Molecular structure of the ATP-bound, phosphorylated human CFTR

Zhe Zhang¹,², Fangyu Liúa,b,¹, and Jue Chena,c,²

¹Laboratory of Membrane Biophysics and Biology, The Rockefeller University, New York, NY 10065; ²Tri-Institutional Training Program in Chemical Biology, The Rockefeller University, New York, NY 10065; and ³Howard Hughes Medical Institute, Chevy Chase, MD 20815

Edited by Richard W. Aldrich, The University of Texas at Austin, Austin, TX, and approved October 30, 2018 (received for review September 4, 2018)

The cystic fibrosis transmembrane conductance regulator (CFTR) is an anion channel important in maintaining proper functions of the lung, pancreas, and intestine. The activity of CFTR is regulated by ATP and protein kinase A-dependent phosphorylation. To understand the conformational changes elicited by phosphorylation and ATP binding, we present here the structure of phosphorylated, ATP-bound human CFTR, determined by cryo-electron microscopy to 3.2-Å resolution. This structure reveals the position of the R domain after phosphorylation. By comparing the structures of human CFTR and zebrafish CFTR determined under the same condition, we identified common features essential to channel gating. The differences in their structures indicate plasticity permitted in evolution to achieve the same function. Finally, the structure of CFTR provides a better understanding of why the G178R, R352Q, L927P, and G970R/D mutations would impede conformational changes of CFTR and lead to cystic fibrosis.

In this work, we report a structure of the phosphorylated, ATP-bound human CFTR channel determined by cryo-EM. The hCFTR structure shares many key features with that of zCFTR determined under the same condition (13). It also reveals several differences that are important to correlate the structure with functional data obtained from hCFTR.

Results

Structure Determination. Based on previous work showing that eliminating ATP hydrolysis prolongs the lifetime of the open CFTR channel (14), we replaced a catalytic residue in NBD2 (E1371) to abolish its ATPase activity. The mutant construct, E1371Q, was treated with PKA to phosphorylate the R domain. The cryo-EM structure was determined in the presence of saturating ATP-Mg2⁺ (9 mM). The final reconstruction, with an estimated overall resolution of 3.2 Å (SI Appendix, Fig. S1 and Table S1), shows well-defined densities for the lasso motif, 12 TM helices, 2 NBDs, and 2 ATP-Mg2⁺ molecules (SI Appendix, Fig. S2). Additional density, likely corresponding to part of the phosphorylated R domain, is also visible (SI Appendix, Fig. S2). With this map, we built a model containing nearly all residues in the TMDs and NBDs, as well as a polyalanine helix of the R domain (Fig. 1B). The structure is refined to excellent geometry and stereochemistry statistics (SI Appendix, Table S1).

With this structure, we can now directly visualize conformational changes within hCFTR upon phosphorylation and ATP binding (Fig. 1). Compared with the dephosphorylated, ATP-free conformation, the two halves of CFTR move toward each other, changing the shape of the molecule from an inverted “V” to an elongated form. The two NBDs form a closed dimer, occluding two ATP molecules at their interface. In CFTR, the

Significance

Mutations in cystic fibrosis transmembrane conductance regulator (CFTR) cause cystic fibrosis, a lethal genetic disease occurring in people of northern European descent. Decades of study have been directed toward a molecular understanding of this ion channel. The structure presented here enables a direct correlation of structure with function, most of which has been characterized in human CFTR.
phosphorylated R domain relocates to the peripheral surface, away from the NBD dimer interface.

**Phosphorylated R Domain.** Structural comparison of the R domain before and after phosphorylation provides a plausible explanation of how the R domain regulates channel activity. When dephosphorylated, the R domain is wedged between TM helices 9, 10, and 12 and extends into the cytosol between the two NBDs (Fig. 2A). Dimerization of the NBDs, which leads to channel opening (14), is hindered by the presence of the R domain. In the phosphorylated, ATP-bound conformation, the R domain-binding surface is completely closed off (Fig. 2B). A helical segment is observed parallel to the membrane surface, interacting with TMs 10 and 11 and residues 34–39 of the lasso motif (Fig. 2B). Because the side chain densities in this region are inadequate, we can’t assign its amino acid sequence. NMR studies of the phosphorylated R domain indicated that residues 766–776, 801–810, and 826–835 are likely to be helical (15). The distance between the observed helix and the first visible residue in TMD2 (T845) is 20 Å (Fig. 1B), in line with all three possibilities. In the dephosphorylated hCFTR structure, a helical segment of the R domain was also observed (Fig. 1A).

Because this helix is connected to TMD2, it is most likely to correspond to residues 825–843 (Fig. 1A). At a lower contour, amorphous densities are observed along the peripheral surface of NBD1 and at the C-terminal region of NBD1 (Figs. 1B and 2B). These densities most likely correspond to parts of the R domain, consistent with studies showing that the R domain is largely unstructured and interacts with NBD1 at multiple locations (15–17). Interaction of the R domain with the lasso motif was previously reported using isolated peptides (18), but the specific residues identified to mediate the interaction are different from those observed in the cryo-EM structure.

These structural observations indicate that the dephosphorylated R domain inhibits channel activity by preventing NBD dimerization. It seems that phosphorylation by PKA stabilizes the R domain away from the NBD interface, permitting the conformational changes necessary for channel opening. This conclusion was inferred from our previous structural study of the zebrafish CFTR, but here we observe parts of the phosphorylated R domain on the outside surface of NBD1. In fact, the functional effects of phosphorylation can be imitated by simply deleting the R domain. Coexpression of the two halves of CFTR
without the R domain results in an ATP-gated channel that opens independent of phosphorylation (19).

**The Ion Permeation Pathway.** Using a probe equal to the size of a dehydrated chloride ion (1.7-Å radius), a passageway into the heart of the channel can be identified (Fig. 3 A and B). Functional studies indicate that this passageway is the path of ion flow in CFTR (9, 12). Comparison of hCFTR in two conformations show that upon NBD dimerization, the pore penetrates 6 Å further toward the extracellular space, reaching a constriction area comprised of residues L102, F337, T338, and N1138 (Fig. 3 A–D). Consistent with this conformational change, residues L102C, F337C, T338C, and N1138C in cysteine-free constructs are accessible to intracellular cysteine reactive reagents only in the channel-open state (20–22).

In the NBD-dimerized conformation, the pore remains connected to the cytosol through a gap between TMs 4 and 6 (Fig. 3B). At the extracellular surface, a small opening is observed between TM 1 and TM 6 (Fig. 3 E and F). The radius of the opening is about 1.2 Å, too small to permit ion or water permeation. Among the residues surrounding this opening, S341 and T338 have been proposed to gate the pore (20, 21). Three positively charged residues, R104, R334, and K335, cluster on the extracellular side of the opening (Fig. 3 E and F). Modification of these residues influences anion conduction (22–24), suggesting that these residues are involved in attracting extracellular anions through electrostatic interactions. It seems possible that this small opening is the extracellular mouth of the pore. However, we suspect a fully conductive channel would require further dilation.

**Structural Comparison Between Zebrafish and Human CFTR.** The structure of the phosphorylated, ATP-bound hCFTR share many features with that of zCFTR (13) (Fig. 4). For example, the two NBDs...
form a “head-to-tail” dimer occluding two ATP molecules, the ion conduction pathway is lined with positively charged residues, and the pore is open to the cytoplasm between TMs 4 and 6. Also similar to zCFTR, TM 8 shows an unusual helix–loop–helix configuration inside the membrane and TM 7 locates outside the ion conduction pathway in a position atypical of ABC transporters. These conserved features emphasize structural aspects central to CFTR function.

One major structural difference between the phosphorylated zCFTR and hCFTR occurs in the TMDs (Fig. 4B). Although the overall configuration of the pore is similar, the positions of many individual residues are not identical. When superimposed, the root-mean-square deviation (rmsd) for all 627 C\textalpha\ positions in TMD is 2.0 Å. The difference is more prominent in the extracellular region of TMs 1, 6, and 12, where the largest displacement of equivalent residues is 6 Å (residues 107–109 of TM 1 and 1129–1134 of TM 12; Fig. 4B, red region). The small extracellular opening observed in hCFTR (Fig. 3E and F) is completely sealed in zCFTR.

A second major structural difference is found in the NBD dimer (Fig. 4C). In CFTR, as well as other ABC transporters in the ABCC subfamily, only one ATPase site is catalytically competent. The degenerate site binds but does not hydrolyze ATP (25, 26). The NBD dimer in zCFTR is asymmetric: The degenerate site is more open than that of the consensus site (Fig. 4C). However, in hCFTR, the NBD dimer is symmetric: Both ATPase sites are completely closed off when ATP is bound. Symmetric ATP binding was also observed for the multidrug resistance protein 1, a closely related ABC transporter in the ABCC subfamily (27). Thus, the asymmetric NBD dimer observed

![Fig. 4. Structural comparison of phosphorylated human and zebrafish CFTR. (A) Superposition of the overall structures. hCFTR is shown in blue, and zCFTR (PDB ID code 5W81) is in yellow. (B) An extracellular view showing the different locations of TMs 1, 6, and 12 (highlighted as cylinders) in the two structures. The extracellular region of TMs 1 and 6, where the residues have the largest displacement between these two structures, are colored in red (residues 107–109 and 1129–1134 of hCFTR, and 107–109 and 1137–1142 of zCFTR). (C) The symmetric NBD dimer in hCFTR versus the asymmetric dimer in zCFTR. The distances shown are between the C\textalpha\ atoms of the conserved lysine in the Walker A motif and the glycine in the signature motif (LSGGQ) at each ATP-binding site. (D) The distance between R104 and E116 in the two structures. (E) The R352/D993 salt bridge is only observed in hCFTR.](https://www.pnas.org/cgi/doi/10.1073/pnas.1815287115)
in zCFTR does not represent a general feature caused by the degenerate ATPase site.

Because of these structural differences, the hCFTR structure provides a better framework to interpret functional data, most of which have been obtained from hCFTR. For example, cysteines introduced at positions 104 and 116 were reported to form a spontaneous disulfide bond (28). Whereas the Ca atoms of these two residues are 11.7 Å apart in zCFTR, in hCFTR the distance is 6.5 Å, within the range of a disulfide bond (Fig. 4D). Another example is R352, whose functional role has been debated. Based on analysis of relative permeability of Cl− to Na+, it was suggested that R352 functions as a selectivity determinant by providing positive charge in bridge is indeed observed (Fig. 4D993 (32). In the structure of phosphorylated hCFTR, this salt bridge becomes accessible to thiol reagents (23). Additional studies show that the charge-swapping mutant R352E/D993R functions similar to the wild-type protein, suggesting that R352 forms a salt bridge with D993 (32). In the structure of phosphorylated hCFTR, this salt bridge is indeed observed (Fig. 4E), thus directly supporting the structural role of R352 in stabilizing the open pore. In zCFTR, the equivalent residues are too far apart to interact with each other.

**Disease-Causing Mutations Interfere with Conformational Changes.** More than 300 mutations have been identified to cause cystic fibrosis (https://cftr2.org/). Some mutations lead to defects in protein synthesis and processing, and others diminish channel activity. Understanding the molecular mechanisms underlying these mutations is key to personalized treatment as different therapy is needed to facilitate folding versus improve function. In addition, these mutations provide valuable information to help us understand the molecular details underlying CFTR gating. Previously, we mapped all of the 53 missense mutations onto the structure of zCFTR and categorized the mutants into four groups based on their structural consequences (12). At that time, we could not explain several mutations, including G178R, G970R, and L927P (12). Now with the structure of hCFTR, we can better rationalize why these mutations cause cystic fibrosis.

In the dephosphorylated, ATP-free conformation, residues G178 and G970 line the intracellular opening where the R domain resides (Fig. 5A). Upon phosphorylation and NBD dimerization, both glycine residues are engaged in close interhelical packing: G178 packs against TM 4 and G970 against TM 10 (Fig. 5A). Replacing either residue with arginine, as found in some cystic fibrosis patients, would be incompatible with the NBD-dimerized conformation. Consistent with this rationale, although both G178R and G970R mutants are expressed at the cell surface, they show little channel activity in iodide efflux assays (33, 34). Furthermore, the G970R mutant has a lower conductance compared with the G970D mutant, suggesting that the larger side chain of the arginine has a larger effect on NBD dimerization (35).

Similarly, L927P may also cause cystic fibrosis by interfering with conformational changes necessary for channel opening (Fig. 5B). The surface expression level of the L927P mutant is 43% that of the wild-type protein, but its channel activity is only 0.1% (36). Comparison of the two structures of CFTR indicates that the extracellular portion of TM 8 rotates about 55° upon phosphorylation and NBD dimerization. L927 is located within the hinge region in TM 8 (Fig. 5B). As proline contains a constrained cyclic side chain, it is unable to adopt many of the main chain conformations that leucine permits. Thus, the L927P mutation would limit the conformational flexibility of the hinge or redirect TM 8 into a nonconductive conformation. The fact that this mutation nearly abolishes the channel activity suggests that the TM 8 hinge is essential for normal gating.

**Summary**

Here, we have reported the structure of human CFTR in the phosphorylated, ATP-bound conformation. This structure reveals a previously unresolved helix belonging to the phosphorylated R domain, thus offering a more complete understanding of how R domain phosphorylation regulates channel activity. Direct comparison of human CFTR in two functional states establishes conformational changes necessary for channel opening. It also provides a better understanding of several disease-causing mutations.

**Methods**

**Protein Expression and Purification.** The E1371Q mutation was introduced to the wild-type construct with QuickChange Site-Directed Mutagenesis (Stratagene). The E1371Q mutant construct was expressed and purified using a protocol previously established for the wild-type CFTR (12). Briefly, the construct was expressed in HEK293S GnT1− cells and purified using a nanobody resin against the GFP tag attached to the C-terminal of CFTR (37). The membrane was solubilized with 1% 2,2-didecylpropane-1,3-bis-o-maltopyranoside (LMNG) and 0.2% cholesteryl hemisuccinate (CHS). The supernatant was applied to the GFP nanobody-coupled affinity resin followed by buffer exchange to 0.06% digitonin. The GFP tag was removed by PreScission protease and the protein was phosphorylated with PKA (NEB P6000L). Phosphorylated CFTR was further purified by gel filtration chromatography using a Superose 6 10/300 column (GE Healthcare) in a buffer containing 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 2 mM DTT, 1 mM MgCl2, 1 mM ATP, and 0.06% digitonin. The GFP tag was removed by PreScission protease and the protein was phosphorylated with PKA (NEB P6000L). Phosphorylated CFTR was further purified by gel filtration chromatography using a Superose 6 10/300 column (GE Healthcare) in a buffer containing 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 2 mM DTT, 1 mM MgCl2, 1 mM ATP, and 0.06% digitonin.

**EM Sample Preparation.** The protein sample was concentrated to 6 mg/mL and supplemented with additional 8 mM ATP, 8 mM MgCl2, and 3 mM fluorinated...
Data Collection and Processing. Data were collected on a 300-kV Titan Krios (FEI) microscope with a K2 summit detector (Gatan) in superresolution mode. Each micrograph was exposed for 10 s at a rate of 8 electrons/pixel per second and fractionated into 50 subframes. The images were normalized with gain reference and binned by 2, resulting in a physical pixel size of 1.03 Å. Contrast transfer function estimation and whole-frame drift correction were carried out using RELION (39). Subsequent data processing, including particle picking, 2D, and 3D classification, was carried out in Relion (40), except that drifts of individual particles were corrected using the program alignparts. Imbigs (41). The ATP-bound zebrafish CFTR (13) (EMD-8782) was used as the initial model for 3D classification. The map generated from Relion 3D refinement was then used as an initial model for further refinements in Frealign (42) using all 677,308 particles. The resolution of the final map is estimated to be 3.2 Å based on the correlation of the two half maps at the 0.143 value (SI Appendix, Fig. S1).

Model Construction and Refinement. We divided the dataset into two halves: one used for model building (the working set) and the other for validation (the free set). To better visualize the side chains for model building, the map calculated from the working set was sharpened using BFACORR.EXE (written by Nikolai Grigorieff) with a resolution limit of 3.2 Å and a b-factor of −50 Å². The final model was built in Chimera (43) and Coot (44). The final structure contains residues 1–390 of TM1D; 391–409, 435–637 of NBD1; 845–889, 900–1173 of TM2D; 1202–1451 of NBD2; and 17 residues of R domain (built as alamines). Phenix (45) and Refmac (46, 47) were used to refine the model in real space and reciprocal space, respectively. Fourier shell correlations (FSCs) were calculated between the model versus the working map, the free map, or the full map, and between the two half maps. Local resolutions were calculated using Blores (48). The geometry of the model was validated using MolProbity (49, 50). Figures were generated with Pymol (https://pymol.org/2), Chimera (43), and HOLE (51).

ACKNOWLEDGMENTS. We thank Mark Ebrahim and Johanna Sotiris of The Evelyn Gross Lipper Cryo-Electron Microscopy Resource Center at The Rockefeller University for assistance in data collection. This work was supported by Cystic Fibrosis Foundation Therapeutics (to J.C.) and the Charles H. Revson Senior Fellowship in Biomedical Science (to Z.Z.). J.C. is a Howard Hughes Medical Institute Investigator.