Published in final edited form as:

Nat Struct Mol Biol. 2005 January ; 12(1): 10–16. doi:10.1038/nsmb881.

Side chain and backbone contributions of Phe508 to CFTR folding

Patrick H Thibodeau^{1,3}, Chad A Brautigam², Mischa Machius^{2,3}, and Philip J Thomas^{1,3}

¹Department of Physiology, The University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75390 USA

²Department of Biochemistry, The University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75390 USA

³Molecular Biophysics Graduate Program, The University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75390 USA

Abstract

Mutations in the cystic fibrosis transmembrane conductance regulator (CFTR), an integral membrane protein, cause cystic fibrosis (CF). The most common CF-causing mutant, deletion of Phe508, fails to properly fold. To elucidate the role Phe508 plays in the folding of CFTR, missense mutations at this position were generated. Only one missense mutation had a pronounced effect on the stability and folding of the isolated domain *in vitro*. In contrast, many substitutions, including those of charged and bulky residues, disrupted folding of full-length CFTR in cells. Structures of two mutant nucleotide-binding domains (NBDs) reveal only local alterations of the surface near position 508. These results suggest that the peptide backbone plays a role in the proper folding of the domain, whereas the side chain plays a role in defining a surface of NBD1 that potentially interacts with other domains during the maturation of intact CFTR.

CFTR is a 1,480-residue polytopic membrane protein belonging to the ATP-binding cassette (ABC) superfamily of proteins, and is composed of two transmembrane domains (TMDs), two NBDs and a regulatory domain (R)^{1,2}. Mutations in CFTR give rise to several diseases, including CF, a disease of abnormal ion secretion across epithelia¹. More than 1,000 individual mutations have been identified that give rise to a spectrum of differing disease severities and symptoms (<http://www.genet.sickkids.on.ca/cftr>)³. The most common mutation is the deletion of a phenylalanine at position 508 (Δ F508) in the N-terminal NBD (NBD1)⁴. The deletion of Phe508 in CFTR gives rise to a temperature-sensitive folding defect evidenced by failure of the full-length protein to mature, retention in the endoplasmic reticulum (ER) and subsequent degradation by the proteasome^{5–11}.

© 2005 Nature Publishing Group

Correspondence should be addressed to P.J.T. (Philip.Thomas@UTSouthwestern.edu).

Note added in proof: Crystal structures of the human F508A missense NBD1 (with solubilizing mutations F429S and H667R) and the corrected Δ F508 NBD1 (with three known suppressor mutations G550E, R553Q and R555K, and the solubilizing mutations F409L, F429S, F433L and H667R) have been reported⁵¹. Consistent with the current study, neither of these mutant structures differ substantially from the murine wild-type structure outside of the flexible regulatory regions, nor does the Δ F508 mutation measurably alter the $\Delta G_{\text{unfolding}}$ of the domain. The *in vivo* yield of soluble Δ F508 protein is decreased relative to both the wild-type and F508A proteins with both solubilizing and suppressor mutations, consistent with a decrease in the efficiency of domain folding as described in this study.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

The folding and maturation of $\Delta F508$ CFTR can be rescued by several treatments in cell culture, although the rescued protein has a reduced efficiency of maturation and a reduced half-life at the plasma membrane^{12–17}. The functional activity of the ‘corrected’ protein is also at least partially rescued by these treatments^{12–15,18}, suggesting that the most predominant biophysical manifestations of this mutation affect the protein in a transient manner during biosynthesis. The ability of the mutant protein to be rescued also suggests that the $\Delta F508$ defect is subtle in nature and that a therapeutic strategy to correct this misfolding might potentially be developed that could benefit the vast majority of patients with CF^{19,20}.

To further develop our understanding of the role Phe508 plays in the proper folding and maturation of CFTR, a systematic series of missense mutations was introduced at the 508 locus in full-length CFTR and NBD1 proteins. The effects of these mutations on the folding of the NBD, its stability and the maturation of full-length CFTR were evaluated. Finally, to assess the specific structural consequences of these mutations, the crystal structures of two mutant NBD1 proteins, F508S, a previously identified non-CF-causing mutation (<http://www.genet.sickkids.on.ca/cftr>), and F508R, a maturation-deficient mutation⁵, were determined. Taken together, the results suggest that position 508 contributes directly to the proper folding of NBD1 and also potentially contributes to associations between domains in CFTR, thus affecting multiple steps along the folding pathway. A hierarchical model for the translation, folding and assembly of CFTR is presented based on these data and prior studies of other related ABC-transporter protein systems^{21,22}.

RESULTS

NBD1 folding and stability

Previous studies have indicated that the Phe508 position contributes to the folding pathway of NBD1, but does not markedly alter the native-state stability of the domain. The $\Delta G_{\text{unfolding}}$ of the $\Delta F508$ NBD1 was similar to that of the wild-type NBD1 but the folding efficiency of the $\Delta F508$ protein was reduced^{23,24}. However, the mechanism by which the deletion affects the folding efficiency of the domain is not understood and could be explained by at least three possibilities: (i) the loss of the Phe508 peptide backbone is responsible for the folding defect, (ii) the loss of the Phe508 side chain underlies the folding defect, or (iii) both. To ascertain whether the change in folding efficiency was due to the loss of the peptide backbone or the loss of the side chain, mutations were introduced at position 508 and the temperature-dependence of the refolding reaction was measured in kinetic partitioning experiments. In this assay, the protein is refolded under conditions where both on- and off-pathway folding reactions compete and the relative efficiency of on-pathway folding is reflected in the fraction of the protein that is soluble. Native-state structure and function were subsequently confirmed by ATP binding. All folded, soluble mutant proteins tested retained nucleotide-binding function.

With the exception of F508W, all measured NBD1 proteins were capable of folding at 4 °C at near-100% efficiency as measured by the production of soluble conformers that were quantified by tryptophan fluorescence intensity or western blotting. In addition, all of the domains exhibited a temperature-dependence of refolding efficiency where overall yield in the soluble fraction decreased as temperature increased. Wild-type protein refolded with near-100% efficiency at 4 °C and was capable of refolding at >90% efficiency at 10 °C. Under the conditions used, as temperature increased >10 °C, the amount of soluble wild-type protein decreased, with ~75% of the total protein soluble at 16 °C and <50% soluble at 22 °C (Fig. 1a). The $\Delta F508$ protein also folded with high efficiency at 4 °C, however the temperature-dependence of the refolding reaction was substantially shifted relative to the wild type and the missense mutants. At 10 °C, the $\Delta F508$ NBD1 had substantially reduced

refolding efficiency, with ~55% soluble at 10 °C, 25% soluble at 16 °C and <20% soluble at 22 °C relative to the wild-type soluble fraction at 4 °C (Fig. 1a).

The F508A, F508M, F508P, F508D, F508Q, F508R and F508S mutant proteins were more similar to the wild type than the Δ F508 protein in their temperature-dependence of refolding (Fig. 1b,c). All of these proteins were capable of refolding at near-100% efficiency at 4 °C and refolded at >75% efficiency at 10 °C. At >10 °C, these mutant proteins refolded with reduced efficiency, which decreased as temperature increased.

A tryptophan was also introduced at position 508 to assess the effects of substitution of a larger hydrophobic residue and to act as a spectral probe to track the folding of the NBD both in the wild-type domain and in a mutant background of W496F. In contrast to all of the other substitutions at the 508 locus, the tryptophan substitution folded poorly at all temperatures, with maximal refolding efficiency being ~35% of the wild-type protein at 4 °C. However, when the F508W mutation was introduced onto the background of W496F, folding of the NBD1 was partially restored and the protein was capable of refolding with higher efficiency (>90% soluble) at 4 °C (Fig. 1b).

The kinetic partitioning experiments address the potential impact that mutations exert on the process of protein folding (the transitions from the denatured to the native state and/or to off-pathway states) and their effects on the native state of the protein (changes in native state stability and/or solubility). To determine whether the reduced folding efficiency of Δ F508 was caused by a reduction of stability or alteration of the folding pathway, native state stabilities were directly determined by extrapolation of the guanidinium hydrochloride (GuHCl)-dependence of unfolding to zero denaturant. As previously reported, the NBD1 wild-type and Δ F508 proteins exhibited nearly identical thermodynamic stabilities as measured by denaturation experiments²³. The $\Delta G_{\text{unfolding}}$ for the wild-type isolated NBD1 was 3.7 kcal mol⁻¹ and the m -value, a measure of the cooperativity of unfolding, directly related to the change in surface area exposure²⁵, was 1.7 kcal mol⁻¹ M⁻¹denaturant (Table 1), reasonable for an isolated domain from a multidomain protein. Similar values were obtained for the Δ F508 NBD1 protein, with a $\Delta G_{\text{unfolding}}$ of 3.6 kcal mol⁻¹ and an m -value of 1.7 kcal mol⁻¹ M⁻¹ denaturant. The missense mutant proteins F508A, F508M, F508P, F508D, F508Q, F508R and F508S had similar $\Delta G_{\text{unfolding}}$ and m -values, 3.4–3.8 kcal mol⁻¹ and 1.5–1.7 kcal mol⁻¹ M⁻¹ denaturant, respectively, highlighting the fact that changes in the bulk or chemical properties of the substituted side chain had little effect on the native-state stabilities of these domains as measured by denaturation with GuHCl (Table 1).

Although the native-state stabilities and folding efficiencies of the substitution mutants were markedly similar to those of the wild-type protein, it is possible that substantial structural perturbation of NBD1 may occur in order to accommodate the changes in side chain character at the 508 locus. Although these changes may not affect the measured folding efficiencies and stabilities, they may directly impact the ability of the NBD to interact with other domains in CFTR *in cis* or with other proteins *in trans*. How does the isolated NBD accommodate such considerable changes in amino acid character at position 508 when this position is critical to the proper biogenesis of the full-length protein, and what are the underlying structural changes associated with these substitutions?

NBD1 structure

To directly assess the structural changes associated with the substitution of residues at position 508, crystal structures of two missense-mutant proteins were determined for the highly similar murine NBD1: F508S, a previously identified non-CF-causing variant, and F508R, a previously described maturation-deficient mutation. The proteins were expressed and purified essentially as described for the wild-type protein and crystallized under

conditions similar to the wild-type protein in the presence of Mg^{2+} and ATP with sodium acetate as the precipitant²⁶. Tetragonal bipyramidal crystals grew for the F508R proteins, whereas the F508S protein spontaneously crystallized as large tetragonal plates. The F508S and F508R crystals diffracted to 2.7 and 3.1 Å, respectively and structures were determined with final R/R_{free} values of 0.207/0.262 and 0.254/0.266, respectively (Table 2).

As might be predicted, based on the biochemical experiments described above, and consistent with the relatively exposed side chain of Phe508 in the previously determined murine wild-type NBD1 crystal structures²⁶, the substitution of both serine and arginine had little effect on the overall structure of NBD1 when compared with that of the wild-type protein (Fig. 2a and Supplementary Fig. 1 online). The organization of the fold is the same for the wild type and both missense mutant proteins, with three subdomains: a β -subdomain, a mixed α/β -core domain and an α -helical subdomain. All of the major structural elements are conserved and the r.m.s. deviations between the C α atoms of the wild type and F508S or F508R structures were <0.33 Å (Fig. 2a)²⁶.

Although the topology and overall fold of the mutant proteins are nearly identical to those of the wild-type proteins, there are several noteworthy differences in the structures of these proteins. Small, local structural changes surrounding the Phe508 locus are evident in the missense mutant protein structures. The aromatic side chain of Phe508 is largely surface-exposed and accessible in all of the wild-type structures and is in close proximity to Trp496 and Met498, both located in the Q loop or γ -phosphate switch^{26,27}. The substitution of the smaller serine for the phenylalanine induces a rotation of Met498 into a region that is occupied by the Phe508 side chain in the wild-type structure (Fig. 2b). This rotation is observed in only one of two monomers in the asymmetric unit when the phenylalanine is replaced by the larger arginine side chain in the F508R structure (Fig. 2c). In both mutants, like the wild type, the side chain of the residue at position 508 is largely surface-exposed and accessible (Fig. 2d).

ATP is bound in both of the mutant protein structures. As with the wild-type protein, no products of ATP hydrolysis were observed in the electron density maps derived from diffraction data from the mutant crystals (Supplementary Fig. 1 online). This is despite the fact that the protein was incubated with ATP and Mg^{2+} for several days during the course of crystallization. The conformation of ATP in the wild-type²⁶ and F508S structures is very similar, with a noncanonical interaction between the NBD and ATP and unusual torsional angles in the ATP molecule (Supplementary Fig. 1 online). In the F508R structure, the two monomers that occur in the asymmetric unit contain ATP in different conformations. In one monomer, the ATP is bound in an orientation similar to that observed in the wild-type and F508S structures. The ATP bound to the other monomer, however, adopts a more conventional orientation (Supplementary Fig. 1 online). In both monomers, the interactions stabilizing the conformation of ATP are influenced by nearby crystal contacts. Specifically, the adenine base of the ATP in the more canonical conformation interacts with the guanidine moiety of Arg508 in the opposing monomer in the asymmetric unit of the crystals.

Examination of the calculated molecular surface of the wild-type, F508S and F508R proteins is revealing. Phe508 contributes to a largely hydrophobic region of the surface²⁶ that presumably contributes to the domain-domain interface between the NBDs and TMDs (circled, Fig. 2d)^{21,22}. Although the serine substitution does not markedly affect the calculated electrostatic molecular surface locally or globally, the surface-exposed arginine side chain exhibits substantial changes owing to the surface-exposed, basic guanidine group. In addition, the physical contours of the NBD1 protein surface are also affected by the substitutions of serine and arginine. Quantification of the surface-accessibility of position 508 reveals that the wild-type and F508S side chains are very similar at 8.5 and 9.6 Å²

respectively. The F508R protein has greater average accessible surface area at position 508, with a value of 16.8 Å².

CFTR folding

The effects of the majority of these mutations on the folding and assembly of CFTR were unknown. To evaluate the ability of the mutant proteins to fold in the context of the full-length CFTR protein, each of the mutations were introduced into a pCMV-CFTR expression vector for expression in HEK 293 cells. All of the mutant proteins expressed in HEK 293 cells as evidenced by the presence of band B, the core glycosylated CFTR conformer, which has not yet reached the Golgi (Fig. 3). Levels of the fully glycosylated band C protein, which is indicative of the post-ER trafficking of CFTR, were markedly reduced in a substantial number of the substitution mutants even though the isolated domain was capable of refolding with little regard for the specific amino acid substitution at position 508 (Fig. 3).

The steady-state band C levels of F508C and F508M were reduced, but closest to those of wild type. Band C levels in F508A, F508G, F508L and F508V as well as the polar amino acid substitutions F508S, F508T, F508N and F508Q were evident, but substantially reduced relative to wild-type band C levels. The known polymorphism F508C and the non-CF-causing variant F508S both showed measurable quantities of band C at steady-state levels, as would be expected for non-CF-causing substitutions. The hydrophobic amino acid substitutions F508I, F508W and F508Y did not produce substantial steady-state levels of band C as measured by western blotting, nor did the ionizable amino acid substitutions F508D, F508E, F508K, F508H or F508R. No band C was seen for the F508P substitution. No endogenous CFTR was detected in HEK 293 cells transfected with a pCMV-GFP expression plasmid.

DISCUSSION

The biosynthesis and maturation of multidomain proteins are complex biological processes. Their folding and maturation could occur by several mechanisms: a purely cotranslational, sequential folding process²⁸, a purely post-translational, global folding process²⁹, or some combination of these two extremes. The cotranslational, sequential model would require individual domains to fold and subsequently assemble to form a functional native state, whereas a global collapse model would require that the protein be held in a folding-competent state until the termination of translation, allowing for all of the protein sequence to be simultaneously exposed and thereby initiating cooperative folding of the protein. The fact that NBD1 can fold as an autonomous unit suggests that cotranslational folding of CFTR is plausible. Data indicate that NBD1 assumes a proteolytically resistant, compact structure early during full-length CFTR biosynthesis in cells³⁰, consistent with the cotranslational folding of the NBD. Given that many homologous microbial ABC transporter systems produce the individual functional units (nucleotide binding and hydrolysis and substrate translocation) on separate polypeptide chains with high sequence similarity, it is likely that CFTR is not unique in having individual domains with the necessary information and at least some capacity to achieve a folded, functional state.

The *in vitro* folding and stability data on the NBD1 of CFTR indicate, as in previous studies, that the deletion of the phenylalanine at position 508 directly impacts the efficiency by which the NBD achieves its native state, while not substantially affecting the stability of the protein that does reach the native state^{23,24}. Notably, missense mutations other than F508W had little effect on either the folding or stability of the isolated NBD, suggesting that the peptide backbone at this locus is critical to NBD1 folding efficiency, whereas the side chain character is largely unimportant. This is also consistent with the Phe508 position having substantial surface exposure as is seen in the murine NBD1 structures presented here and

elsewhere²⁶, as well as a growing number of structures from homologous proteins in which the Phe508-analogous residues are highly surface-accessible and exposed^{21,31,32}. Most notably, the crystal structures of F508S and F508R both indicate that substitutions for Phe508 do not substantially impact the structure of NBD1, providing further evidence for the high tolerance for substitution at this position in the isolated domain.

The structures of NBD1 proteins also suggest a potential mechanism for the deleterious effects of the F508W substitution, as the phenylalanine side chain, although partially surface-exposed and accessible, interacts with surrounding residues. The nearest atom distances from both Trp496 and Met498 to Phe508 are ~4 Å. The additional physical size of the tryptophan side chain thus may not be accommodated by the local protein structure. However, when a second substitution, W496F, was introduced, the folding of the domain was rescued. Given the close proximity of both residues, the W496F substitution probably resolves a steric clash between the substituted tryptophan at position 508 and other local residues, consistent with the refolded protein reaching a native or near-native-state structure *in vitro*.

Full-length CFTR was considerably more sensitive to substitutions at position 508, failing to traffic when all but small hydrophobic or polar residues were introduced. Potential explanations for the increased sensitivity in full-length CFTR come from genetic and biochemical studies of other ABC transporters^{33–36} and the two ABC transport systems whose TMDs and NBDs have been studied structurally, BtuCD and MsbA. In both of these structures the residue analogous to Phe508 lies at the interface between the NBDs and the TMDs and contributes to the interaction between these domains (Fig. 4a)^{21,22}. Changes in the amino acid character at this position may therefore directly impact the ability of these domains to assemble into the functional transport unit. This model is also consistent with previous functional and biochemical studies on bacterial transporter systems that map the interactions between the α -helical subdomain of the NBDs to regions in the cytoplasmic loops of the TMDs^{35,36}.

Similarly, NBD-TMD interactions have been reported to occur during CFTR biosynthesis. Biochemical experiments³⁷ indicated that although CFTR TMD1 could integrate into the membrane in the absence of other CFTR domains, TMD1 stability was marginal, allowing for the movement of transmembrane spans into and out of the membrane. The presence of NBD1, however, stabilized the integration of TMD1 in the membrane. Additional recent studies with CFTR and the homologous ABC transporter protein P-gp suggest that the proper folding and docking of NBD1 may induce changes in the transmembrane domains that are required for the proper maturation of these proteins and provide a mechanism by which mutations in cytosolic domains alter protein-chaperone interactions in the lumen of the ER^{38,39}. These findings suggest that NBD1 and the TMDs are capable of interacting and that binding of NBD1 with protein components in TMD1 may induce structural changes that stabilize the integration and conformation of transmembrane spans^{37,38}.

Coupled with the biophysical and cell biological studies presented here, the structures of the mutant NBD1s suggest the role Phe508 plays in CFTR folding and how mutations at this position can affect the folding and assembly of CFTR. In a hierarchical model for the biosynthesis of CFTR (Fig. 4b), the translation and integration of TMD1 is followed by the translation and folding of NBD1, consistent with the evidence for early cotranslational folding³⁰ and previous models^{40,41}. The folded NBD1 then docks with the integrated TMD and both the NBD and the TMD are stabilized by this interaction. Subsequent steps, not evaluated in this study, would then allow for the translation, folding and assembly of domains C-terminal to NBD1, including the translation, folding and assembly of TMD2 and NBD2, which may also be affected by the misfolding of NBD1 (ref. 30).

Which steps in the biosynthesis of CFTR are affected by mutations at position 508? This study indicates that at least two steps in the CFTR folding pathway are affected by position 508. The initial defect associated with the $\Delta F508$ mutation is probably the folding of NBD1, as suggested by the reduced folding efficiency of the isolated $\Delta F508$ NBD. Consistent with the *in vitro* results, the *in vivo* production of soluble murine $\Delta F508$ NBD is reduced several-fold relative to wild-type levels under identical expression conditions, even though the soluble, folded fractions of both wild-type and mutant proteins show similar behavior with respect to purification and solubility characteristics (data not shown). However, changes in the local surface character of the NBD may also contribute to the efficiency of subsequent steps in CFTR biogenesis, such as tertiary assembly of the individual domains. In this regard, previous studies have demonstrated that $\Delta F508$ CFTR is cotranslationally ubiquitinated⁴², suggesting that early missteps, before the completion of translation, may be sensed as aberrant and thus targeted for degradation. Such a mechanism is consistent either with recognition of the disruption of NBD1 folding, as seen in the *in vitro* experiments, or with a disruption of later steps of domain assembly. The relatively low tolerance for substitutions in full-length CFTR indicates that although loss of the peptide backbone is important, the contributions of the side chain to subsequent biogenic assembly steps, such as the assembly of NBD1 with other CFTR domains, are also key to recognition by the quality control machinery.

At present, only symptomatic treatments have been developed for CF. Therapeutic strategies that target the underlying cause of the disease—that is, those directed at correction of $\Delta F508$ CFTR folding—may need to address the effects of the mutation at multiple points in the life cycle of CFTR. A small molecule ligand that promotes formation of the native state of NBD1 may also need to facilitate interactions between domains in CFTR to fully rescue the $\Delta F508$ folding defect. Moreover, the stabilization of the appropriate domain-domain interface also is likely to be functionally important, as the binding and hydrolysis of ATP in the cytosolic NBDs are coupled to the regulation of activity of the TMDs^{21,22}. The use of ligand-binding energy to promote an unfavorable folding reaction has been achieved, so far, only for mutations that destabilize the native state^{43,44}. Because the $\Delta F508$ mutation exerts its effects during the process of folding, use of this paradigm for the development of CF therapeutics is a more difficult, and as yet unmet, challenge.

The studies presented here provide information regarding the mechanisms multidomain membrane proteins use to achieve a folded, functional state. These data are consistent with a model for the hierarchical folding of CFTR, whereby the formation of individual domains precedes the final assembly of the multidomain protein complex. Careful study of the cotranslational nature of folding and assembly may facilitate a better understanding of the mechanisms by which these complicated proteins fold and are recognized by the cellular quality control machinery, and the steps that need to be corrected for effective therapeutic intervention to alleviate disease when the process goes awry.

METHODS

In vivo maturation assays

Expression plasmids of full-length, wild-type and $\Delta F508$ CFTR (pCMV-CFTR-Not6.2) as well as pCMV-GFP were a gift from J. Rommens (The Hospital for Sick Children, Toronto) and were mutagenized using the QuikChange site-directed mutagenesis kit (Stratagene). DNA (1 μ g) was transfected into HEK 293 cells (American Type Culture Collection) using the Fugene-6 transfection reagent (Roche) following manufacturer's protocols. The cells were harvested 48–72 h after transfection, washed with PBS, and lysed in 0.5 ml RIPA (20 mM Tris-HCl, 150 mM NaCl, 0.1% (w/v) SDS, 0.5% (w/v) deoxycholate, 1.0% (v/v) IGEPAL CA-630, Complete protease inhibitor tablets (Roche), pH 8.0). The cell lysates

were cleared by centrifugation and aliquots of the supernatant were subsequently electrophoresed on a 6% (w/v) Tris-glycine gel, and transferred for western blotting with anti-CFTR M3A7 (Upstate). Cells were routinely maintained in DMEM (Invitrogen) supplemented with 10% (v/v) ^{ca}If serum, 50 $\mu\text{g ml}^{-1}$ penicillin, and 50 units ml^{-1} streptomycin using standard culture techniques.

Human NBD1 expression and purification

The NBD1 coding sequences, spanning residues 389–655, were amplified from the pCMV-CFTR constructs described above, cloned into the pET28a expression vector (Novagen), and transformed into BL21 (DE3) bacteria for protein expression and purification as described²⁴. The protein was purified using His-Bind resin (Novagen) following the manufacturer's protocols for purification under denaturing conditions using GuHCl. The purified NBD1 protein was precipitated by dialysis (100 mM Tris and 2 mM EDTA, pH 8.0) and stored at -20°C .

NBD1 refolding

The purified NBD1 proteins were refolded essentially as described²³. The proteins were refolded by rapid dilution to between 1 and 20 μM final concentration into refolding buffer (R-buffer: 375 mM L-arginine, 200 mM GuHCl, 100 mM Tris, 2 mM EDTA, 1 mM DTT, pH 8.0), vortexed briefly, and then incubated overnight at the desired temperature. Folding was monitored by a blue shift and increase in intensity of tryptophan fluorescence emission from 340–350 nm to 325–330 nm, when excited at 280 nm. The folding yield experiments were completed at 1 μM final protein concentration for all of the NBD1 proteins measured and quantified by tryptophan fluorescence and western blotting.

In vitro stability

The GuHCl-induced denaturation of each of the purified, refolded proteins was completed to assess the relative thermodynamic stabilities of each of the NBD1 proteins. Emission spectra of tryptophan fluorescence were collected and corrected for both pre- and post-transition slopes and the transition region was used to calculate the equilibrium constant (K_{eq}) for the unfolding reaction across the transition region. Linear regression was carried out to extrapolate the $\Delta G_{\text{unfolding}}$ at zero denaturant and the m -values⁴⁵.

Murine NBD1 expression, purification and crystallization

Murine cDNA, a gift from S. Muallem (University of Texas Southwestern Medical Center at Dallas), was used as a template to amplify NBD1, residues 389–673, which was subsequently cloned into the pSmt3 expression vector⁴⁶, a gift from C. Lima (Cornell University, New York). The Smt3-NBD1 fusion proteins were expressed in BL21 (DE3) codon-plus cells. Cultures were grown, after inoculation, to an A_{600} of 1.5–2.0 at 37 $^{\circ}\text{C}$, shifted to 15 $^{\circ}\text{C}$, induced with 750 μM IPTG and allowed to express for 20 h. The cells were harvested and lysed by sonication (50 mM Tris, 150 mM NaCl, 100 mM L-arginine, 5 mM MgCl_2 , 2 mM ATP, 1 mM β -mercaptoethanol, 12.5% (v/v) glycerol, and 0.2% (v/v) IGEPAL CA-630, pH 7.6). Purification of soluble mNBD1 was completed essentially as described using standard nickel-affinity and size-exclusion chromatography techniques²⁶. The purified murine NBD1 proteins were crystallized using the hanging-drop method against a well solution of 2.5–4.0 M sodium acetate, pH 7.5 at 5 $^{\circ}\text{C}$. The crystals were transferred directly from the mother liquor into liquid propane and stored in liquid nitrogen.

Structure determination

X-ray diffraction data from crystals of the mNBD1 mutants were collected at the Structural Biology Center of the Advanced Photon Source of the Argonne National Laboratory. The

data were indexed, integrated and scaled using HKL2000 (ref. 47). Table 2 shows the statistics for the diffraction data. The structure of F508R was determined using the molecular replacement protocols available in CNS version 1.1 (ref. 48). The search model was a single monomer from the structure of mNBD1 (PDB entry 1R0X)²⁶ stripped of water molecules, ions and ATP. The structure of F508S was determined using the mNBD1 structure (PDB entry 1R0W)²⁶ stripped of water molecules, ions and ATP as the starting model. Subsequent rigid-body refinement provided satisfactory starting models for both structures. The models were refined using the simulated annealing, conjugate gradient minimization, grouped *B*-factor, and individual *B*-factor refinement protocols available in CNS version 1.1 (ref. 48). The model statistics are given in Table 2. Structural figures were generated using PyMOL (<http://www.pymol.org>), XtalView⁴⁹ and GRASP⁵⁰ and rendered in POV-Ray (<http://www.povray.org>).

Coordinates

The atomic coordinates and structure factors for the F508S and F508R NBD1 structures have been deposited in the Protein Data Bank (accession codes 1XF9 and 1XFA, respectively).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank H. Lewis and M. Kearins for helpful advice regarding production of the two NBD1 crystal forms and members of the Thomas lab and the Structural Biology Lab for helpful suggestions and constructive criticism. This work was supported by grants from the US National Institutes of Health (NIH) (DK49835), the Cystic Fibrosis Foundation (THOMAS01GO) and Welch Foundation (I-1284) to P.J.T. and by a position on the NIH Training Grant, Mechanisms of Drug Action and Disposition (GM07062), awarded to P.H.T.

References

1. Sheppard DN, Welsh MJ. Structure and function of the CFTR chloride channel. *Physiol Rev.* 1999; 79:S23–S45. [PubMed: 9922375]
2. Riordan JR, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science.* 1989; 245:1066–1073. [PubMed: 2475911]
3. Tsui LC. The spectrum of cystic fibrosis mutations. *Trends Genet.* 1992; 8:392–398. [PubMed: 1279852]
4. Tsui LC. The cystic fibrosis transmembrane conductance regulator gene. *Am J Respir Crit Care Med.* 1995; 151:S47–S53. [PubMed: 7533605]
5. Cheng SH, et al. Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell.* 1990; 63:827–834. [PubMed: 1699669]
6. Lukacs GL, et al. Conformational maturation of CFTR but not its mutant counterpart (Δ F508) occurs in the endoplasmic reticulum and requires ATP. *EMBO J.* 1994; 13:6076–6086. [PubMed: 7529176]
7. Denning GM, Ostedgaard LS, Welsh MJ. Abnormal localization of cystic fibrosis transmembrane conductance regulator in primary cultures of cystic fibrosis airway epithelia. *J Cell Biol.* 1992; 118:551–559. [PubMed: 1379244]
8. Sharma M, Benharouga M, Hu W, Lukacs GL. Conformational and temperature-sensitive stability defects of the Δ F508 cystic fibrosis transmembrane conductance regulator in post-endoplasmic reticulum compartments. *J Biol Chem.* 2001; 276:8942–8950. [PubMed: 11124952]
9. Ward CL, Omura S, Kopito RR. Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell.* 1995; 83:121–127. [PubMed: 7553863]

10. Jensen TJ, et al. Multiple proteolytic systems, including the proteasome, contribute to CFTR processing. *Cell*. 1995; 83:129–135. [PubMed: 7553864]
11. Kartner N, Augustinas O, Jensen TJ, Naismith AL, Riordan JR. Mislocalization of Δ F508 CFTR in cystic fibrosis sweat gland. *Nat Genet*. 1992; 1:321–327. [PubMed: 1284548]
12. Zhang XM, et al. Organic solutes rescue the functional defect in F508 CFTR. *J Biol Chem*. 2003; 278:51232–51242. [PubMed: 14532265]
13. Denning GM, et al. Processing of mutant cystic fibrosis transmembrane conductance regulator is temperature-sensitive. *Nature*. 1992; 358:761–764. [PubMed: 1380673]
14. Sato S, Ward CL, Krouse ME, Wine JJ, Kopito RR. Glycerol reverses the misfolding phenotype of the most common cystic fibrosis mutation. *J Biol Chem*. 1996; 271:635–638. [PubMed: 8557666]
15. Yang H, et al. Nanomolar affinity small molecule correctors of defective Δ F508-CFTR chloride channel gating. *J Biol Chem*. 2003; 278:35079–35085. [PubMed: 12832418]
16. Brown CR, Hong-Brown LQ, Biwersi J, Verkman AS, Welch WJ. Chemical chaperones correct the mutant phenotype of the Δ F508 cystic fibrosis transmembrane conductance regulator protein. *Cell Stress Chaperones*. 1996; 1:117–125. [PubMed: 9222597]
17. Dormer RL, et al. Correction of Δ F508-CFTR activity with benzo(c)quinolizinium compounds through facilitation of its processing in cystic fibrosis airway cells. *J Cell Sci*. 2001; 114:4073–4081. [PubMed: 11739639]
18. Howard M, et al. Mammalian osmolytes and *S*-nitrosoglutathione promote Δ F508 cystic fibrosis transmembrane conductance regulator (CFTR) protein maturation and function. *J Biol Chem*. 2003; 278:35159–35167. [PubMed: 12837761]
19. Thomas PJ, Ko YH, Pedersen PL. Altered protein folding may be the molecular basis of most cases of cystic fibrosis. *FEBS Lett*. 1992; 312:7–9. [PubMed: 1385213]
20. Zeitlin PL. Therapies directed at the basic defect in cystic fibrosis. *Clin Chest Med*. 1998; 19:515–525. [PubMed: 9759553]
21. Locher KP, Lee AT, Rees DC. The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science*. 2002; 296:1091–1098. [PubMed: 12004122]
22. Chang G. Structure of MsbA from *Vibrio cholera*: a multidrug resistance ABC transporter homolog in a closed conformation. *J Mol Biol*. 2003; 330:419–430. [PubMed: 12823979]
23. Qu BH, Strickland EH, Thomas PJ. Localization and suppression of a kinetic defect in cystic fibrosis transmembrane conductance regulator folding. *J Biol Chem*. 1997; 272:15739–15744. [PubMed: 9188468]
24. Qu BH, Thomas PJ. Alteration of the cystic fibrosis transmembrane conductance regulator folding pathway. *J Biol Chem*. 1996; 271:7261–7264. [PubMed: 8631737]
25. Myers JK, Pace CN, Scholtz JM. Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci*. 1995; 4:2138–2148. [PubMed: 8535251]
26. Lewis HA, et al. Structure of nucleotide-binding domain 1 of the cystic fibrosis transmembrane conductance regulator. *EMBO J*. 2004; 23:282–293. [PubMed: 14685259]
27. Karpowich N, et al. Crystal structures of the MJ1267 ATP binding cassette reveal an induced-fit effect at the ATPase active site of an ABC transporter. *Structure*. 2001; 9:571–586. [PubMed: 11470432]
28. Chen W, Helenius J, Braakman I, Helenius A. Cotranslational folding and calnexin binding during glycoprotein synthesis. *Proc Natl Acad Sci USA*. 1995; 92:6229–6233. [PubMed: 7541532]
29. Jansens A, van Duijn E, Braakman I. Coordinated nonvectorial folding in a newly synthesized multidomain protein. *Science*. 2002; 298:2401–2403. [PubMed: 12493918]
30. Du K, Sharma L, Lukacs GL. The Δ F508 cystic fibrosis mutation impairs domain-domain interactions and arrests post-translational folding of CFTR. *Nat Struct Mol Biol*. 2005; 12:17–25. [PubMed: 15619635]
31. Yuan YR, et al. The crystal structure of the MJ0796 ATP-binding cassette. Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter. *J Biol Chem*. 2001; 276:32313–32321. [PubMed: 11402022]

32. Gaudet R, Wiley DC. Structure of the ABC ATPase domain of human TAP1, the transporter associated with antigen processing. *EMBO J.* 2001; 20:4964–4972. [PubMed: 11532960]
33. Ames GF, et al. Purification and characterization of the membrane-bound complex of an ABC transporter, the histidine permease. *J Bioenerg Biomembr.* 2001; 33:79–92. [PubMed: 11456221]
34. Liu PQ, Ames GF. *In vitro* disassembly and reassembly of an ABC transporter, the histidine permease. *Proc Natl Acad Sci USA.* 1998; 95:3495–3500. [PubMed: 9520394]
35. Mourez M, Hofnung M, Dassa E. Subunit interactions in ABC transporters: a conserved sequence in hydrophobic membrane proteins of periplasmic permeases defines an important site of interaction with the ATPase subunits. *EMBO J.* 1997; 16:3066–3077. [PubMed: 9214624]
36. Wilken S, Schmees G, Schneider E. A putative helical domain in the MalK subunit of the ATP-binding-cassette transport system for maltose of *Salmonella typhimurium* (MalFGK2) is crucial for interaction with MalF and MalG. A study using the LacK protein of *Agrobacterium radiobacter* as a tool. *Mol Microbiol.* 1996; 22:655–666. [PubMed: 8951813]
37. Tector M, Hartl FU. An unstable transmembrane segment in the cystic fibrosis transmembrane conductance regulator. *EMBO J.* 1999; 18:6290–6298. [PubMed: 10562541]
38. Chen EY, Bartlett MC, Loo TW, Clarke DM. The $\Delta F508$ mutation disrupts packing of the transmembrane segments of the cystic fibrosis transmembrane conductance regulator. *J Biol Chem.* 2004; 279:39620–39627. [PubMed: 15272010]
39. Loo TW, Bartlett MC, Clarke DM. Introduction of the most common cystic fibrosis mutation ($\Delta F508$) into human P-glycoprotein disrupts packing of the transmembrane segments. *J Biol Chem.* 2002; 277:27585–27588. [PubMed: 12070134]
40. Meacham GC, et al. The Hdj-2/Hsc70 chaperone pair facilitates early steps in CFTR biogenesis. *EMBO J.* 1999; 18:1492–1505. [PubMed: 10075921]
41. Xiong X, Bragin A, Widdicombe JH, Cohn J, Skach WR. Structural cues involved in endoplasmic reticulum degradation of G85E and G91R mutant cystic fibrosis transmembrane conductance regulator. *J Clin Invest.* 1997; 100:1079–1088. [PubMed: 9276724]
42. Sato S, Ward CL, Kopito RR. Cotranslational ubiquitination of cystic fibrosis transmembrane conductance regulator *in vitro*. *J Biol Chem.* 1998; 273:7189–7192. [PubMed: 9516408]
43. Miller SR, Sekijima Y, Kelly JW. Native state stabilization by NSAIDs inhibits transthyretin amyloidogenesis from the most common familial disease variants. *Lab Invest.* 2004; 84:545–552. [PubMed: 14968122]
44. Fan JQ, Ishii S, Asano N, Suzuki Y. Accelerated transport and maturation of lysosomal α -galactosidase A in Fabry lymphoblasts by an enzyme inhibitor. *Nat Med.* 1999; 5:112–115. [PubMed: 9883849]
45. Pace CN, Shaw KL. Linear extrapolation method of analyzing solvent denaturation curves. *Proteins.* 2000; (Suppl 4):1–7. [PubMed: 11013396]
46. Mossessova E, Lima CD. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell.* 2000; 5:865–876. [PubMed: 10882122]
47. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 1997; 276:307–326.
48. Brunger AT, et al. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D.* 1998; 54:905–921. [PubMed: 9757107]
49. McRee DE. A visual protein crystallographic software system for X11/Xview. *J Mol Graph.* 1992; 10:44–46.
50. Nicholls A, Sharp KA, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins.* 1991; 11:281–296. [PubMed: 1758883]
51. Lewis HA, et al. Impact of the $\Delta 508$ mutation in NBD1 of human CFTR on domain folding and structure. *J Biol Chem.* Nov 3.2004 Epub ahead of print. 10.1074/jbc.M410968200

Figure 1.

NBD1 folding efficiency as a function of folding temperature. NBD1 proteins were refolded at multiple temperatures and assessed for their ability to adopt native, soluble structures by tryptophan fluorescence. **(a)** The wild-type and Δ F508 proteins refolded with different efficiencies as temperature increased >4 °C. The Δ F508 was unable to refold with high efficiency at intermediate and high temperatures, whereas the wild type maintained a higher folding efficiency at these elevated temperatures. Substitution mutations at position 508 had little effect on the folding efficiency of NBD1 *in vitro*. **(b,c)** The hydrophobic substitutions **(b)** and the charged and polar substitutions **(c)** are superimposed on the wild-type and Δ F508 protein folding efficiencies. **(b)** The F508W mutant, the only mutant that deviated markedly from the wild type, was rescued by the introduction of a second missense mutation, W496F.

Figure 2.

Structure of NBD1 proteins. **(a)** A smoothed trace of the main chain of the three proteins. Wild type, green; F508S variant, orange; F508R variant, blue. The r.m.s. deviations between the wild-type and F508S or F508R structures are ~ 0.3 Å. The side chain of Phe508 from the wild-type structure is red. **(b,c)** Electron density maps for the region near the 508 locus. Carbon atoms, green; oxygen, red; nitrogen, blue; sulfur, yellow; phosphorus, pink. The positions of the carbon atoms of Met496 **(b)** and Phe508 **(b,c)** from the wild-type structure are shown in orange for reference. The $2F_o - F_c$ electron density map (contoured at 1σ) calculated with the F508S data at a resolution of 2.7 Å superposed on the final F508S model. **(c)** The $2F_o - F_c$ map (contoured at 1σ) calculated from the F508R data at a resolution of 3.1 Å superposed on the final F508R model. **(d)** Molecular surfaces of the NBD1 proteins. The position of the 508 residue is circled for reference. Red, regions of negative electrostatic potential; blue, positive regions (range from -10 to $+10$ kT).

Figure 3. Maturation of full-length CFTR mutants. Maturation of full-length CFTR was monitored as a function of the formation of the upper molecular mass band, band C, indicative of post-ER trafficking. Band B, a marker for core-glycosylated ER resident protein is indicative of protein expression, but not of proper folding, assembly and post-ER trafficking. Closed arrowheads, CF-causing mutations; open arrowheads, known non-CF-causing variants. (F508S has been associated with congenital bilateral absence of (the vas deferens, but not CF.)

Figure 4.

ABC transporter structure and CFTR biogenesis. (a) Crystal structure of the BtuCD ABC-transport systems²¹. The *Escherichia coli* vitamin D transporter system BtuCD structures (PDB entry 1L7V) are shown with the Phe508-analogous residue Leu96 shown in red spheres at the NBD-TMD interfaces. The BtuC transmembrane proteins are blue and the BtuD NBDs are yellow. Two views of the BtuCD complex are shown rotated about the vertical axis by $\sim 90^\circ$. (b) Hierarchical folding of CFTR. Step 1, TMD1 is translated and inserted into the membrane. Pale blue indicates the reduced stability of TMD1 in the absence of NBD1. Step 2, NBD1 is translated and folds into a native or near-native state. The blurred image of the mNBD1 structure indicates the attainment of a native or near-native state, which is most likely stabilized by interactions with additional CFTR domains. Step 3, NBD1 docks against TMD1. This event probably leads to the stabilization of both NBD1 and TMD1, as shown by the change in blue color in the TMD and the sharpening of the NBD1 structure. This is followed by the translation, folding and assembly of the domains C-terminal to NBD1. Mutations that putatively affect each step are in parentheses. The NBDs are represented by the mNBD1 structure and are oriented relative to the NBD

dimer and TMD–NBD complex seen in BtuCD with the assumption that CFTR is monomeric with a functional NBD1–NBD2 heterodimer.

Table 1

Stability of wild-type and mutant NBD proteins

Protein	$\Delta G_{\text{unfolding}}$ (kcal mol ⁻¹)	$\Delta\Delta G_{\text{unfolding}}$ (kcal mol ⁻¹)	m -value (kcal mol ⁻¹ M ⁻¹)
Wild type	3.7 ± 0.1	0	1.7
ΔF508	3.6 ± 0.1	0.1	1.7
F508A	3.6 ± 0.2	0.1	1.6
F508M	3.5 ± 0.1	0.1	1.6
F508P	3.5 ± 0.3	0.2	1.6
F508D	3.6 ± 0.1	0.1	1.6
F508Q	3.5 ± 0.2	0.2	1.6
F508R	3.4 ± 0.3	0.3	1.6
F508S	3.8 ± 0.2	-0.1	1.6

Table 2

Data collection and refinement statistics

	F508S	F508R
Data collection		
Space group	<i>P</i> 4 ₂ 2	<i>I</i> 4 ₁ 22
Cell dimensions (Å)		
<i>a</i>	170.45	139.99
<i>b</i>	170.45	139.99
<i>c</i>	109.07	278.72
Resolution (Å)	40.4–2.7 (2.75–2.7)	44.3–3.1 (3.15–3.1)
<i>R</i> _{sym}	0.078 (0.431)	0.071 (0.987)
<i>I</i> / <i>σI</i>	22.9 (2.3)	34.8 (2.5)
Completeness (%)	98.8 (95.6)	99.9 (100)
Redundancy	7.7 (3.3)	11.9 (10.7)
Refinement		
Resolution (Å)	40.4–2.7	44.3–3.1
No. reflections	42,697	25,581
<i>R</i> _{work} / <i>R</i> _{free}	0.207/0.262	0.254/0.266
No. atoms		
Protein	8,291	4,168
Ligand/ion	155	66
Water	156	–
<i>B</i> -factors		
Protein	52.7	85.9
Ligand/ion	74.1	119.2
Water	47.5	–
R.m.s deviations		
Bond lengths (Å)	0.010	0.009
Bond angles (°)	1.4	1.3

Values in parentheses are for the highest-resolution shell.