The first clinical-development success was VX-770/Ivacaftor in 2012: this first-in-class small-molecule CFTR modulator became available to CF patients with the G551D mutation.[16] G551D is a class-3 variant, and thus has normal cell-surface expression but with a gating defect. VX-770 modulates CFTR as so-called potentiator: upon binding, the open probability of the channel pore is increased by destabilizing the rigid conformation of G551D. Unfortunately, VX-770 is unable to rescue variants stranded in the ER and the concept emerged that combinations of 2 or more modulators may be needed to counteract different defects. For F508del, a small-molecule corrector may rescue the protein from ER to plasma membrane, while a second compound could potentiate plasma-membrane localized CFTR[17,18] (Figure 2). The first clinically approved corrector was VX809; it improves misfolded F508del CFTR towards a state competent for export from ER to the Golgi complex and beyond. VX-770 antagonized the activity of corrector VX-809 in cultured cells,[19,20] but in patients the combination of corrector and potentiator fortunately leads to reduced disease burden. In the next iteration, VX-809 was substituted for VX-661, which showed fewer side effects and drug interactions.[21] The addition of VX-445, a second type of corrector acting differently on CFTR, led to development of ETI (elexacaftor-tezacaftor-ivacaftor) a triple combination of two correctors and a potentiator, which proved a turning point in CF therapy.[22,23]

Structural Organization

A large body of biochemical and biophysical experiments accumulated a wealth of information how the domains are functionally organized within CFTR.[24–27] Also, computational models based on structures of other ABC transporters contributed enormously to our understanding of CFTR structure

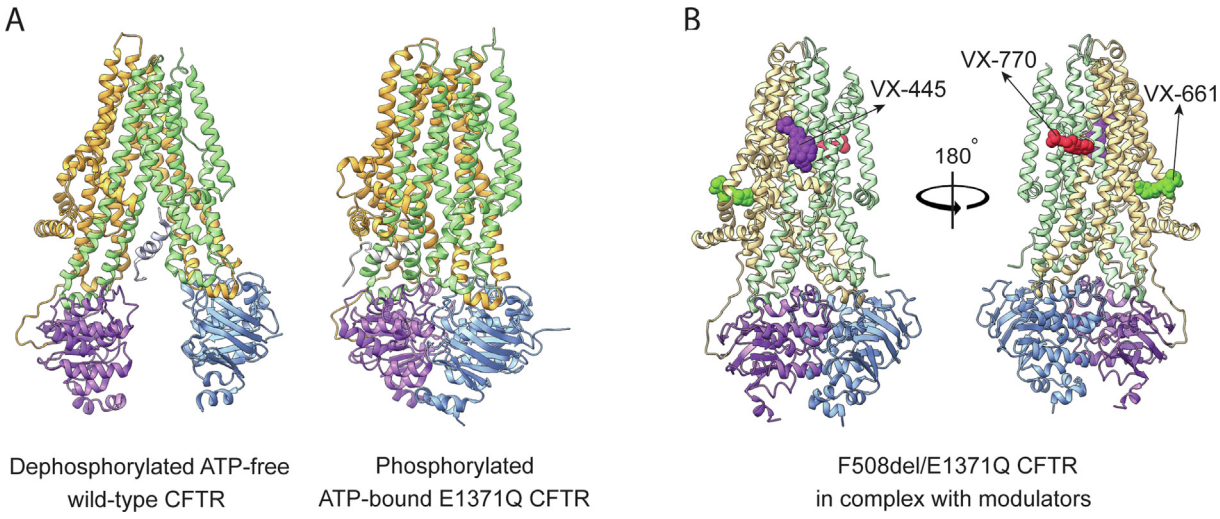


Figure 3. Cryo-EM structures of CFTR. (A) Cryo-EM structures of dephosphorylated ATP-free CFTR (PDB:5UAK),[33] phosphorylated ATP-bound E1371Q CFTR (PDB:6MSM)[34] and (B) F508del/E1371Q CFTR in complex with modulators (PDB:8EIQ).[76] Complexed VX-661, VX-445 and VX-770 are indicated in green, purple, red, respectively.

in the pre-cryo-EM but also the cryo-EM era.[28–30] These earlier insights were deepened to the atomic level by the medium–high resolution cryo-EM structures of zebrafish[31,32] and human CFTR[33,34] (Figure 3A) by the lab of Jue Chen, and by chicken CFTR thermostabilized by mutations[35] in apo (without nucleotide bound) and phosphorylated ATP-bound states. Although human and zebrafish CFTR have a limited amino acid sequence identity of 55%, the structures are nevertheless very similar and the principal features like the architecture of the chloride conductance pathway, separation of NBDs by R, and the structural explanation of patient mutants are the same. The cryo-EM structure of thermally stabilized chicken CFTR is different in several respects. Firstly, the kink found in TM8 of the zebrafish and human structures, which displaces TM7 from the position it has in other ABCC transporters, is not present in the chicken structure. Secondly, the distance between the NBDs is in between that of the open and closed forms of the other structures. Thirdly, R not only protrudes into TM9, TM10, and TM12 but also intertwines with TM3 and TM4. Fourth, as chickens have a body temperature of 39–43 °C, chicken NBD1 is much more stable, which is caused by inclusion of 4 proline residues and I539T, which is an F508del-suppressor mutation.[36,37]

In the full protein, the two TMD-NBD halves are connected in a head-to-tail manner via R that, with a length of ~200 amino acids, is a unique and atypically long linker for ABC transporters.[38] Phosphorylation of R by PKA is required for CFTR channel function, underscoring that R does more than just physically connecting NBD1 and TMD2. The descending and ascending limbs of the intracellular loops (ICLs) linking the TMs largely are not loops but α-helical extensions of the TMs; they feature a short coupling helix (CH) connecting the helices at the tip of the loops. Interactions between the TMDs and NBDs are brought about in cis by CH1 of TMD1 with NBD1 and by CH3 of TMD2 with NBD2. Through a domain swap of α-helices in the TMDs, TM4 and TM5 of TMD1 reach over in trans to interact via CH2 with NBD2, and TM10 and TM11 in TMD2 do so via CH4 with NBD1 (Figure 1). Thus, CH1 and CH4 interdomain joints connect TMD1 and TMD2 with NBD2, and CH2 and CH3 in TMD1 and TMD2, respectively, with NBD1. These interdomain associations are instrumental in transmitting the conformational changes brought about by ATP binding and hydrolysis in the cytoplasmic NBDs, to motions of TMs in the membrane that are needed for channel opening and closing. Contacts between NBDs and coupling helices of ICLs that bind in cis (mediated by CH1 and CH3) are likely weaker than interactions in trans (mediated by CH2 and CH4), because the latter are inserted deeper into the NBDs with a larger contact surface.

The structure of apo-hCFTR resembles that of an ABC transporter in inward-facing conformation, in which the R region (including a C-terminal helix in CFTR that likely corresponds with residues 825–843) is wedged between TM9, TM10, and TM12 of TMD2 and extends into the cytoplasm between the NBDs, thereby physically precluding an interaction between NBD1 and NBD2. The electron density of R however is too low to assign secondary structure. In the phosphorylated ATP-bound state of hCFTR, the two halves of the protein have moved toward each other, and the NBDs form a closed dimer. R becomes completely unstructured and aligns along the surface of NBD1 suggesting interactions with

NBD1 and making it likely that R swayed away from the NBD interface allowing structural changes in the TMDs underlying the opening of the channel. R is a typical intrinsically disordered region as it can bind both NBDs of CFTR, its N-terminus, C-terminus, an ICL of TMD2, and other proteins.[39–45]

Expression of two CFTR halves without R or of a single construct where CFTR is cut into two halves, between R and elbow helix 2 of TMD2 still yields a channel with some functionality even without phosphorylation, suggesting that apo R is inhibiting opening of the channel.[46,47] A complicating factor in the analysis of expression experiments with truncations is the instability of TMD2[48,49] when expressed alone without TMD1 and NBD1. TMD2 stability may be affected by the absence of preceding sequence jeopardizing proper insertion of TM7 and therefore TMD2 into the ER membrane, even though TM7 is a highly effective signal peptide (see below). Thus, the role of R for CFTR maturation remains to be established. R contains over 10 serine/threonine PKA phosphorylation sites. The phosphorylation status of a single amino acid residue can already be subject to many cellular variables and expression experiments with targeted phosphomimicking mutants also have their limitations because such mutants cannot recapitulate the kinetic aspect of phosphorylation. Moreover, changing serine to glutamate or threonine to aspartate or alanine residues are at most, approximations of phosphorylated serine/threonine or the non-phosphorylated status. Accordingly, it has been difficult to correlate folding and channel activity in cells with precise knowledge of phosphorylation in terms of stoichiometry and topology, in particular which residues are phosphorylated when. For that matter the intuitively straightforward scenario where phosphorylation removes R from its occluding inhibitory position between the NBDs to make the difference between active and inactive conformation, may need fine-tuning.[50]

The Elbow and Lasso Helices

The new structures also revealed a previously unrecognized feature in the N-terminal cytoplasmic region preceding TM1 called the lasso, which embraces and interacts with TM2 and TM6 of TMD1 and TM10 and TM11 of TMD2 in CFTR. The lasso is unique to members of the ABCC subfamily of ABC transporters and has been found in structures of CFTR, multidrug resistance protein 1 (MRP1)[51,52] and sulfonylurea receptor 1 (SUR1).[53,54] Two short helical regions in CFTR proximal to the elbow helix of TMD1 are partially inserted in the membrane where corresponding sequence in MRP1 and SUR1 connects TMD1 with the third TMD present in most other ABCC proteins, the penta-helical TMD0 at the N-terminus of TMD1.[55]

The lasso is important in SUR1 for intermolecular interactions with pancreatic Kir2.6 K+ channel.[54] In MRP1 it serves a role in folding[56,57] and in CFTR the lasso binds to syntaxins[58–63] and actin cytoskeleton.[64,65] The mechanistic role of these interactions in CFTR function is not well understood, and for CFTR this part of the protein is critical for function as indicated by the presence of CF-causing mutations found in the lasso.[66,67] A highly conserved negatively charged stretch with acidic side chains within lasso helix 2 (residues 46–61), interacts with N-proximal residues in the R region, and is possibly involved in CFTR gating.[39,68] Interactions between the lasso and R are independent of R phosphorylation and ATP binding[68] and are not visible in the cryo EM structure of apo CFTR.[34] Phosphorylation of R however does increase interactions of the lasso with the C-terminus.[69] An additional structural element in the N-terminal part of TMD1 is the elbow helix which directly precedes TM1 like TM7 does in TMD2, a short helix that sits on or is partially inserted in the cytoplasmic leaflet of the plasma membrane and presumably stabilizes type-IV ABC transporters.[4]

In all cryo-EM structures, whether CFTR is phosphorylated and ATP bound or not, the ion conductance path appears too narrow to facilitate gating of chloride ions. This may reflect differences in stability of the various CFTR conformations, with the closed channel being most amenable to purification and imaging. Alternatively, the use of detergents in purification protocols may affect binding of nucleotide to the NBDs or CFTR TMD conformation, and thereby may bias certain conformations.

Corrector and Potentiator Binding Sites on CFTR

The standard of care for over 175 CF mutations (https://www.cff.org/sites/default/files/2022-02/Trikafta-Approved-Mutations.pdf) is the EMA and FDA-approved Trikafta formulation consisting of ETI: potentiator VX-770 (Ivacaftor) in combination with type-I corrector VX-661 (Elexacaftor) and the newer corrector VX-445 (Tezacaftor), which turned out to have potentiator activity as well.[70,71] Although ETI fails to improve NBD1 folding specifically in the most frequently occurring F508del mutant, it does boost folding of the entire protein (domain assembly) and thereby rescues mutant CFTR from degradation[49,72] The corrector compounds increase the number of CFTR molecules on the cell surface and thereby increase chloride and bicarbonate permeation through the plasma membrane due to their effects on CFTR folding and stability.[16,66,72,73] An overview of compounds approved for clinical use is displayed in Table 1.

Earlier work with CFTR truncations revealed that VX-809 targeted TMD1 and NBD1.[74,75] Cryo-EM structures of CFTR-F508del in complex with the

Table 1 Compounds approved for treatment of Cystic Fibrosis.

FDA/EMA approved	Compounds	Developed for	References
Trikafta	Elexacaftor Tezacaftor Ivacaftor	F508del	23,179
Symdeko	Tezacaftor Ivacaftor	F508del	180
Orkambi	Lumacaftor Ivacaftor	F508del	181
Ivacaftor	Ivacaftor	G551D	16,182

ETI modulators have been solved, which in parallel with molecular docking studies cemented the independent binding sites of VX-661 and VX-445 and potentiator VX-770 on CFTR.[76] The type-I correctors VX-661 and VX-809 insert the central 1,3-benzodioxol-5-yl-cyclopropane carboxamide moiety in a hydrophobic pocket in TMD1, while their variable parts are in contact with lipids. Baatallah and Hudson using computational methods independently discovered a potential, additional binding site of VX-809 in NBD1.[77,78] The type-I correctors are thought to stabilize a weak spot consisting of lasso helix 2, elbow helix 1, TM1, and TM6, which is home to several class-2 mutations, and allosterically couples to NBD1, ICL1, and ICL4.[78]

Cryo-EM images uncovered that corrector VX-445 mainly interacts with TM11 forming a pocket with residues in TM2, TM10, and a proximal part of the lasso (Figure 3B). Molecular docking and molecular dynamics simulations in combination with secondary-site mutational analysis in F508del also revealed a potential binding site for VX-445 in the lasso and NBD1. The latter of this corresponds with the earlier predicted VX-809/VX-661 binding pocket[78] (Figure 3B). The combination of VX-445 and VX-661 synergistically enhances conformation and ER-to-Golgi transport than VX-809 alone and functionally acts additively.[76,79] Whether VX-445 binding to TMD2 or NBD1 can explain its potentiator activity remains to be investigated.

The dual function of VX-445 suggests that a binary classification of CFTR modulators as potentiators and correctors may be too simple especially since drug hunting expeditions may yield new therapeutic compounds with unanticipated modes of action on CFTR. After all, the root cause of impaired functionality is a structure-based defect, whether recognized by the cellular quality-control system or not, and whether due to ER retention and degradation or channel dysfunction. Of note, all CFTR – corrector structures have been obtained on a protein that is already folded, while the correctors in the cell are acting on the vectorial trajectory of CFTR that is in the process of folding but has not reached its final functional form. This in contrast to potentiators that act acutely on folded, domain-assembled CFTR. For a precise understanding, additional structures may be needed to fully appreciate the mode of action of corrector compounds.

Cryo-EM structures of phosphorylated ATP-bound E1371K CFTR[80] and CFTR-F508del/E1371Q,[76] and in silico molecular docking studies[81] revealed a binding site of VX-770 (and GLPG1837) in a pocket lined by TM4, TM5, and TM8 which concurs with the kink region in TM8, a unique feature in CFTR of many species (Figure 3B). The break in TM8 displaces TM7 from the position found in other ABCC proteins and VX-770 is thought to stabilize the open conformation.[16,82] Another binding site of VX-770 elsewhere in CFTR has been reported using a photoactivatable version of VX-770[83] and hydrogen/deuterium exchange – mass spectrometry of thermostabilized CFTR.[84]

Folding Pathway of CFTR

For any protein to be functional, it requires a proper three-dimensional structure. This always is more than a single structure, as a protein is metastable and needs to adopt a range of conformations to perform its function. The CFTR channel opens and closes, regulated by phosphorylation and ATP binding and hydrolysis (see above), each phase affecting subdomain structures and domain interactions in both the membrane and cytosol. Because of their inherent dynamic nature, proteins at best are as stable and rigid as they need to be but not more. This is illustrated by the finding that proteins, in all species examined, are expressed at a level just below their solubility: 'life on the edge of solubility'.[85,86] The challenge for cells is to fold each protein and keep it stable for as long as is needed. These two processes have different requirements, as discussed for individual domains below.

CFTR is amongst the most-studied proteins passing through the secretory pathway, and is often considered to be unstable. Early work showed that appreciable amounts of overexpressed wild-type CFTR were degraded from the ER by the proteasome[87,88]. Later studies reported a >3-times higher percentage of CFTR that received the complex-glycosylated form, to ~80% in 2 hours,[36,72,89] a common number for other proteins in kinetic studies,[90] such as the rapidly folding influenza virus hemagglutinin.[91] Explanations for

the low folding efficiencies in early work are the activation of proteasomal and authophagic degradation by cell stress induced by high expression levels in CHO or BHK cells, and lengthy amino-acid depletion for radioactive labelling.[92] Moreover, a wealth of steady-state Western-blot data on wild-type CFTR shows the fraction of complex glycosylated wild-type CFTR at >80%. The 20% of protein in the ER in steady-state represents the fraction of protein that is still folding in the ER, again similar to other proteins using the secretory pathway.[90] CFTR folding is as robust and effective as that of less complex proteins.

CFTR starts folding during translation, which is far from trivial for a polytopic transmembrane protein that pairs large domains in cytosol and ER membrane with small ER-lumenal loops. Fortunately, translation rate of CFTR is relatively low, at 2.7 residues per second, with an average time of over 9 min to translate the 1,480 amino acids.[87] This allows for initial vectorial folding, from N- to C-terminus, domain by domain, because secondary and tertiary structure form spontaneously at sub-millisecond timescales.[93]

While average translation of CFTR is slow, local translation rate was shown to affect the outcome of folding and function.[94,95] The genetic code contains information regarding protein translation rate and provides an additional mode of regulation for protein synthesis. Missense mutations including synonymous and non-synonymous SNPs can affect translation rate as well.[95] The translation rate can be modified locally by tRNA abundance, mRNA secondary structures, or specific peptide sequences that cause translation stalling.[96] The slowing down or pausing of translation at distinct sites creates a time window for protein maturation, such as domain folding, domain assembly, and chaperone interactions.[97] In CFTR, slowing translation rate altered conformation and function of wild-type protein,[98] and enhanced expression and function of several CF-related mutants including F508del and G85E.[94] Yet, changing translation kinetics is equally likely to negatively impact the fate of newly synthesized proteins.[96,99–101] Abundant codon optimization in the first nucleotide-binding domain (NBD1) of CFTR resulted in the aggregation of the full-length protein.[102] Together these results show that the relationship between translation rate and folding is complex one, which requires further study.

TMD1

The first CFTR domain to be translated is the N-terminal TMD1; the 6 TM helices are co-translationally translocated into the ER membrane. Despite major technological advances, much was already known a decade ago.[103] After synthesis of the long cytosolic N-terminus of 80 residues, the first 2 TM helices each can act as signal peptide for targeting to the ER membrane.[104–106] TM1 or TM2 binds SRP for targeting of the complex of CFTR nascent chain with ribosome and mRNA to a translocon in the ER membrane. Although TM1 is the first TM to emerge from the ribosome, it promotes translocation of only 25% of nascent chains because of its two charged residues, Glu92 and Lys95. The remaining 75% are translocated by TM2, the optimal, lead signal peptide in CFTR.[104,105] While this occurs during the continued translation of CFTR and formally is considered co-translational, TM1 is translocated later than its translation would allow. TM1 hence in effect is translocated via a post-translational TM2-mediated pathway. TM1 and TM2 then are inserted loop-wise into the ER membrane as soon as both are present, with the positively charged residues upstream of TM1 and downstream of TM2 ensuring translocation of the soluble loop between these TMs into the ER lumen (the positive-inside rule).[107–109] This forms ECL1 (extracellular loop 1), positions TM1 and TM2 in the ER membrane, and promotes insertion of the TM3-TM4 and TM5-TM6 loops with proper topology.[106] TM3-TM4 and TM5-TM6 are thought to be inserted by a similar mechanism, especially since ECL2 and ECL3 are short loops between TM3-TM4 and TM5-TM6, respectively.[103] Mechanistic studies have used a helical hairpin pair formed by TM3 and TM4 alone, with TM3 as inefficient signal peptide, and without further physiological TMD1 context.[103,109–111] Similarly, the TM5-TM6 pair alone is translocated inefficiently into the ER membrane from mammalian microsomes or yeast, while the highly charged, isolated TM6 tends to be fully translocated rather than be inserted in the ER membrane.[112] Yet, the Von-Heijne lab showed that TM1-TM6 all insert well into ER membranes, having sufficient length and hydrophobicity, and without requiring flanking sequence, in contrast to 9 out of the 12 TMs of ABC transporter Pg-P and 2 TMs in CFTR TMD2.[113] Whereas CFTR often is considered to be much less stable than other ABC-transporters, this comparison showed the contrast: CFTR prevailed.[113] All published data are consistent with a co-translational insertion of TM1 through TM6 leading to co-translational folding of TMD1 into a stable independent domain.[72,114] Whether the TM3-TM4 and TM5-TM6 pairs are inserted immediately upon their synthesis remains to be established.

In functional, native CFTR, ICL1 and ICL2 of TMD1 associate with NBD1 and NBD2, respectively (Figure 1). Yet, TMD1 folds independent of the NBDs, whether expressed on its own or attached to other domains up to newly synthesized full-length CFTR.[114,115,72] Isolated TMD1 is not insensitive to intracellular degradation,[116,117] and proper insertion into the membrane is key. The cytosolic N- and C-termini of TMD1 assemble with ICL1 into a relatively protease-resistant, compact conformation that has protease

accessibility only in the cytosolic N-terminus and in ICL2.[72,115] This early association of cytosolic helices and TMs keeps the loop-wise-inserted TMs and ICLs in register and in the membrane, stabilizing TMD1 already during the co-translational insertion into the ER membrane.

TMD1 folds into a stable domain independent of other CFTR domains, as shown by its stability in cells when expressed in isolation[72] and its similar conformation whether attached to other CFTR domains or not.[72,114] Despite this independence, ICL1 of TMD1 associates with the N-terminal subdomain of the next domain, NBD1, as soon as this subdomain is synthesized. The contact is independent of the rest of NBD1 and of the presence of F508,[72] and is likely native. Another contact of TMD1 with NBD1 was found in 3D structure models[118] and peptide-interaction assays.[119] Peptides including the first 20 residues and the last 20 (the elbow helix) of the cytosolic N-terminus of TMD1 bind purified NBD1.[119] In the model, the TMD1 N-terminus folds around NBD1. Indeed, the TMD1 N-terminus is more protected from protease when the N-terminal subdomain of NBD1 is tethered to the C-terminus of TMD1.[72] This conformation solves the conundrum of the location of the (in parts amphipathic) N-terminus between its translation and the time it can embrace its native interaction partners in mature CFTR, which include TMD2.[72]

NBD1

NBD1 folds during synthesis as well and as independently as TMD1.[36,114,120] The individual domain folds well in eukaryotic and prokaryotic cells[121,122,36] and has complex topology that makes it difficult to refold *in vitro*.[121] The N-terminal subdomain that binds ICL1 does not need the interaction with ICL1 and folds as soon as it emerges from the ribosomal tunnel, stabilized by ATP binding.[120] Next is the synthesis of the α-helical subdomain in NBD1, the domain that contains the F508 residue missing in most CF patients. Conformational freedom of the α-helical subdomain is limited by the association of the N-terminal NBD1 subdomain to ICL1, which may support the folding process. NBD1 then is completed by synthesis of the ß-sheet core, which integrates with the already compact N-terminal subdomain to form the α/β-core subdomain[120,123] and bind the γ-phosphate of its ATP molecule and the Mg ion, which may support N-C assembly and NBD1 stability. Limited proteolysis indeed yields a rather stable protease-protected NBD1 fragment of ~25 kDa,[114,124,125,36] which results from cleavage in the N- and C-terminally localized intrinsically disordered regions RI (regulatory insertion) and RE (regulatory extension) in NBD1,[115] implying that both α-helical subdomain and α/β-core subdomain are rather compactly folded. NBD1 is stabilized by an ATP molecule that is not hydrolyzed during the functional cycle,[121,126,127] but was recently shown to be important for regulating CFTR activity as chloride channel.[128] The ATP-bound N-terminal subdomain needs some to become competent for integration with the C-terminal ß-sheet core. This time is provided by the relatively slow translation of the α-helical subdomain, as accelerating its translation by ~20 synonymous nucleotide substitutions increases aggregation of NBD1.[102]

R region

The C-terminus of NBD1 is called RE. It contains the protease-accessible cleavage site in CFTR that generates the C-terminal boundary of ~25-kDa NBD1-fragment N1a[115] and RE forms a debatable boundary with the R region, which is translated next. In structures of purified NBD1, RE associates with NBD1[41] but in full-length CFTR structures this is not obvious. R is intrinsically disordered, but should be called intrinsically polyamorous instead, as it was shown to associate with (and lend conformation from) N- and C-termini of CFTR, with NBD1, NBD2, TMD2,[39–45,129] and the STAS domain of SLC 26 proteins,[130] all regulated by the phosphorylation status of R during its functional cycle. CFTR in the ER is phosphorylated on R already (our unpublished observations); whether this plays a role in CFTR folding remains to be determined.

TMD2

The next domain to be translated is TMD2. It resembles TMD1 in many respects.[103] Similar net hydrophobicity, similar number of charges, both are co-translationally inserted into the ER membrane, both can be expressed independent of other CFTR domains. Yet, TMD2 contains imperfect TMs that make it more aggregation prone and less cell-biologically stable than TMD1. TM7 is a fairly perfect TM helix and signal peptide that targets TMD2 to the ER membrane, in TMD2-only constructs, and to guarantee reinsertion of TMD2 after release of the ribosome from the ER-membrane-localized translocon during synthesis of NBD1 and R. TM8, while important for function, is less ideal: it is kinked in the membrane and forms together with TM7 and part of R the most aggregation-prone area of CFTR.[129] TM8 and TM12 were identified as the only two TMs in CFTR that were incapable of inserting into the ER membrane on their own.[113] TM7 likely drags TM8 along in a loop-wise insertion of the two TMs, likely stabilized by glycosylation of ECL4 between TM7 and TM8, the only glycosylated extracellular loop in CFTR. We recently identified a compact, protease-resistant, aggregation-prone structure including R and TM7-TM8. TM9-TM12 are likely inserted by the translocon in similar manner as TM3-TM6, perhaps as 2 looped pairs with start- transfer and stop-transfer sequences flanking the TMs. TM9, TM10, and TM11 insert well as individual TMs, whereas TM12 requires TM9-TM11 or NBD2 attached.[113] The importance of structures

both C-terminal of TMD2 (NBD2) and N-terminal (the aggregation-prone fragment including TM7-TM8 and part of R) resembles the packing of the TMD1 N- and C-termini onto ICL1. Indeed, TMD1 and TMD2 inserted into the ER membrane alone or in context of full-length CFTR are relatively protease resistant already immediately after synthesis: the only protease accessibility in TMD1 is in ICL2, and in TMD2 in ICL3, with ICL1 and ICL4 protected.[72,115] This suggests that TMD1 and TMD2 go through similar insertion, folding, and compacting processes.

NBD2

NBD2 is the last domain translated; it completes CFTR. This domain has similar structure as NBD1, with an N-terminal subdomain, an α-helical subdomain, and a C-terminal ß-sheet core that integrates its ß-strands with those of the N-terminal subdomain. Its stability is much lower though, requiring stabilizing mutations for in-vitro experiments,[131] and lacking a stable, protease-resistant fold in absence of assembly with other CFTR domains.[49,115116,125,132] The ATP molecule bound to NBD2 does less to stabilize NBD2 but is hydrolyzed when NBD1 and NBD2 dimerize upon displacement of R from the NBD-dimer interface by its phosphorylation. ATP binding to NBD2 and hydrolysis regulate the gating of the channel. Striking is the stability and residual functionality of CFTR lacking NBD2 altogether: the truncated protein reaches the cell surface and displays minimal activity; gating is affected as there is no ATP to hydrolyze.[116,133,134]

Post-translational domain assembly

Translation termination after NBD2 closes the co-translational phase of folding. TMD1, NBD1, and TMD2 are folded as if they are expressed in isolation, in rather stable form, NBD2 is not (Figure 4). Because purified NBD2 with solubilizing mutations[131] allowed determination of its structure [PDB:6UK1], NBD2 may well resemble a molten globule at this early folding phase,[135,136] a looser tertiary structure with native secondary structure in place (Figure 4). TMD1 and NBD1 have assembled during synthesis, but the other domains need the post-translational phase to assemble. While NBD1 does not change detectably, TMD1, TMD2, and NBD2 acquire their native protease resistance at the time of export from the ER,[115] suggesting a cooperative process. ICL4 docks onto NBD1, forming a crucial functional interface.[137–140,18,24,25] Although the α-helical subdomain of NBD1 constitutes the co-translationally formed docking site for ICL4 in TMD2, there is no indication

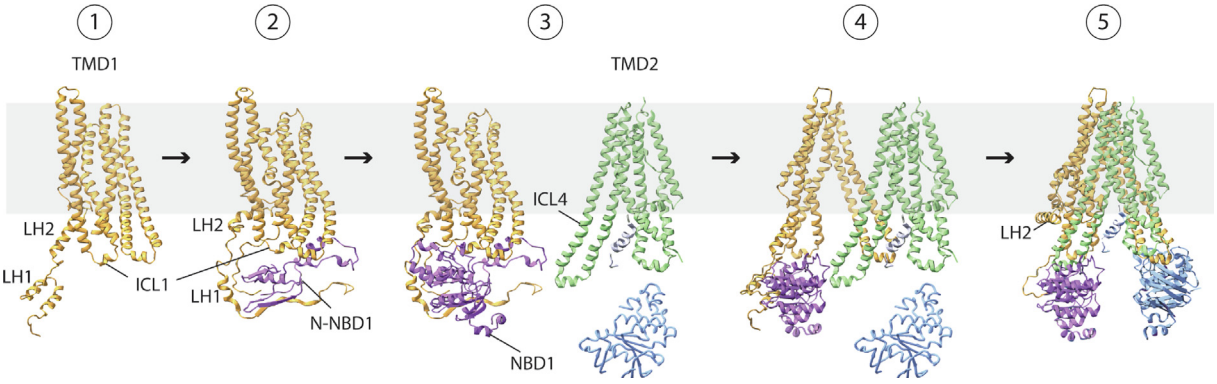


Figure 4. Model of folding pathway CFTR. TMD1 is the most N-terminal domain, translated first and co-translationally inserted into the ER membrane (image 1, based on structural model).[178] Its far N-terminus lacks native interaction partners,[178] whereas the N-terminal LH2 and the C-terminus of TMD1 associate with ICL1 into a stable domain structure.[72] This packing is improved by corrector VX-809 and, through a different mechanism, by expression of the N-terminal subdomain of NBD1, associating with ICL1 and the N-terminus of TMD1 (image 2, based on structural model.[116] The N-terminus changes conformation for interactions including LH1 with NBD1.[116] Translation of NBD1 continues and NBD1 folds co-translationally,[114,115,36] in association with ICL1 and the N-terminus of TMD1[72,118] (image 3, TMD1-NBD1 based on the same structural model as image 2).[116] R (only a C-terminal helix visible in grey) and TMD2 are translated next. TMD2 is inserted into the ER membrane and folds during translation as well, resulting in a CFTR nascent chain with TMD1, NBD1, TMD2 folded as if expressed alone,[114,115] and only TMD1-NBD1 assembled. NBD2 synthesis completes CFTR but the domain cannot fold well at this stage[113]; it remains protease sensitive (image 3, R-TMD2-NBD2 based on cryo-EM structure PDB:5UAK[32]; NBD2 was modified to indicate its non-native state in images 3 and 4). For NBD2 folding, assembly of TMD1, NBD1, and TMD2 is needed. Proteolytic protection of TMD1 ICL2, TMD2 ICL3, and NBD2 is simultaneous, likely cooperative (image 4, based on PDB:5UAK[32]).[115] Completion of CFTR folding and domain assembly includes the N-terminus of TMD1, which becomes protease resistent by wrapping around TMD1 and TMD2 (image 5, PDB:5UAK).[33] TMD1 in orange/wheat, NBD1 in pink/purple, R in grey, TMD2 in green, and NBD2 in blue.

that they already assemble during synthesis. Instead, all findings point to a post-translational association, probably because CFTR structures and structural models demonstrate that association of both ICL1 and ICL4 with NBD1 requires their assembly to avoid steric hindrance. Next to ICL1-ICL4 assembly, ICL2 and ICL3 assemble in the cytosol, together and with NBD2, and TM intramembrane interactions are rearranged, resulting in the native CFTR structure (Figure 4). The only major change during domain assembly is in the N-terminus of TMD1. Those 80 residues interact with residues translated throughout CFTR synthesis and in part unavailable until domain assembly has been completed. It remains to be determined whether the range of conformations in the structural models and cryo-EM structures (Figure 4) all are populated and if so, whether they are populated transiently and in sequence or in an equilibrium, perhaps related to function.

The complexity of the assembly process is limited, because TMD1, NBD1, and TMD2 do not change conformation but rather become more packed, with proteolytic protection of ICL2 and ICL3 at the same time as their docking onto NBD2.[115] Compared to ABC transporters where TMD1 interacts exclusively with NBD1 and TMD2 exclusively with NBD2, ICL2 and ICL4 need to swap position, with ICL2 binding NBD2 and ICL4 binding NBD1. Yet, the two TMDs in fully assembled, active CFTR, still are recognizable from the "top" of the structure, from the outside of the cell, as 2 assembled halves of 6 TMs each.[33] It therefore should not be too difficult to assemble these 4 already folded domains, despite the dynamic nature of the assembled protein, with constantly moving ICLs and attached NBDs in the functional cycle. Why then does it take an average of 30–60 min.[36,115,141] The delay most likely is caused by the involvement of cytosolic and intramembrane chaperones/translocons [see 103 and below], because all proteins that follow the secretory pathway appear to have a similar half time of exit from the ER/ secretion.[90]

Upon its proper folding, CFTR passes the quality-control system of molecular chaperones and leaves the ER on a path to the cell surface. Most components of this sentence are not clear. What determines which pathway CFTR travels, via the Golgi complex or bypassing it? Do distinct exit pathways exist or do all start with COPII coating of ER membranes? What is required for exit from the ER? Release from chaperones or entry into a COPII-coated bud or both? Which quality-control factors are involved in the triage decision? And what is "quality"?

What constitutes proper folding? Classes 3 and 4 of CFTR missense mutations reach the cell surface but are not functional, so quality does not equate functional quality. For wild-type CFTR we do not know whether it is the closed-channel conformation that leaves the ER –intuitively preferable– or the open-channel conformation, or the active channel, as was reported.[142] The lack of evidence for NBD2 using NBD1 as direct template favors the closed channel as endpoint of folding in the ER. NBD1 folds well unless it cannot bind ATP or contains a missense mutation that prevents its folding, whereas NBD2 folds when it assembles with ICLs in TMD1 and TMD2.[115] While the 2 processes are linked, they can be modulated independently (see below: A trade-off...). Related questions are the location and interactions of R during domain folding and assembly in the ER. Immediately after synthesis, CFTR contains some phosphorylated residues in NBD1 or R[143] but whether these are sufficient to dislodge R from the NBDs and allow ATP hydrolysis and channel opening is not clear.

With many questions remaining, it has become clear that the folding process of CFTR is rather resilient and effective. The translation phase is optimally exploited for the folding of 3 out of the 4 structured domains and a first domain-assembly step. The second phase thereby is limited to the assembly of 4 domains plus some rearranging and the embrace of the N-terminus. Despite the many inherent instabilities in CFTR, its functional structure is amazingly stable, tolerating many openings and closings before its collapse.

Folding of Missense Disease Mutants

Over 4,000 CF-related CFTR variants have been reported[8] (http://www.genet.sickkids.on.ca/Home.html),[8] many of which are missense mutations giving rise to amino-acid substitutions or short internal deletions in all domains, along the length of the entire molecule. Missense variants often are synthesized and translocated into the ER suggesting that co-translational folding is relatively robust and insensitive to this type of mutations. Post-translational folding to a mature conformation and biosynthetic transport from ER to the plasma membrane however may differ enormously between mutants and with wild-type CFTR.[9] From retention in the ER and destruction by ER-related degradative pathways for F508del, to wild-type-like expression of G551D on the plasma membrane. Much effort went into uncovering the mechanisms underlying the molecular defects for a limited number of missense mutations and research particularly focused on the most prevalent mutation F508del. We here discuss what the molecule-driven approaches targeted on the F508del variant have taught us about the pathway of domain assembly of CFTR and what the remaining challenges are that we face to fully correct the defect.

Soon after the discovery of the CFTR gene and its most common disease-causing variant F508del, it became apparent that this absence of a single amino acid leads to a local destabilization of the subdomain,[121,122,144] which in turn leads to overall

misfolding, ER retention and consequent degradation. Soon after, several lines of evidence suggested that the destabilization caused by F508del can be counteracted.[121,145] First, lowering the temperature of cells expressing F508del CFTR leads to a partial restoration of its cell-surface localization[146]; second, intragenic suppressor mutations rescued F508del NBD1 folding,[121] third were chemical chaperones with the same mild rescue effect.[147] The at least partial rescue by these different maneuvers suggested that, energetically, the defect may be relatively mild. Third, studies to identify so-called intragenic second-site suppressors were successful in identifying multiple single amino-acid substitutions that in cis rescue F508del CFTR.[148–150] Continued efforts eventually led to direct evidence that F508del was causing two distinct defects: F508del causes both a (local) destabilization of the mutant NBD1 domain, and F508del prevents domain assembly by disrupting the interface with ICL4 in TMD2. These defects in stability and assembly of F508del-CFTR can be corrected independently. The suppressor mutants demonstrate that improving folding and exit from the ER is no guarantee for improved function and stability. It should be noted that these second-site-suppressor variants in cis with F508del still showed a much-reduced ion conductance as well as decreased cell-surface residency when compared to wild-type CFTR. Thus, while both genetic and exogenous treatments can rescue F508del-CFTR to the membrane, its native structure is still more similar to F508del than wild type.[36]

A Trade-off between Folding Process, Stability, and Function

Channel function of CFTR requires flexibility of the TMDs because their TMs are lining the ion-conductance pathway and must be capable of following the conformational alterations induced by dimerization of the NBDs in the cytoplasm. Residue D924 causes a discontinuity in TM8, which creates 2 helical breaks within the membrane. These liberate main-chain atoms from their secondary structure, thereby accommodating an interaction with Cl^- ions along the ion conductance pathway as seen in ClC chloride channels[151] and a potassium channel.[152] Breaks in transmembrane helices can generate hinges in the protein, which may provide requisite flexibility for gating. In addition, the conductance pore is lined with positively charged residues in TMs. These functional requirements are not necessarily compatible with a stable fold of a membrane protein. Indeed, CFTR easily misfolds in cells and is prone to aggregation, while repeated opening and closing of the channel during the gating cycle contribute to its limited stability. Contrary to expectations, this is not caused by physicochemical properties in TMD1 of which the most C-terminal transmembrane helix TM6 contains three positively charged amino acids that are important for CFTR function as chloride channel, but hamper integration of TM6 into the ER membrane (see above) together with its relatively short length.[113] TM6 has a predicted ΔG of 4.76 kcal/mol for membrane insertion (https://dgpred.cbr.su.se/index.php?p=home), which makes it the predicted most unstable TM of CFTR.[153] Instead, the instability of CFTR can be traced back to TM7, TM8, and the connecting ECL4 in TMD2, and part of R, which form a protease-resistant, folded, aggregation-prone structure of 42 kDa.[129] Net charge and distribution of charged residues of TMD1 and TMD2, and hydrophobicity predictions for TMD1 and TMD2 were similar. Other biophysical properties of TMD2 such as TM length,[113,154] and structure must explain its aggregation propensity. Correct translocation of the first TMs of TMD2, TM7 and TM8, is crucial for proper CFTR folding and insertion.[106] The two N-linked glycans in ECL4 between TM7 and TM8, are added efficiently and likely contribute to the stable insertion of TM7 and TM8. TM7 can function as signal peptide and is the most hydrophobic transmembrane helix of CFTR with a predicted ΔG for insertion –3.36 kcal/mol, ensuring strong anchoring in the bilayer during synthesis. In contrast, TM8 is not stably integrated and needs TM7 for proper translocation and insertion.[106,155] D924 located at the breaks in TM8 has a high propensity to be in a loop and reduces its hydrophobicity.

Clearly, functionality requires not only dynamics but also imperfection in stability such as charged side chains in the TMs inside the membrane. CFTR stability then is a challenge, and is ensured through a wealth of interactions in the TMs and ICLs. Stability however needs to be distinguished from the folding and membrane-insertion processes. Translocons can insert imperfect TMs into the ER membrane, but whether they are retained in the membrane or pass through to the ER lumen initially depends on their interactions with translocons and chaperones, and eventually on interactions within CFTR. TM6, the most difficult TM of CFTR to insert then can become stable part of the better TMD in CFTR. TM7 is the most perfect TM and an effective signal peptide, easy to insert, but then also the worst for CFTR stability, co-responsible for aggregation sensitivity of CFTR.

The same is true for the NBDs: human NBD1 requires ATP binding to become stable enough for function, and NBD2 needs solubilizing mutations or assembly with other domains for (functional) stability. The F508del mutation renders NBD1 more unstable, albeit without much of a change in conformation. The melting temperature of the purified F508del domain is 36 °C, 5 °C lower than that of wild-type NBD1 and its half life in cells is lower as well.[36,122,141] It is no surprise that this difference generally has been considered to be

the cause of the temperature sensitivity of F508del-CFTR in cells (at 27 °C, F508del CFTR reaches the cell surface and is partially active).[146] The contrary is true, though (Figure 5): whereas low temperature indeed stabilizes F508del NBD1 to a similar level as wild-type NBD1, synthesis and folding at low temperature does not rescue the mutant domain, neither *in vitro* in isolation nor in CFTR in cells. Instead, domain assembly is slightly improved, which leads to some release from the ER to the cell-surface, where the low-temperature rescued F508del CFTR gains stability more significantly, resulting in clear rescue in steady-state experiments.[146] F508del NBD1 therefore is not temperature sensitive for folding, but once folded, is more stable at lower temperature. This disconnect between NBD1 and NBD2 in limited proteolysis has been seen in steady state upon chaperone manipulation[156] and in kinetic experiments with corrector drugs.[49,115,132] NBD1 folding is not rescued but domain assembly is, and this leads to stabilization of NBD2, making CFTR folding rather robust. Domain assembly compensates largely for domain defects.

Genotype-phenotype Expansion

Up to now, clinical data of 804 variants from CF patients have been included in the CFTR2 project (https://CFTR2.org). Sosnay et al initially annotated this database with functional Cl^- conductance- and Western blot steady-state expression (of ER and complex glycosylated forms as proxy for folding) data of the 65 most abundant variants expressed in a cell model.[12] Bihler et al.[9] extended the database with results for another set of 650 transfected variants and determined the functional response of transfected variants in the presence of the ETI compounds. These studies are extremely useful for personalized medicine to determine whether and to what extent a particular variant can be rescued by the modulators. This type of information has been contributing to FDA-EMA approval of Trikafta prescription for rare CF-causing variants.

The significance of such studies cannot be underestimated since the 2 most abundant CFTR missense variants F508del and G551D together are responsible for nearly 95% of all CF cases. The vast majority of other CF-causing variants are represented by small groups of patients, which renders their study more challenging were it only because of scarcity of clinical and functional data and limitations on scalability of standard assays for measuring protein biogenesis and function. New technologies are needed to analyze larger number of samples more efficiently that arise for instance from combinatorial testing of corrector compounds for precision medicine of rare variants.

An example of a promising approach to break down one of these barriers is the implementation of deep mutational scanning by Schlebach et al.[157] for analyzing the effect of clinically used correctors on cell surface residence of nearly 130 variants of the CFTR2 database expressed in HEK293T cells. Each variant was provided with a triple-HA tag in extracellular loop 4 for FACS analysis of plasma membrane fluorescence of single cells and a unique molecular identifier. The latter is needed to quantitate enrichment after Illumina sequencing of sorted clones to calculate the relative surface staining per variant with and without corrector treatment. Not only is this approach much more efficient than standard Western-blot methods used as proxy for folding and biosynthetic transport, but outcomes also correlate well with each other. FACS-based assays however do not discriminate between the high mannose- and complex-glycosylated form of CFTR, a glycoprotein subject to these post-translational modifications. Thus, without additional information they do not report directly on the trafficking route from ER to the cell surface which could be relevant for mechanistical reasons given that CFTR can reach the plasma membrane via non-conventional secretion,[158,159] nor on the efficiency of ER to Golgi transport. Nevertheless, data generated in this type of 'omic' approaches can be expected to contribute greatly to development of improved computational methods for predicting structural consequences of variants of unknown significance and their responses to small molecule modulators.

Exciting developments in machine learning and artificial intelligence are also starting to find their application in CFTR research. Balch and colleagues creatively adapted Variation Specific Profiling (VSP), a geo-statistical concept from oil and gas exploration to predict phenotype from genotype for CFTR.[160] In oil prospecting a discrete number of bore holes (50 ore more) are drilled at various places (x and y coordinates) in a terrain to assess confidently oil abundance (z coordinate). The three coordinates of the members constituting the sparse collection of boreholes are then used to infer the probable distribution and spatial covariance of oil reserves across the landscape with high confidence using statistical methods. The rationale being that the z value of positions in close proximity correlate better to each other than those at more distant spatial locations.

Anglès et al extended VSP and employed it on the set of most abundant pathogenic mutants of the CFTR2 database with well-characterized chloride conductance and ER-to-Golgi transport phenotypes, the equivalent of the boreholes.[161] By linking the position of a variant on the linear amino-acid sequence with Golgi transport and with function, the investigators created a continuous landscape of sequence-function variables that can be related to structure of the CFTR fold. This strategy allows predicting function-and-structure parameters for all amino acids in the CFTR sequence.

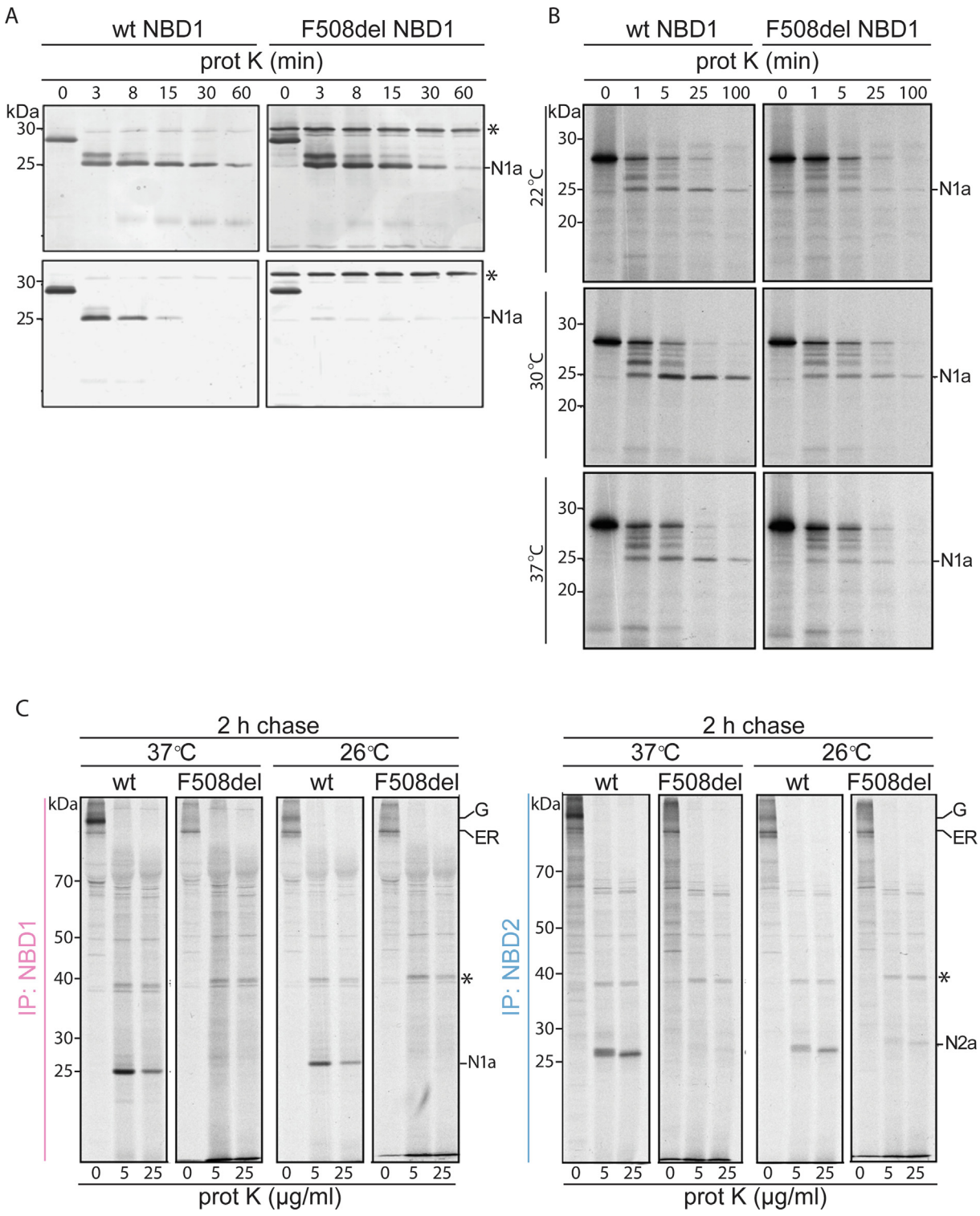


Figure 5. Low temperature rescues F508del domain assembly but not NBD1 folding. (A) The F508del mutation strongly destabilizes NBD1 at room temperature (bottom panels) as evidenced by the loss of the N1a fragment during limited proteolysis with proteinase K (Prot K); on ice (top panels), F508del and wild-type NBD1 were equally stable; (B) F508del-NBD1 produced in vitro is not rescued by translation at reduced temperature, as N1a is lost also at 22 °C. (C) Analysis of temperature effects on wild-type and F508del CFTR 2 h after radiolabeling shows that F508del NBD1 is not rescued at 26 °C (lack of protease-resistant N1a fragment). Domain assembly however is somewhat salvaged, as seen by the small amount of N2a fragment[115] (right panels). Asterisk denotes a background band not related to CFTR.

Innovative subsequent installments included pairing the spatial-covariance-analysis approach with chloride-conductance assays as well as transport experiments at 37 °C, and at 27 °C with or without the ETI compounds for the 63 most abundant CFTR2 variants.[162] The lower temperature is permissive for biosynthetic transport of F508del to the plasma membrane and permits the analysis of peripheral quality control at the plasma membrane. Entry of the experimental data in their spatial-covariance-analysis platform yielded novel insights to the authors into how variation in the genome affects the thermodynamic basis of the CFTR fold in vivo. Using this approach helped them assess the distinct structural determinants for quality control of CFTR at birth (in the ER) and death (instigated at the plasma membrane). This revealed an energetically destabilized region containing the ER-export motif in NBD1, which when mutated could be rescued only poorly by current corrector compounds. The vast amount of ER-to-Golgi transport data that came available through the extended annotation of CFTR2 by Bihler et al.[9] and the deep-scanning-mutagenesis experiments of McKee et al.[157] will provide an excellent opportunity to test and validate predictions of the VSP approach.

Last year investigators from Deepmind published AlphaMissense, a machine learning algorithm for forecasting pathogenic missense protein variants and whether a gene is essential for cell survival and fitness.[163] AlphaMissense combines AlphaFold methodology to predict protein structure from sequence, with evolutionary conservation to predict functional phenotypes and according to the investigators does so better than other variant effect predicting algorithms. Datasets from AlphaMissense also include pathogenicity predictions for CF. McDonald et al.[164] calibrated the output of AlphaMissense against CFTR metrics from the CFTR2 database,[9] a collection of 100 missense variants from their deep mutational scanning study[157] and the Golgi to ER transport index of the Balch lab.[161] The AM pathogenicity index had a relatively weak inverse correlation ($r = -0.37$) with the cell surface expression in the DMS study. Slightly better correlations were found with the trafficking index parameters ($r = -0.46$)[161] either at 37 °C or 27 °C. If we take these metrics as proxy for foldedness of CFTR variants than correlations are relatively modest. Adding the % Golgi dataset of the 560 missense mutations from Bihler et al.[9] might increase the correlation between foldedness and pathogenicity prediction. Perhaps inclusion of the % Golgi dataset of the 560 missense mutations from the Bihler study might increase the correlation between foldedness and pathogenicity prediction. More extensive benchmarking with other disease protein variants however is needed to fully appreciate the usefulness of AlphaMissense for understanding folding function questions.

Perspective

Many aspects of CFTR's folding trajectory have been uncovered and could be put to practical use in that this scientific advance facilitated development of clinical compounds that improved the lives of most CF patients. On the other hand, many questions still need to be answered before we fully understand how CFTR, and multispan membrane proteins in general, reach their mature form in vivo. An appreciable hurdle to answer this question is the relative inaccessibility of a sizeable part of CFTR because 12 TMs are embedded in the membrane. The cellular mechanism for inserting the TMs of CFTR is not understood and several scenarios can be entertained. Is this achieved solely by sec61, are all TMs inserted by the same translocon, or are additional translocons involved that can deal with the non-identical chemical environments to which the distinct TMs are exposed? Steady progress has been made with the discovery of additional translocons such as the EMC complex[165–168] and PAT, GEL, and BOS complexes[169,170] that cooperate with sec61 for insertion of the large number of different multispan membranes in the ER membrane.

Another major unanswered question is the molecular nature of the sequence and structural ques within CFTR that signal recognition for quality control by the ER. In a similar vein, what are the ER resident molecules-complexes that sense misfolded CFTR and direct it for degradative pathways. It could be anticipated that generic mechanisms in the ER are delegated to perform this task. An unexpected protein that recently entered the main stage of ER quality control for membrane proteins is the signal peptidase complex whose known role is co-translational cleavage of targeting peptide form secretory proteins. Host SPC however also has a post-translational role that is capitalized on by flaviviruses for processing their polyproteins[171] suggesting that it has other functions beyond cleavage of signal sequences. Zanotti et al discovered that SPC can cleave post-translationally cryptic sites in TMs of misfolded membrane proteins[172] in HEK293T cells (which do not express CFTR). It will be interesting to investigate whether SPC is also involved in recognition and disposal of class II CFTR mutants. The membrane chaperone calnexin, which for many years has been known to bind to monoglucosylated glycoproteins has also been implicated in CFTR homeostasis. Its role however has been somewhat enigmatic in part because perturbation of the calnexin interactions with CFTR does not impact turnover of CFTR.[173] Early studies showed that calnexin via its transmembrane domain could bind to MHCI and proteolipid protein in a glycan-independent manner.[174,175] Recently Bloemeke et al showed that a specific motif in the

calnexin transmembrane domain interacts within the membrane with many non-glycosylated misfolded membrane proteins, and elucidated the molecular basis for this interaction.[176] Given these new insights it will be opportune to revisit the role of calnexin in CFTR proteostasis.

Another important question in CFTR biogenesis concerns the coordination of the co-translational stage in CFTR folding and post-translational assembly of the domains. These are operationally considered as two discrete stages where the appearance of an ER band by SDS-PAGE is considered to reflect a co-translational event and a complex glycosylated Golgi form as a proxy for post-translational processes. There is good reason to consider that the distinction is less black and white, and that assembly already starts during translation. To investigate the transition from co- to post-translational processes we need to refine the existing in-vivo-folding assays or perhaps devise new ones. New experimental approaches and additional structures may also shed light on the elusive role of the non-structured R region and of its extensive post-translational modifications[177] on the folding trajectory of CFTR. Collectively these avenues may lead to a deeper understanding of CFTR folding pathways and of hierarchy and/or cooperativity of domain assembly, which we anticipate will contribute to the development of better small modulators that also improve for instance misfolded NBD1, which has been refractory to correctors or other mutations for which ETI does not offer relief.

CRediT authorship contribution statement

Peter van der Sluijs: Writing – review & editing, Writing – original draft, Supervision, Resources, Funding acquisition, Conceptualization. **Hanneke Hoelen:** Writing – review & editing, Project administration, Investigation, Formal analysis, Data curation. **Andre Schmidt:** Writing – review & editing. **Ineke Braakman:** Writing – review & editing, Writing – original draft, Supervision, Resources, Funding acquisition, Conceptualization.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by The Cystic Fibrosis Foundation, Innovation fund for Chemistry-LIFT, the Netherlands Cystic Fibrosis Foundation, Stichting Zeldzame Ziekten Fonds via Stichting Muco & Friends and the Netherlands Organization for Health Research and Development (Zon-MW), UK Cystic Fibrosis Trust, TKI Holland. We are grateful to Dr. Sana Ansari for artwork and editing, and we thank members of the Braakman-Van der Sluijs lab for discussions.

Received 23 February 2024;
Accepted 23 April 2024;
Available online xxxx

Keywords:
Cystic Fibrosis;
endoplasmic reticulum;
co-translational folding;
post-translational domain assembly;
stability

References

1. Welsh, M.J., Smith, A.E., (1993). Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. *Cell* **73**, 1251–1254.
2. Smith, J.J., Welsh, M.J., (1992). cAMP stimulates bicarbonate secretion across normal, but not cystic fibrosis airway epithelia. *J. Clin. Invest.* **89**, 1148–1153.
3. Lee, M.G., Wigley, W.C., Zeng, W., Noel, L.E., Marino, C. R., Thomas, P.J., et al., (1999). Regulation of Cl-/HCO3- exchange by cystic fibrosis transmembrane conductance regulator expressed in NIH 3T3 and HEK 293 cells. *J. Biol. Chem.* **274**, 3414–3421.
4. Thomas, C., Tampé, R., (2020). Structural and mechanistic principles of ABC transporters. *Annu. Rev. Biochem* **89**, 605–636.
5. Ong, T., Ramsey, B.W., (2023). Cystic fibrosis: A review. *J. Am. Med. Assoc.* **329**, 1859–1871.
6. Grasemann, H., Ratjen, F., (2023). Cystic fibrosis. *N. Engl. J. Med.* **389**, 1693–1707.
7. Pique, L., Graham, S., Pearl, M., Kharrazi, M., Schrijver, I., (2017). Cystic fibrosis newborn screening programs: implications of the CFTR variant spectrum in nonwhite patients. *Genet. Med.* **19**, 36–44.
8. Ideozu, J.E., Liu, M., Riley-Gillis, B.M., Paladugu, S.R., Rahimov, F., Krishnan, P., et al., (2023). Diversity of CFTR variants across ancestries characterized using 454,727 UK biobank whole exome sequences. *medRxiv.* 2023.03.10.23287070.
9. Bihler, H., Sivachenko, A., Millen, L., Bhatt, P., Patel, A.T., Chin, J., et al., (2024). In vitro modulator responsiveness of 655 CFTR variants found in people with cystic fibrosis. *J. Cyst. Fibros.*.
10. Hobbs, H.H., Russell, D.W., Brown, M.S., Goldstein, J.L., (1990). The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein. *Annu. Rev. Genet.* **24**, 133–170.
11. Rowe, S.M., Miller, S., Sorscher, E.J., (2005). Cystic fibrosis. *N. Engl. J. Med.* **352**, 1992–2001.

12. Sosnay, P.R., Siklosi, K.R., Van Goor, F., Kaniecki, K., Yu, H., Sharma, N., et al., (2013). Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. *Nat. Genet.* **45**, 1160–1167.
13. Okiyoneda, T., Barrière, H., Bagdány, M., Rabeh, W.M., Du, K., Höhfeld, J., et al., (2010). Peripheral protein quality control removes unfolded CFTR from the plasma membrane. *Science* **329**, 805–810.
14. Riordan, J.R., Rommens, J.M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., et al., (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* **245**, 1066–1073.
15. Balch, W.E., Roth, D.M., Hutt, D.M., (2011). Emergent properties of proteostasis in managing cystic fibrosis. *Cold Spring Harb. Perspect. Biol.* **3**
16. Van Goor, F., Hadida, S., Grootenhuis, P.D., Burton, B., Cao, D., Neuberger, T., et al., (2009). Rescue of CF airway epithelial cell function in vitro by a CFTR potentiator, VX-770. *PNAS* **106**, 18825–18830.
17. Rabeh, W.M., Bossard, F., Xu, H., Okiyoneda, T., Bagdany, M., Mulvihill, C.M., et al., (2012). Correction of both NBD1 energetics and domain interface is required to restore ΔF508 CFTR folding and function. *Cell* **148**, 150–163.
18. Mendoza, J.L., Schmidt, A., Li, Q., Nuvaga, E., Barrett, T., Bridges, R.J., et al., (2012). Requirements for efficient correction of ΔF508 CFTR revealed by analyses of evolved sequences. *Cell* **148**, 164–174.
19. Cholon, D.M., Quinney, N.L., Fulcher, M.L., Esther Jr., C. R., Das, J., Dokholyan, N.V., et al., (2014). Potentiator ivacaftor abrogates pharmacological correction of ΔF508 CFTR in cystic fibrosis. *Sci. Transl. Med.* **6**, 246ra96.
20. Veit, G., Avramescu, R.G., Perdomo, D., Phuan, P.W., Bagdany, M., Apaja, P.M., et al., (2014). Some gating potentiators, including VX-770, diminish ΔF508-CFTR functional expression. *Sci. Transl. Med.* **6**, 246ra97.
21. Schwarz, C., Sutharsan, S., Epaud, R., Klingsberg, R.C., Fischer, R., Rowe, S.M., et al., (2021). Tezacaftor/ivacaftor in people with cystic fibrosis who stopped lumacaftor/ivacaftor due to respiratory adverse events. *J. Cyst. Fibros.* **20**, 228–233.
22. Rowe, S.M., Daines, C., Ringshausen, F.C., Kerem, E., Wilson, J., Tullis, E., et al., (2017). Tezacaftor-ivacaftor in residual-function heterozygotes with cystic fibrosis. *N. Engl. J. Med.* **377**, 2024–2035.
23. Keating, D., Marigowda, G., Burr, L., Daines, C., Mall, M. A., McKone, E.F., et al., (2018). VX-445-tezacaftor-ivacaftor in patients with cystic fibrosis and one or two Phe508del alleles. *N. Engl. J. Med.* **379**, 1612–1620.
24. He, L., Aleksandrov, A.A., Serohijos, A.W., Hegedus, T., Aleksandrov, L.A., Cui, L., et al., (2008). Multiple membrane-cytoplasmic domain contacts in the cystic fibrosis transmembrane conductance regulator (CFTR) mediate regulation of channel gating. *J. Biol. Chem.* **283**, 26383–26390.
25. Serohijos, A.W., Hegedus, T., Aleksandrov, A.A., He, L., Cui, L., Dokholyan, N.V., et al., (2008). Phenylalanine-508 mediates a cytoplasmic-membrane domain contact in the CFTR 3D structure crucial to assembly and channel function. *PNAS* **105**, 3256–3261.
26. Wang, W., Linsdell, P., (2012). Relative movements of transmembrane regions at the outer mouth of the cystic fibrosis transmembrane conductance regulator channel pore during channel gating. *J. Biol. Chem.* **287**, 32136–32146.
27. Gao, X., Bai, Y., Hwang, T.C., (2013). Cysteine scanning of CFTR's first transmembrane segment reveals its plausible roles in gating and permeation. *Biophys. J .* **104**, 786–797.
28. Mornon, J.P., Hoffmann, B., Jonic, S., Lehn, P., Callebaut, I., (2015). Full-open and closed CFTR channels, with lateral tunnels from the cytoplasm and an alternative position of the F508 region, as revealed by molecular dynamics. *Cell. Mol. Life Sci.* **72**, 1377–1403.
29. Callebaut, I., Hoffmann, B., Lehn, P., Mornon, J.P., (2017). Molecular modelling and molecular dynamics of CFTR. *Cell. Mol. Life Sci.* **74**, 3–22.
30. Simhaev, L., McCarty, N.A., Ford, R.C., Senderowitz, H., (2017). Molecular dynamics flexible fitting simulations identify new models of the closed state of the cystic fibrosis transmembrane conductance regulator protein. *J. Chem. Inf. Model.* **57**, 1932–1946.
31. Zhang, Z., Chen, J., (2016). Atomic structure of the cystic fibrosis transmembrane conductance regulator. *Cell* **167**, 1586–1597.e9.
32. Zhang, Z., Liu, F., Chen, J., (2017). Conformational changes of CFTR upon phosphorylation and ATP binding. *Cell* **170**, 483–491.e8.
33. Liu, F., Zhang, Z., Csanády, L., Gadsby, D.C., Chen, J., (2017). Molecular structure of the human CFTR ion channel. *Cell* **169**, 85–95.e8.
34. Zhang, Z., Liu, F., Chen, J., (2018). Molecular structure of the ATP-bound, phosphorylated human CFTR. *PNAS* **115**, 12757–12762.
35. Fay, J.F., Aleksandrov, L.A., Jensen, T.J., Cui, L.L., Kousouros, J.N., He, L., et al., (2018). Cryo-EM visualization of an active high open probability CFTR anion channel. *Biochemistry* **57**, 6234–6246.
36. Hoelen, H., Kleizen, B., Schmidt, A., Richardson, J., Charitou, P., Thomas, P.J., et al., (2010). The primary folding defect and rescue of ΔF508 CFTR emerge during translation of the mutant domain. *PLoS One* **5**, e15458.
37. Aleksandrov, A.A., Kota, P., Cui, L., Jensen, T., Alekseev, A.E., Reyes, S., et al., (2012). Allosteric modulation balances thermodynamic stability and restores function of ΔF508 CFTR. *J. Mol. Biol.* **419**, 41–60.
38. Ford, R.C., Marshall-Sabey, D., Schuetz, J., (2020). Linker domains: why ABC transporters 'live in fragments no longer'. *Trends Biochem. Sci* **45**, 137–148.
39. Naren, A.P., Cormet-Boyaka, E., Fu, J., Villain, M., Blalock, J.E., Quick, M.W., et al., (1999). CFTR chloride channel regulation by an interdomain interaction. *Science* **286**, 544–548.
40. King, S.A., Sorscher, E.J., (2000). R-domain interactions with distal regions of CFTR lead to phosphorylation and activation. *Biochemistry* **39**, 9868–9875.
41. Baker, J.M., Hudson, R.P., Kanelis, V., Choy, W.Y., Thibodeau, P.H., Thomas, P.J., et al., (2007). CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. *Nat. Struct. Mol. Biol.* **14**, 738–745.
42. Seavilleklein, G., Amer, N., Evagelidis, A., Chappe, F., Irvine, T., Hanrahan, J.W., et al., (2008). PKC phosphorylation modulates PKA-dependent binding of the R domain to other domains of CFTR. *Am. J. Physiol. Cell Physiol.* **295**, C1366–C1375.

43. Kanelis, V., Hudson, R.P., Thibodeau, P.H., Thomas, P. J., Forman-Kay, J.D., (2010). NMR evidence for differential phosphorylation-dependent interactions in WT and DeltaF508 CFTR. *EMBO J.* **29**, 263–277.
44. Wang, G., Duan, D.D., (2012). Regulation of activation and processing of the cystic fibrosis transmembrane conductance regulator (CFTR) by a complex electrostatic interaction between the regulatory domain and cytoplasmic loop 3. *J. Biol. Chem.* **287**, 40484–40492.
45. Bozoky, Z., Krzeminski, M., Muhandiram, R., Birtley, J.R., Al-Zahrani, A., Thomas, P.J., et al., (2013). Regulatory R region of the CFTR chloride channel is a dynamic integrator of phospho-dependent intra- and intermolecular interactions. *PNAS* **110**, E4427–E4436.
46. Csanády, L., Chan, K.W., Seto-Young, D., Kopsco, D.C., Nairn, A.C., Gadsby, D.C., (2000). Severed channels probe regulation of gating of cystic fibrosis transmembrane conductance regulator by its cytoplasmic domains. *J. Gen. Physiol.* **116**, 477–500.
47. Ostedgaard, L.S., Zabner, J., Vermeer, D.W., Rokhlina, T., Karp, P.H., Stecenko, A.A., et al., (2002). CFTR with a partially deleted R domain corrects the cystic fibrosis chloride transport defect in human airway epithelia in vitro and in mouse nasal mucosa in vivo. *PNAS* **99**, 3093–3098.
48. Du, K., Lukacs, G.L., (2009). Cooperative assembly and misfolding of CFTR domains in vivo. *Mol. Biol. Cell* **20**, 1903–1915.
49. Hillenaar, T., Beekman, J., van der Sluijs, P., Braakman, I., (2022). Redefining hypo- and hyper-responding phenotypes of CFTR mutants for understanding and therapy. *Int. J. Mol. Sci.* **23**
50. Winter, M.C., Welsh, M.J., (1997). Stimulation of CFTR activity by its phosphorylated R domain. *Nature* **389**, 294–296.
51. Johnson, Z.L., Chen, J., (2017). Structural basis of substrate recognition by the multidrug resistance protein MRP1. *Cell* **168**, 1075–1085.e9.
52. Johnson, Z.L., Chen, J., (2018). ATP binding enables substrate release from multidrug resistance protein 1. *Cell* **172**, 81–89.e10.
53. Li, N., Wu, J.X., Ding, D., Cheng, J., Gao, N., Chen, L., (2017). Structure of a pancreatic ATP-sensitive potassium channel. *Cell* **168**, 101–110.e10.
54. Lee, K.P.K., Chen, J., MacKinnon, R., (2017). Molecular structure of human KATP in complex with ATP and ADP. *Elife* **6**
55. Alam, A., Locher, K.P., (2023). Structure and mechanism of human ABC transporters. *Annu. Rev. Biophys.* **52**, 275–300.
56. Bakos, E., Evers, R., Szakács, G., Tusnády, G.E., Welker, E., Szabó, K., et al., (1998). Functional multidrug resistance protein (MRP1) lacking the N-terminal transmembrane domain. *J. Biol. Chem.* **273**, 32167–32175.
57. Westlake, C.J., Qian, Y.M., Gao, M., Vasa, M., Cole, S.P., Deeley, R.G., (2003). Identification of the structural and functional boundaries of the multidrug resistance protein 1 cytoplasmic loop 3. *Biochemistry* **42**, 14099–14113.
58. Naren, A.P., Nelson, D.J., Xie, W., Jovov, B., Pevsner, J., Bennett, M.K., et al., (1997). Regulation of CFTR chloride channels by syntaxin and Munc18 isoforms. *Nature* **390**, 302–305.
59. Naren, A.P., Quick, M.W., Collawn, J.F., Nelson, D.J., Kirk, K.L., (1998). Syntaxin 1A inhibits CFTR chloride channels by means of domain-specific protein-protein interactions. *PNAS* **95**, 10972–10977.
60. Bilan, F., Thoreau, V., Nacfer, M., Dérand, R., Norez, C., Cantereau, A., et al., (2004). Syntaxin 8 impairs trafficking of cystic fibrosis transmembrane conductance regulator (CFTR) and inhibits its channel activity. *J. Cell Sci.* **117**, 1923–1935.
61. Gee, H.Y., Tang, B.L., Kim, K.H., Lee, M.G., (2010). Syntaxin 16 binds to cystic fibrosis transmembrane conductance regulator and regulates its membrane trafficking in epithelial cells. *J. Biol. Chem.* **285**, 35519–35527.
62. Sabirzhanova, I., Boinot, C., Guggino, W.B., Cebotaru, L., (2018). Syntaxin 8 and the Endoplasmic Reticulum Processing of ΔF508-CFTR. *Cell. Physiol. Biochem.* **51**, 1489–1499.
63. Arora, K., Liyanage, P., Zhong, Q., Naren, A.P., (2021). A SNARE protein Syntaxin 17 captures CFTR to potentiate autophagosomal clearance under stress. *FASEB J.* **35**, e21185.
64. Thelin, W.R., Chen, Y., Gentzsch, M., Kreda, S.M., Sallee, J.L., Scarlett, C.O., et al., (2007). Direct interaction with filamins modulates the stability and plasma membrane expression of CFTR. *J. Clin. Invest.* **117**, 364–374.
65. Playford, M.P., Nurminen, E., Pentikäinen, O.T., Milgram, S.L., Hartwig, J.H., Stossel, T.P., et al., (2010). Cystic fibrosis transmembrane conductance regulator interacts with multiple immunoglobulin domains of filamin A. *J. Biol. Chem.* **285**, 17156–17165.
66. Sabusap, C.M., Joshi, D., Simhaev, L., Oliver, K.E., Senderowitz, H., van Willigen, M., et al., (2021). The CFTR P67L variant reveals a key role for N-terminal lasso helices in channel folding, maturation, and pharmacologic rescue. *J. Biol. Chem.* **296**, 100598
67. Wong, S.L., Awatade, N.T., Astore, M.A., Allan, K.M., Carnell, M.J., Slapetova, I., et al., (2022). Molecular dynamics and functional characterization of I37R-CFTR lasso mutation provide insights into channel gating activity. *iScience* **25**, 103710.
68. Fu, J., Ji, H.L., Naren, A.P., Kirk, K.L., (2001). A cluster of negative charges at the amino terminal tail of CFTR regulates ATP-dependent channel gating. *J. Physiol.* **536**, 459–470.
69. Hwang, T.C., Yeh, J.T., Zhang, J., Yu, Y.C., Yeh, H.I., Destefano, S., (2018). Structural mechanisms of CFTR function and dysfunction. *J. Gen. Physiol.* **150**, 539–570.
70. Laselva, O., Bartlett, C., Gunawardena, T.N.A., Ouyang, H., Eckford, P.D.W., Moraes, T.J., et al., (2021). Rescue of multiple class II CFTR mutations by elexacaftor +tezacaftor+ivacaftor mediated in part by the dual activities of elexacaftor as both corrector and potentiator. *Eur. Respir. J.* **57**
71. Shaughnessy, C.A., Zeitlin, P.L., Bratcher, P.E., (2021). Elexacaftor is a CFTR potentiator and acts synergistically with ivacaftor during acute and chronic treatment. *Sci. Rep.* **11**, 19810.
72. Kleizen, B., van Willigen, M., Mijnders, M., Peters, F., Grudniewska, M., Hillenaar, T., et al., (2021). Co-translational folding of the first transmembrane domain of ABC-transporter CFTR is supported by assembly with the first cytosolic domain. *J. Mol. Biol.* **433**, 166955

73. Okiyoneda, T., Veit, G., Dekkers, J.F., Bagdany, M., Soya, N., Xu, H., et al., (2013). Mechanism-based corrector combination restores ΔF508-CFTR folding and function. *Nat. Chem. Biol.* **9**, 444–454.
74. Loo, T.W., Bartlett, M.C., Clarke, D.M., (2013). Corrector VX-809 stabilizes the first transmembrane domain of CFTR. *Biochem. Pharmacol.* **86**, 612–619.
75. Loo, T.W., Clarke, D.M., (2017). Corrector VX-809 promotes interactions between cytoplasmic loop one and the first nucleotide-binding domain of CFTR. *Biochem. Pharmacol.* **136**, 24–31.
76. Fiedorczuk, K., Chen, J., (2022). Mechanism of CFTR correction by type I folding correctors. *Cell* **185**, 158–168. e11.
77. Hudson, R.P., Dawson, J.E., Chong, P.A., Yang, Z., Millen, L., Thomas, P.J., et al., (2017). Direct binding of the corrector VX-809 to human CFTR NBD1: Evidence of an allosteric coupling between the binding site and the NBD1:CL4 interface. *Mol. Pharmacol.* **92**, 124–135.
78. Baatallah, N., Elbahnsi, A., Mornon, J.P., Chevalier, B., Pranke, I., Servel, N., et al., (2021). Pharmacological chaperones improve intra-domain stability and inter-domain assembly via distinct binding sites to rescue misfolded CFTR. *Cell. Mol. Life Sci.* **78**, 7813–7829.
79. Veit, G., Roldan, A., Hancock, M.A., Da Fonte, D.F., Xu, H., Hussein, M., et al., (2020). Allosteric folding correction of F508del and rare CFTR mutants by elexacaftor-tezacaftor-ivacaftor (Trikafta) combination. *JCI Insight* **5**
80. Liu, F., Zhang, Z., Levit, A., Levring, J., Touhara, K.K., Shoichet, B.K., et al., (2019). Structural identification of a hotspot on CFTR for potentiation. *Science* **364**, 1184–1188.
81. Yeh, H.I., Qiu, L., Sohma, Y., Conrath, K., Zou, X., Hwang, T.C., (2019). Identifying the molecular target sites for CFTR potentiators GLPG1837 and VX-770. *J. Gen. Physiol.* **151**, 912–928.
82. Jih, K.Y., Hwang, T.C., (2013). Vx-770 potentiates CFTR function by promoting decoupling between the gating cycle and ATP hydrolysis cycle. *PNAS* **110**, 4404–4409.
83. Laselva, O., Qureshi, Z., Zeng, Z.W., Petrotchenko, E.V., Ramjeesingh, M., Hamilton, C.M., et al., (2021). Identification of binding sites for ivacaftor on the cystic fibrosis transmembrane conductance regulator. *iScience* **24**, 102542.
84. Byrnes, L.J., Xu, Y., Qiu, X., Hall, J.D., West, G.M., (2018). Sites associated with Kalydeco binding on human cystic fibrosis transmembrane conductance regulator revealed by hydrogen/deuterium exchange. *Sci. Rep.* **8**, 4664.
85. Tartaglia, G.G., Pechmann, S., Dobson, C.M., Vendruscolo, M., (2007). Life on the edge: a link between gene expression levels and aggregation rates of human proteins. *Trends Biochem. Sci* **32**, 204–206.
86. Vecchi, G., Sormanni, P., Mannini, B., Vandelli, A., Tartaglia, G.G., Dobson, C.M., et al., (2020). Proteome-wide observation of the phenomenon of life on the edge of solubility. *PNAS* **117**, 1015–1020.
87. Ward, C.L., Kopito, R.R., (1994). Intracellular turnover of cystic fibrosis transmembrane conductance regulator. Inefficient processing and rapid degradation of wild-type and mutant proteins. *J. Biol. Chem.* **269**, 25710–25718.
88. Ward, C.L., Omura, S., Kopito, R.R., (1995). Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell* **83**, 121–127.
89. Varga, K., Jurkuvenaite, A., Wakefield, J., Hong, J.S., Guimbellot, J.S., Venglarik, C.J., et al., (2004). Efficient intracellular processing of the endogenous cystic fibrosis transmembrane conductance regulator in epithelial cell lines. *J. Biol. Chem.* **279**, 22578–22584.
90. Braakman, I., Hebert, D.N., (2013). Protein folding in the endoplasmic reticulum. *Cold Spring Harb. Perspect. Biol.* **5**, a013201
91. Braakman, I., Hoover-Litty, H., Wagner, K.R., Helenius, A., (1991). Folding of influenza hemagglutinin in the endoplasmic reticulum. *J. Cell Biol.* **114**, 401–411.
92. Vabulas, R.M., Hartl, F.U., (2005). Protein synthesis upon acute nutrient restriction relies on proteasome function. *Science* **310**, 1960–1963.
93. Myers, J.K., Oas, T.G., (2002). Mechanism of fast protein folding. *Annu. Rev. Biochem* **71**, 783–815.
94. Oliver, K.E., Rauscher, R., Mijnders, M., Wang, W., Wolpert, M.J., Maya, J., et al., (2019). Slowing ribosome velocity restores folding and function of mutant CFTR. *J. Clin. Invest.* **129**, 5236–5253.
95. Rauscher, R., Bampi, G.B., Guevara-Ferrer, M., Santos, L.A., Joshi, D., Mark, D., et al., (2021). Positive epistasis between disease-causing missense mutations and silent polymorphism with effect on mRNA translation velocity. *PNAS* **118**
96. Cassaignau, A.M.E., Cabrita, L.D., Christodoulou, J., (2020). How does the ribosome fold the proteome? *Annu. Rev. Biochem* **89**, 389–415.
97. Bartoszewski, R.A., Jablonsky, M., Bartoszewska, S., Stevenson, L., Dai, Q., Kappes, J., et al., (2010). A synonymous single nucleotide polymorphism in DeltaF508 CFTR alters the secondary structure of the mRNA and the expression of the mutant protein. *J. Biol. Chem.* **285**, 28741–28748.
98. Kirchner, S., Cai, Z., Rauscher, R., Kastelic, N., Anding, M., Czech, A., et al., (2017). Alteration of protein function by a silent polymorphism linked to tRNA abundance. *PLoS Biol.* **15**, e2000779
99. Kimchi-Sarfaty, C., Oh, J.M., Kim, I.W., Sauna, Z.E., Calcagno, A.M., Ambudkar, S.V., et al., (2007). A “silent” polymorphism in the MDR1 gene changes substrate specificity. *Science* **315**, 525–528.
100. Zhou, M., Guo, J., Cha, J., Chae, M., Chen, S., Barral, J. M., et al., (2013). Non-optimal codon usage affects expression, structure and function of clock protein FRQ. *Nature* **495**, 111–115.
101. Komar, A.A., Samatova, E., Rodnina, M.V., (2023). Translation rates and protein folding. *J. Mol. Biol.*, 168384.
102. Kim, S.J., Yoon, J.S., Shishido, H., Yang, Z., Rooney, L. A., Barral, J.M., et al., (2015). Protein folding. Translational tuning optimizes nascent protein folding in cells. *Science* **348**, 444–448.
103. Kim, S.J., Skach, W.R., (2012). Mechanisms of CFTR folding at the endoplasmic reticulum. *Front. Pharmacol.* **3**, 201.
104. Lu, Y., Xiong, X., Helm, A., Kimani, K., Bragin, A., Skach, W.R., (1998). Co- and posttranslational translocation mechanisms direct cystic fibrosis transmembrane

conductance regulator N terminus transmembrane assembly. *J. Biol. Chem.* **273**, 568–576.
105. Skach, W.R., (2000). Defects in processing and trafficking of the cystic fibrosis transmembrane conductance regulator. *Kidney Int.* **57**, 825–831.
106. Carveth, K., Buck, T., Anthony, V., Skach, W.R., (2002). Cooperativity and flexibility of cystic fibrosis transmembrane conductance regulator transmembrane segments participate in membrane localization of a charged residue. *J. Biol. Chem.* **277**, 39507–39514.
107. Hartmann, E., Rapoport, T.A., Lodish, H.F., (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. *PNAS* **86**, 5786–5790.
108. von Heijne, G., Gavel, Y., (1988). Topogenic signals in integral membrane proteins. *Eur. J. Biochem.* **174**, 671–678.
109. von Heijne, G., (1989). Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature* **341**, 456–458.
110. Mulvihill, C.M., Deber, C.M., (2012). Structural basis for misfolding at a disease phenotypic position in CFTR: comparison of TM3/4 helix-loop-helix constructs with TM4 peptides. *Biochim. Biophys. Acta* **1818**, 49–54.
111. Krainer, G., Treff, A., Hartmann, A., Stone, T.A., Schenkel, M., Keller, S., et al., (2018). A minimal helical-hairpin motif provides molecular-level insights into misfolding and pharmacological rescue of CFTR. *Commun Biol.* **1**, 154.
112. Tector, M., Hartl, F.U., (1999). An unstable transmembrane segment in the cystic fibrosis transmembrane conductance regulator. *EMBO J.* **18**, 6290–6298.
113. Enquist, K., Fransson, M., Boekel, C., Bengtsson, I., Geiger, K., Lang, L., et al., (2009). Membrane-integration characteristics of two ABC transporters, CFTR and P-glycoprotein. *J. Mol. Biol.* **387**, 1153–1164.
114. Kleizen, B., van Vlijmen, T., de Jonge, H.R., Braakman, I., (2005). Folding of CFTR is predominantly cotranslational. *Mol. Cell* **20**, 277–287.
115. Im, J., Hillenaar, T., Yeoh, H.Y., Sahasrabudhe, P., Mijnders, M., van Willigen, M., et al., (2023). ABC-transporter CFTR folds with high fidelity through a modular, stepwise pathway. *Cell. Mol. Life Sci.* **80**, 33.
116. Cui, L., Aleksandrov, L., Chang, X.B., Hou, Y.X., He, L., Hegedus, T., et al., (2007). Domain interdependence in the biosynthetic assembly of CFTR. *J. Mol. Biol.* **365**, 981–994.
117. Ren, H.Y., Grove, D.E., De La Rosa, O., Houck, S.A., Sopha, P., Van Goor, F., et al., (2013). VX-809 corrects folding defects in cystic fibrosis transmembrane conductance regulator protein through action on membrane-spanning domain 1. *Mol. Biol. Cell* **24**, 3016–3024.
118. Hoffmann, B., Elbahnsi, A., Lehn, P., Décout, J.L., Pietrucci, F., Mornon, J.P., et al., (2018). Combining theoretical and experimental data to decipher CFTR 3D structures and functions. *Cell. Mol. Life Sci.* **75**, 3829–3855.
119. Wang, W., He, Z., O'Shaughnessy, T.J., Rux, J., Reenstra, W.W., (2002). Domain-domain associations in cystic fibrosis transmembrane conductance regulator. *Am. J. Physiol. Cell Physiol.* **282**, C1170–C1180.
120. Khushoo, A., Yang, Z., Johnson, A.E., Skach, W.R., (2011). Ligand-driven vectorial folding of ribosome-bound human CFTR NBD1. *Mol. Cell* **41**, 682–692.
121. Qu, B.H., Strickland, E.H., Thomas, P.J., (1997). Localization and suppression of a kinetic defect in cystic fibrosis transmembrane conductance regulator folding. *J. Biol. Chem.* **272**, 15739–15744.
122. Qu, B.H., Thomas, P.J., (1996). Alteration of the cystic fibrosis transmembrane conductance regulator folding pathway. *J. Biol. Chem.* **271**, 7261–7264.
123. Lewis, H.A., Buchanan, S.G., Burley, S.K., Conners, K., Dickey, M., Dorwart, M., et al., (2004). Structure of nucleotide-binding domain 1 of the cystic fibrosis transmembrane conductance regulator. *EMBO J.* **23**, 282–293.
124. Aleksandrov, L., Mengos, A., Chang, X., Aleksandrov, A., Riordan, J.R., (2001). Differential interactions of nucleotides at the two nucleotide binding domains of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* **276**, 12918–12923.
125. Du, K., Sharma, M., Lukacs, G.L., (2005). The DeltaF508 cystic fibrosis mutation impairs domain-domain interactions and arrests post-translational folding of CFTR. *Nat. Struct. Mol. Biol.* **12**, 17–25.
126. Aleksandrov, L., Aleksandrov, A.A., Chang, X.B., Riordan, J.R., (2002). The first nucleotide binding domain of cystic fibrosis transmembrane conductance regulator is a site of stable nucleotide interaction, whereas the second is a site of rapid turnover. *J. Biol. Chem.* **277**, 15419–15425.
127. Basso, C., Vergani, P., Nairn, A.C., Gadsby, D.C., (2003). Prolonged nonhydrolytic interaction of nucleotide with CFTR's NH2-terminal nucleotide binding domain and its role in channel gating. *J. Gen. Physiol.* **122**, 333–348.
128. Hwang, T.C., Braakman, I., van der Sluijs, P., Callebaut, I., (2023). Structure basis of CFTR folding, function and pharmacology. *J. Cyst. Fibros.* **22** (Suppl 1), S5–S11.
129. Kleizen, B., de Mattos, E., Papaioannou, O., Monti, M., Tartaglia, G.G., van der Sluijs, P., et al., (2023). Transmembrane helices 7 and 8 confer aggregation sensitivity to the cystic fibrosis transmembrane conductance regulator. *Int. J. Mol. Sci.* **24**
130. Ko, S.B., Zeng, W., Dorwart, M.R., Luo, X., Kim, K.H., Millen, L., et al., (2004). Gating of CFTR by the STAS domain of SLC26 transporters. *Nat. Cell Biol.* **6**, 343–350.
131. Vernon, R.M., Chong, P.A., Lin, H., Yang, Z., Zhou, Q., Aleksandrov, A.A., et al., (2017). Stabilization of a nucleotide-binding domain of the cystic fibrosis transmembrane conductance regulator yields insight into disease-causing mutations. *J. Biol. Chem.* **292**, 14147–14164.
132. van Willigen, M., Vonk, A.M., Yeoh, H.Y., Kruisselbrink, E., Kleizen, B., van der Ent, C.K., et al., (2019). Folding-function relationship of the most common cystic fibrosis-causing CFTR conductance mutants. *Life Sci Alliance* **2**
133. Wang, Y., Loo, T.W., Bartlett, M.C., Clarke, D.M., (2007). Modulating the folding of P-glycoprotein and cystic fibrosis transmembrane conductance regulator truncation mutants with pharmacological chaperones. *Mol. Pharmacol.* **71**, 751–758.
134. Loo, T.W., Clarke, D.M., (2008). Mutational analysis of ABC proteins. *Arch. Biochem. Biophys.* **476**, 51–64.

135. Dolgikh, D.A., Gilmanshin, R.I., Brazhnikov, E.V., Bychkova, V.E., Semisotnov, G.V., Venyaminov, S., et al., (1981). Alpha-Lactalbumin: compact state with fluctuating tertiary structure? *FEBS Lett.* **136**, 311–315.
136. Ohgushi, M., Wada, A., (1983). 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett.* **164**, 21–24.
137. Chen, E.Y., Bartlett, M.C., Loo, T.W., Clarke, D.M., (2004). The DeltaF508 mutation disrupts packing of the transmembrane segments of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* **279**, 39620–39627.
138. Mense, M., Vergani, P., White, D.M., Altberg, G., Nairn, A. C., Gadsby, D.C., (2006). In vivo phosphorylation of CFTR promotes formation of a nucleotide-binding domain heterodimer. *EMBO J.* **25**, 4728–4739.
139. He, L., Kota, P., Aleksandrov, A.A., Cui, L., Jensen, T., Dokholyan, N.V., et al., (2013). Correctors of ΔF508 CFTR restore global conformational maturation without thermally stabilizing the mutant protein. *FASEB J.* **27**, 536–545.
140. Farinha, C.M., King-Underwood, J., Sousa, M., Correia, A.R., Henriques, B.J., Roxo-Rosa, M., et al., (2013). Revertants, low temperature, and correctors reveal the mechanism of F508del-CFTR rescue by VX-809 and suggest multiple agents for full correction. *Chem. Biol.* **20**, 943–955.
141. Lukacs, G.L., Mohamed, A., Kartner, N., Chang, X.B., Riordan, J.R., Grinstein, S., (1994). Conformational maturation of CFTR but not its mutant counterpart (delta F508) occurs in the endoplasmic reticulum and requires ATP. *EMBO J.* **13**, 6076–6086.
142. Pasyk, E.A., Foskett, J.K., (1995). Mutant (delta F508) cystic fibrosis transmembrane conductance regulator Cl- channel is functional when retained in endoplasmic reticulum of mammalian cells. *J. Biol. Chem.* **270**, 12347–12350.
143. Seibert, F.S., Chang, X.B., Aleksandrov, A.A., Clarke, D. M., Hanrahan, J.W., Riordan, J.R., (1999). Influence of phosphorylation by protein kinase A on CFTR at the cell surface and endoplasmic reticulum. *Biochim. Biophys. Acta* **1461**, 275–283.
144. Protasevich, I., Yang, Z., Wang, C., Atwell, S., Zhao, X., Emtage, S., et al., (2010). Thermal unfolding studies show the disease causing F508del mutation in CFTR thermodynamically destabilizes nucleotide-binding domain 1. *Protein Sci.* **19**, 1917–1931.
145. Qu, B.H., Strickland, E., Thomas, P.J., (1997). Cystic fibrosis: a disease of altered protein folding. *J. Bioenerg. Biomembr.* **29**, 483–490.
146. Denning, G.M., Anderson, M.P., Amara, J.F., Marshall, J., Smith, A.E., Welsh, M.J., (1992). Processing of mutant cystic fibrosis transmembrane conductance regulator is temperature-sensitive. *Nature* **358**, 761–764.
147. Brown, C.R., Hong-Brown, L.Q., Biwersi, J., Verkman, A. S., Welch, W.J., (1996). Chemical chaperones correct the mutant phenotype of the delta F508 cystic fibrosis transmembrane conductance regulator protein. *Cell Stress Chaperones* **1**, 117–125.
148. Teem, J.L., Berger, H.A., Ostedgaard, L.S., Rich, D.P., Tsui, L.C., Welsh, M.J., (1993). Identification of revertants for the cystic fibrosis delta F508 mutation using STE6-CFTR chimeras in yeast. *Cell* **73**, 335–346.
149. Chang, X.B., Cui, L., Hou, Y.X., Jensen, T.J., Aleksandrov, A.A., Mengos, A., et al., (1999). Removal of multiple arginine-framed trafficking signals overcomes misprocessing of delta F508 CFTR present in most patients with cystic fibrosis. *Mol. Cell* **4**, 137–142.
150. Thibodeau, P.H., Richardson 3rd, J.M., Wang, W., Millen, L., Watson, J., Mendoza, J.L., et al., (2010). The cystic fibrosis-causing mutation deltaF508 affects multiple steps in cystic fibrosis transmembrane conductance regulator biogenesis. *J. Biol. Chem.* **285**, 35825–35835.
151. Doyle, D.A., Morais Cabral, J., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., et al., (1998). The structure of the potassium channel: molecular basis of K+ conduction and selectivity. *Science* **280**, 69–77.
152. Dutzler, R., Campbell, E.B., Cadene, M., Chait, B.T., MacKinnon, R., (2002). X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity. *Nature* **415**, 287–294.
153. Hessa, T., Meindl-Beinker, N.M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., et al., (2007). Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* **450**, 1026–1030.
154. Pitonzo, D., Yang, Z., Matsumura, Y., Johnson, A.E., Skach, W.R., (2009). Sequence-specific retention and regulated integration of a nascent membrane protein by the endoplasmic reticulum Sec61 translocon. *Mol. Biol. Cell* **20**, 685–698.
155. Hämmerle, M.M., Aleksandrov, A.A., Chang, X.B., Riordan, J.R., (2000). A novel CFTR disease-associated mutation causes addition of an extra N-linked oligosaccharide. *Glycoconj. J.* **17**, 807–813.
156. Rosser, M.F., Grove, D.E., Chen, L., Cyr, D.M., (2008). 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* **19**, 4570–4579.
157. McKee, A.G., McDonald, E.F., Penn, W.D., Kuntz, C.P., Noguera, K., Chamness, L.M., et al., (2023). General trends in the effects of VX-661 and VX-445 on the plasma membrane expression of clinical CFTR variants. *Cell Chem. Biol.* **30**, 632–642.e5.
158. Yoo, J.S., Moyer, B.D., Bannykh, S., Yoo, H.M., Riordan, J.R., Balch, W.E., (2002). Non-conventional trafficking of the cystic fibrosis transmembrane conductance regulator through the early secretory pathway. *J. Biol. Chem.* **277**, 11401–11409.
159. Gee, H.Y., Noh, S.H., Tang, B.L., Kim, K.H., Lee, M.G., (2011). Rescue of ΔF508-CFTR trafficking via a GRASP-dependent unconventional secretion pathway. *Cell* **146**, 746–760.
160. Wang, C., Balch, W.E., (2018). Bridging genomics to phenomics at atomic resolution through variation spatial profiling. *Cell Rep.* **24**, 2013–2028.e6.
161. Anglès, F., Wang, C., Balch, W.E., (2022). Author correction: Spatial covariance analysis reveals the residue-by-residue thermodynamic contribution of variation to the CFTR fold. *Commun Biol.* **5**, 494.

162. Wang, C., Anglès, F., Balch, W.E., (2022). Triangulating variation in the population to define mechanisms for precision management of genetic disease. *Structure* **30**, 1190–1207.e5.
163. Cheng, J., Novati, G., Pan, J., Bycroft, C., Žemgulytė, A., Applebaum, T., et al., (2023). Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* **381**, eadg7492
164. McDonald, E.F., Oliver, K.E., Schlebach, J.P., Meiler, J., Plate, L., (2024). Benchmarking AlphaMissense pathogenicity predictions against cystic fibrosis variants. *PLoS One* **19**, e0297560
165. Shurtleff, M.J., Itzhak, D.N., Hussmann, J.A., Schirle Oakdale, N.T., Costa, E.A., Jonikas, M., et al., (2018). The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins. *Elife* **7**
166. Chitwood, P.J., Juszkiewicz, S., Guna, A., Shao, S., Hegde, R.S., (2018). EMC is required to initiate accurate membrane protein topogenesis. *Cell* **175**, 1507–1519. e16.
167. Lakshminarayan, R., Phillips, B.P., Binnian, I.L., Gomez-Navarro, N., Escudero-Urquijo, N., Warren, A.J., et al., (2020). Pre-emptive quality control of a misfolded membrane protein by ribosome-driven effects. *Curr. Biol.* **30**, 854–864.e5.
168. Wu, H., Smalinskaitė, L., Hegde, R.S., (2023). EMC rectifies the topology of multipass membrane proteins. *Nat. Struct. Mol. Biol.*.
169. Smalinskaitė, L., Kim, M.K., Lewis, A.J.O., Keenan, R.J., Hegde, R.S., (2022). Mechanism of an intramembrane chaperone for multipass membrane proteins. *Nature* **611**, 161–166.
170. Sundaram, A., Yamsek, M., Zhong, F., Hooda, Y., Hegde, R.S., Keenan, R.J., (2022). Substrate-driven assembly of a translocon for multipass membrane proteins. *Nature* **611**, 167–172.
171. Zhang, R., Miner, J.J., Gorman, M.J., Rausch, K., Ramage, H., White, J.P., et al., (2016). A CRISPR screen defines a signal peptide processing pathway required by flaviviruses. *Nature* **535**, 164–168.
172. Zanotti, A., Coelho, J.P.L., Kaylani, D., Singh, G., Tauber, M., Hitzenberger, M., et al., (2022). The human signal peptidase complex acts as a quality control enzyme for membrane proteins. *Science* **378**, 996–1000.
173. Loo, M.A., Jensen, T.J., Cui, L., Hou, Y., Chang, X.B., Riordan, J.R., (1998). Perturbation of Hsp90 interaction with nascent CFTR prevents its maturation and accelerates its degradation by the proteasome. *EMBO J.* **17**, 6879–6887.
174. Margolese, L., Waneck, G.L., Suzuki, C.K., Degen, E., Flavell, R.A., Williams, D.B., (1993). Identification of the region on the class I histocompatibility molecule that interacts with the molecular chaperone, p88 (calnexin, IP90). *J. Biol. Chem.* **268**, 17959–17966.
175. Swanton, E., High, S., Woodman, P., (2003). Role of calnexin in the glycan-independent quality control of proteolipid protein. *EMBO J.* **22**, 2948–2958.
176. Bloemeke, N., Meighen-Berger, K., Hitzenberger, M., Bach, N.C., Parr, M., Coelho, J.P., et al., (2022). Intramembrane client recognition potentiates the chaperone functions of calnexin. *EMBO J.* **41**, e110959.
177. Pankow, S., Bamberger, C., Yates 3rd., J.R., (2019). A posttranslational modification code for CFTR maturation is altered in cystic fibrosis. *Sci. Signal.* **12**
178. Mornon, J.P., Lehn, P., Callebaut, I., (2009). Molecular models of the open and closed states of the whole human CFTR protein. *Cell. Mol. Life Sci.* **66**, 3469–3486.
179. Middleton, P.G., Mall, M.A., Dřevínek, P., Lands, L.C., McKone, E.F., Polineni, D., et al., (2019). Elexacaftor-tezacaftor-ivacaftor for cystic fibrosis with a single Phe508del allele. *N. Engl. J. Med.* **381**, 1809–1819.
180. Taylor-Cousar, J.L., Munck, A., McKone, E.F., van der Ent, C.K., Moeller, A., Simard, C., et al., (2017). Tezacaftor-ivacaftor in patients with cystic fibrosis homozygous for Phe508del. *N. Engl. J. Med.* **377**, 2013–2023.
181. Boyle, M.P., Bell, S.C., Konstan, M.W., McColley, S.A., Rowe, S.M., Rietschel, E., et al., (2014). A CFTR corrector (lumacaftor) and a CFTR potentiator (ivacaftor) for treatment of patients with cystic fibrosis who have a phe508del CFTR mutation: a phase 2 randomised controlled trial. *Lancet Respir. Med.* **2**, 527–538.
182. Accurso, F.J., Rowe, S.M., Clancy, J.P., Boyle, M.P., Dunitz, J.M., Durie, P.R., et al., (2010). Effect of VX-770 in persons with cystic fibrosis and the G551D-CFTR mutation. *N. Engl. J. Med.* **363**, 1991–2003.