2021

Structural Study of Disease Relevant ABC Transporters-Cystic Fibrosis Transmembrane Conductance Regulator and ABCA4

Fangyu Liu

STRUCTURAL STUDY OF DISEASE RELEVANT ABC TRANSPORTERS—CYSTIC FIBROSIS TRANSMEMBRANE CONDUCTANCE REGULATOR AND ABCA4

A Thesis Presented to the Faculty of
The Rockefeller University
in Partial Fulfillment of the Requirements for
the degree of Doctor of Philosophy

by
Fangyu Liu
June 2021

STRUCTURAL STUDY OF DISEASE RELEVANT ABC TRANSPORTERS—CYSTIC FIBROSIS TRANSMEMBRANE CONDUCTANCE REGULATOR AND ABCA4

Fangyu Liu, Ph.D.
The Rockefeller University 2021

ATP-binding cassette (ABC) transporters are primary transporters that utilize the energy from ATP binding and hydrolysis to transport substrates across membrane against their concentration gradients [1]. Structurally, canonical ABC transporters consist of four subunits—two transmembrane domains (TMDs) which form the substrate transport pathway and two nucleotide binding domains (NBDs) which dimerize upon ATP binding to provide energy for substrate transport. Most mammalian ABC transporters are exporters, with three exceptions: SUR—a regulatory protein for the K_{ATP} channel, cystic fibrosis transmembrane conductance regulator (CFTR)—an chloride channel and ABCA4 (aka the Rim protein and ABCR)—an retinylidene-PE importer [1-3]. In addition to their unique functional properties, both CFTR and ABCA4 are very important in human health. Mutations in CFTR cause cystic fibrosis, a lethal disease with a prevalence of 1 in 2,500 in Caucasian populations [4, 5]. Over 800 mutations have been identified in ABCA4 to associate with various types of retinal disease [6], including the Stargardt disease (also known as juvenile macular degeneration), the most common form of inherited macular degeneration [7, 8].

To understand how those two proteins function, I mainly took a structural approach to capture the structures of CFTR and ABCA4 in different functional states. In correlating the structures with functional data, we now have a deeper mechanistic understanding of these two important ABC transporters.

The function of CFTR is regulated by ATP and phosphorylation. Once phosphorylated, ATP binding opens the CFTR channel and ATP hydrolysis closes it [9]. First, we determined by cryo-electron microscopy (cryo-EM), the structure of dephosphorylated human CFTR in the absence of ATP (Chapter 2). With this structure, in conjunction with the functional studies performed by our collaborator (Prof. Laszlo Csanady and Prof. David C. Gadsby), we were able to propose a mechanism of how phosphorylation regulates CFTR. In addition, we identified a structural feature distinguishing CFTR from all other ABC transporters, which likely forms the structural basis for CFTR's channel function.

Next, we determined the structure of CFTR in the phosphorylated, ATP-bound state (Chapter 3 and 4). By comparing the ATP-free and -bound structures, we identified the nature of conformational changes that lead to channel opening. These structures also allow us to map many disease-causing mutants and explain how they lead to the malfunctioning of CFTR.

To understand how small molecules, called potentiators, interact with CFTR and increase its open probability, we determined the structures of CFTR in complex with 2 potentiators—ivacaftor and GLPG1837 (Chapter 5). Interestingly, both small molecules bind to the same pocket inside the transmembrane region. These studies identified a hotspot on CFTR for rational drug design.

Finally, I also studied ABCA4, the only known importer in mammalian ABC transporters. To understand how ABCA4 functions, I determined the structures of ABCA4 in the absence and presence of ATP (Chapter 6). Based on these structures, we propose a rudimentary transport mechanism for ABCA4. Future work will be carried out in the Chen lab to test this model.

Dedication

Dedicated to my family members for their unconditional love and support

ACKNOWLEDGMENTS

First, I would like to give my deepest gratitude to my thesis advisor, Prof. Jue Chen. I got interested in studying disease-related membrane proteins while I was in college and applied to the Tri-Institutional Chemical Biology Program because of her. As I started my thesis in her lab, she was not only a great science mentor but also guided me when I got lost while making decisions in life. She has been tremendously supportive in letting me to do the science that excites me the most and in the way I wanted, while also devoting a huge amount of time sitting down with me and giving me a hand when I got stuck. From Jue, I learned how strong, dedicated and creative a successful scientist should be, so I can work towards that standard throughout my career.

Moreover, I would like to thank my thesis committee members, Prof. Seth A. Darst, Prof. Roderick MacKinnon, and Prof. Scott C. Blanchard. They always asked valuable scientific questions and provided suggestions to guide me to think more critically and work towards the goal from different routes. I would also like to thank Prof. David C. Gadsby. He was not only a great collaborator who helped us tremendously on starting the CFTR functional experiments, but also a great mentor who was always open to questions and discussions.

I have learned enormously from the people I met during my graduate life, both inside and outside the lab. Inside the lab, first and foremost, I would like to thank my mentor in the Chen lab, who is now a professor at Peking University in China, Dr. Zhe Zhang. He provided me with systematic training on the skillsets used in the lab and was extremely patient whenever I had questions. Additionally, I would like to thank Dr. Hongtu Zhao and Dr. Michael Oldham from the Chen lab and Dr. Xiao Tao, Dr. Chen Zhao, Dr. Chia-Hsueh Lee, Dr. Ji Sun, and Dr. Weiwei Wang from MacKinnon lab for spending hours with me to discuss scientific problems that I encountered. I would also like to thank Dr. Youngjin Kim, a former lab member of the Chen lab, who is now a professor in Korea, for being a great and caring friend. I would also like to demonstrate my appreciation to the girls in the lab—Ana-Maria Tanasescu, Donna Tallent, Dr. Natalie Omattage and Sarah McCarry for always being sweet and open to talking and helping. Moreover, thanks to Yuxi Zhang and Yiming Niu. They brought me tons of laughter with their young spirit. Outside the lab, I would like to thank my close friend Chuying Xia, who is currently a graduate student at Weill Cornell Medicine. We met in the first year of graduate school. I was very lucky to possess a great friend like her in an unfamiliar city. I would also like to thank my classmate who is a close friend as well—Putianqi Wang; we have spent lots of hours together figuring out future plans after graduation. Moreover, I would like to express my gratitude towards Dr. Elizabeth Campbell and Yanti Amos, who are my favorite cardio sculpt and yoga instructors. There are always invigorating and positive. In their classes, I can temporarily forget all the problems and just focus on the practices.

Finally, I would like to express my greatest appreciation to my family, especially my parents—Ms. Honglei Mao and Mr. Ning Liu. They have always been supportive about my decisions. They offered me guidance along the way and taught me to always look on the bright side.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: Introduction

1.1 Overview of ABC transporters

1.1.1 Function of ABC transporters

Cells are basic functional, structural, and biological units of all known organisms. They consist of cytoplasm enclosed within plasma membrane, which contains important biomolecules and organelles to carry out important functions such as energy production, nutrient uptake, and cell replication. Living organisms can be subdivided into two large categories: prokaryotes, which are made of a single cell, and eukaryotes, in which many cells work together. Thriving of both prokaryotes and eukaryotes rely on efficient nutrient uptake and extrusion of toxic products across the membrane. Simple diffusion of substances across the membrane is extremely slow and relies on the existence of concentration gradients; however, under many circumstances, organisms live under nutrient-poor situations which require them to take up as many nutrients as possible in the shortest amount of time. To cope with that over their period of evolution, organisms developed transport systems to transport substances across the lipid bilayers. One type of transport system is the ATP-binding cassette (ABC) transporter system which uses the primary energy source of cell-ATP to translocate substrates across membranes. ABC transporters usually transport their substrates unidirectionally. In prokaryotes, ABC transporters can function as both importers and exporters to take up nutrients from the environment and pump toxic materials out. In eukaryotes, ABC transporters mainly act as exporters to pump out foreign materials that are potentially hazardous to cell function [10].

1.1.2 Structural mechanism of archaetypical ABC transporters

ABC transporters have a conserved architecture which consists minimally of four domains: two nucleotide-binding domains (NBDs) that contain revolutionarily conserved sequence motifs and two transmembrane domains (TMDs) that form substrate translocation pathways. In some cases, additional domains exist to confer regulatory functions. For prokaryotic importers, periplasmic binding proteins are required for ligand delivery to those importers [11].

The most common transport model for substrate translocation by ABC transporters is an "alternating-access" model. In this model, transporters can alternatively open to outside or inside of the membrane. High affinity binding sites only exist on one side to prevent substrate backflow. As a substrate binds to its high affinity site, ATP binding or hydrolysis alter the conformation of the protein to achieve transport. In the ATP-free state, two NBDs are separated from other, and in the presence of ATP, ATP induces NBDs to dimerize together through conserved sequence motifs such as walker A, walker B, and signature motifs. This conformational change is communicated to the TMDs through coupling helices that are located in the loops between transmembrane helices. The TMDs are more structurally heterogeneous with distinct sets of folds for differentiation of substrate specificity [10].

1.2 General knowledge of CFTR

1.2.1 Mutations on CFTR cause cystic fibrosis

Cystic fibrosis (CF) is the most common lethal genetic disorder in populations of Northern European descent, affecting one out of every 2,500 newborns [4, 5]. It was first recognized as a separate disease entity in 1938 [12]. Later on, CF was recognized to be a genetic disease transmitted in an autosomal recessive manner, and the disease-causing gene CFTR was first identified and cloned in 1989 [5, 13, 14]. CF is characterized by accumulation of dehydrated mucus, recurrent bacterial infection, and other life-threatening consequences including organ failure. Cystic fibrosis primarily affects the lungs, but the disorder also impacts the pancreas, intestine, liver, kidneys, and sweat glands [15].

Over the years, sequencing information from 89,052 patients has now identified a total of 352 mutations that cause cystic fibrosis (http://cftr2.org/). Some mutations cause defective CFTR biosynthesis, resulting in diminished or absent protein production. Others are missense mutations that produce misfolded or dysfunctional protein. Among those, the most common mutation is ΔF508, where the amino acid F508 is deleted from the amino acid sequence. It is responsible for about 70% of all CF disease and results in intracellular retention and rapid degradation of the channel [16]. Other mutants, such as R117H and G551D, are expressed on the cell membrane but do not gate properly [17].

1.2.2 Functional studies on CFTR

When CFTR was identified in 1989, the mutant gene was found to encode an ABC transporter homologue with a highly charged cytoplasmic domain (R domain) containing several consensus sites for phosphorylation by protein kinase A. By that time, it was unclear whether CFTR directly served as an ion channel or a regulator for a chloride channel [14]. In the 1990s, CFTR was shown to be a cAMP-regulated chloride channel which conducts chloride ions in either direction with a conductance of ~10 pS [18-20].

The discovery of the CFTR gene spurred a huge amount of functional studies performed on the protein it encoded for. For example, in 1990, critical mutations such as ΔF508, ΔI507, K464M, F508R, and S549I were characterized biochemically to show that defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis [16]. In 2005, the tight coupling between nucleotide binding and the CFTR channel opening was established [21]. Over the years, vast amounts of functional studies were carried out to pinpoint the residues that were important for CFTR gating, selectivity, and conductance, so researchers were able to propose models for CFTR regulation and architecture [22-38].

However, full-length CFTR is resistant to crystallography due to its low expression level and unstable biochemical behavior upon purification. Obtaining the structures of CFTR would be very meaningful for the mechanistical understanding of the disease-causing protein and provide structural templates for developing the next generation of cystic fibrosis therapeutics.

1.2.3 Structural study of zCFTR

Before I began my thesis, with the advantages brought on by mammalian expression system and the resolution evolution of cryo-microscopy (cryo-EM), the first structure of CFTR (zebrafish analog) was determined to 3.7 Å resolution by Dr. Zhe Zhang in the Chen Lab after the initial discovery of CFTR gene more than 20 years ago [39]. The zCFTR structure was determined in its dephosphorylated, ATP-free form, which corresponds to zCFTR in its closed state. Under this state, the overall structure of CFTR resembles a typical ABC transporter in the inward-facing conformation, and the cytosolic NBDs are separated with the R domain wedged in between.

Several structural features that likely set CFTR apart as an ion channel were observed. First, a large vestibule was observed that opens to the cytosol, and the electrostatic surface of the tunnel is highly positive. Combined with the mutagenesis studies performed in the field, Dr. Zhe Zhang and Prof. Jue Chen were able to assign the ion conducting pathway of CFTR and observe a single gate for this pathway that is located near the extracellular side. Second, a discontinuous helix TM8 is observed as a special feature of CFTR. The helical breaks free main-chain atoms from their secondary structure hydrogen bonds and renders them available to interact with ions and water along the ion conduction pathway. Mutation on the TM8 loop (L927P) causes gating defects and leads to cystic fibrosis [40], suggesting that it may play a critical role in CFTR's channel activities.

Moreover, Dr. Zhe Zhang and Prof. Jue Chen interpreted and classified 53 missense CF-causing mutations by mapping these mutations onto the structure. Four classes were proposed: pore construction mutations, folding mutations, ATPase site mutations, and NBD/MTD interface mutations. The most prevalent mutation, ΔF508, is classified to be an NBD/TMD interface mutations. F508 is in nucleotide binding domain and interacts with a transmembrane domain through hydrophobic interactions. The structure clarifies how it impairs the assembly of the entire protein and thereby depletes cells of functional CFTR channels.

To understand the full gating cycle of CFTR, it is essential to capture structures of CFTR in different conformational states. Moreover, because zebrafish and human CFTR only share a 55% sequence identity, and most functional studies were of human CFTR, we reasoned that the structure of human CFTR is essential to directly correlate structure with function.

1.3 General knowledge of CFTR drugs

When cystic fibrosis was first recognized as a disease entity in the 1930s, the life expectancy was only approximately six months. Over the past eight decades, especially after the discovery of the CFTR gene, basic and clinical research enabled the lengthening of the average survival age of patients to around 47 years.

The first generation of CFTR drugs focused on symptomatic relief, including pancreatic enzyme supplements to aid digestion, antibiotics to prevent and treat infection, mucus-thinning drugs to clear the airway, and lung transplants. Since 2012, therapies have been developed to target the CFTR protein directly. Small-molecule CFTR modulators included correctors that increased the abundance of CFTR at the cell surface and potentiators that increased the ion flux of mutant CFTR [41-43]. Currently, two correctors (lumacaftor and tezacaftor) and one potentiator (ivacaftor), all developed by Vertex Pharmaceuticals, are available to patients [44-46]. In addition, many other candidates to enhance the function of CFTR are in the drug discovery pipeline [42, 47, 48]. All of these CFTR modulators were discovered through intensive high-throughput screening and iterative medicinal chemistry optimization. Rational drug discovery has not been feasible, owing to the lack of structural information.

To address this issue, we reasoned it would be meaningful to pinpoint the binding site and binding modes of CFTR drugs by solving the structure of CFTR in complex with these drugs through cryo-EM. This information would provide us and the field templates with a rational design of the next generation of CFTR drugs.

1.4 ABCR and Stargardt disease

1.4.1 Physiological role of ABCA4

The visual cycle initializes with the photoisomerization of 11 cis-retinal to all-trans retinal in photoreceptors. This event allows for the conversion of photochemical signal to electrochemical signal which can be processed by the brain. All-trans retinal is then released by photoreceptors and recycled to 11 cis-retinal for the next round of the visual cycle. Chemically, all-trans retinal is a reactive vitamin A aldehyde, which, when not recycled properly, may lead to retinal photodamage and consequently various types of retinal degenerative diseases such as Stargardt's disease and age-related macular degeneration. The elimination of all-trans retinal is carried out in two steps: translocation and reduction. After photoisomerization, all-trans retinal is released to the photoreceptor disc membranes (photoreceptor discs are specialized organelles in the photoreceptors where light perception happens). ABCA4 is responsible for translocating all-trans retinal from the intradiscal space to the cytoplasmic space so all-trans retinal can be reduced to all-trans-retinol by all-trans-retinol dehydrogenase (RDH) [49]. In 1997, ABCA4 was identified to be a causal gene of Stargardt disease (also known as juvenile macular degeneration), which is the most common form of inherited macular degeneration [7, 8]. In the same year, patients with AMD were also found to have a higher chance of harboring mutations in ABCA4 [50]. In 1998, through linkage analysis, mutations on ABCA4 was reported to be able to cause other eye diseases such as autosomal recessive retinitis pigmentosa and cone-rod dystrophy [51]. Presently, over 800 mutations associated with various types of retinal diseases have been identified on ABCA4 [6](http://www.hgmd.cf.ac.uk/ac/all.php).

1.4.2 Architecture of ABCA transporters

Structurally, ABCA4 belongs to the ABCA subfamily, which includes twelve members that share a high degree of sequence conservation and are generally involved in trafficking lipids in different tissues. Mutations of ABCA genes lead to various types of genetic disorders. For example, mutations in ABCA1 can lead to Tangier disease [52] and ABCA3 mutations can cause neonatal surfactant deficiency [53]. In addition to the two TMDs and NBDs commonly found in other ABC transporters, members of the ABCA subfamily are characterized by having two large extracellular domains (ECDs) and two regulatory domains (RDs) whose functions are still unclear. Previously, the structure of a close family member, ABCA1, was solved in its ATP-free conformation with an overall resolution of 4.1 Å [54]. This structure demonstrates that the two ECDs form an elongated hydrophobic tunnel and provides a structural interpretation of disease-related mutations. However, for ABCA4, the only structural understanding we have is through the 18 Å electron microscopic reconstruction [55]. Functionally, ABCA4 is the only recognized importer in mammals; therefore, it is interesting to investigate how ABCA4 functions by capturing it in different functional states.

1.5 Dissertation summary

This dissertation includes all of my projects in the Chen lab, which aim at understanding the structure-function relationship of disease-related ABC transporter systems. Except for the first chapter, each chapter is a summary of each project. The first chapter provides an overview of the importance of ABC transporters, giving necessary background knowledge for the readers to understand the importance of the questions we are interested in.

Chapter 2 focuses on the structural understanding of human CFTR in its closed state. As mentioned in 1.2.3, previously solved zebrafish CFTR gave a comprehensive overview of how CFTR can cause cystic fibrosis; however, as a highly relevant protein in causing human disease, the observations needed to be confirmed in human CFTR. This project did confirm the previous observations. In addition, we collaborated with Prof. Laszlo Csanady and Prof. Dr. David C. Gadsby to propose a regulatory mechanism for CFTR. Dr. Zhe Zhang and I contributed equally to this project.

Chapter 3 and 4 summarize the work on the cryo-EM structures of zebrafish and human CFTR in their pre-/post-open state in order to fully understand the full-gating cycle of CFTR. The understanding of the full cycle may aid future drug development in favoring CFTR's open state to treat cystic fibrosis. In chapter 2, Dr. Zhe Zhang collected, processed, and interpreted the data, and I assisted in the data processing. In chapter 3, Dr. Zhe Zhang and I contributed equally.

Chapter 5 describes the structural complexes between CFTR and CFTR potentiators (one class of CFTR drugs) ivacaftor (FDA-approved) and GLPG1837 (an investigational drug), respectively. This project revealed a common drug binding pocket on CFTR for potentiators and provided a template for future development of improved potentiators. Dr. Anat Levit and Prof. Brian Shoichet verified the binding mode of two potentiators with computational methods. Jesper Levring tested the potentiating activity of GLPG1837 in lipid bilayers. Dr. Zhe Zhang and I contributed equally to this project.

Chapter 6 explores the protein ABCA4 whose mutations can lead to various types of vision disease such as Stargardt macular degeneration. Structurally, it is interesting to explore how ABCA4 acts as the only known mammalian importer given its high similarity with its subfamily members. By solving ABCA4 in two conformations, we were able to propose an initial hypothesis for ABCA4 transport which can be tested in the future.

CHAPTER 2: Cryo-EM structure of dephosphorylated human CFTR

2.1 Summary

After the zebrafish CFTR structure was solved on 2017, it was still not known whether human CFTR would look the same. As protein whose mutations directly lead to the human disease cystic fibrosis, obtaining a structure of human CFTR is was critical. In this project, we solved the structure of dephosphorylated human CFTR without nucleotides to 3.9 Å by electron cryomicroscopy (cryo-EM). Through direct comparison with zebrafish CFTR, we found that they were in close resemblance. In addition, the human CFTR structure reveals a previously unresolved helix belonging to the R domain docked inside the intracellular vestibule, precluding channel opening. Through collaboration with Professor Laszlo Csanady and Professor Dr. David C. Gadsby, we propose that PKA phosphorylation of the R domain is enabled by its infrequent spontaneous disengagement, which also explains residual ATPase and gating activity of dephosphorylated CFTR. From comparison with MRP1, a feature distinguishing CFTR from all other ABC transporters is the helix-loop transition in transmembrane helix 8, which likely forms the structural basis for CFTR's channel function.

This chapter (including this summary) is adopted from my work co-authored with Dr. Zhe Zhang, Prof. Laszlo Csanady, Prof. Dr. David C. Gadsby and Prof. Jue Chen, published in *Cell.* 2017 [56].

2.2 Results

2.2.1 Structural determination of CFTR in its dephosphorylated, closed conformation

The human CFTR construct was not initially chosen for structural study due to its imperfect chromatographic behavior—it had an obvious shoulder peak preceding the main peak of CFTR (Fig. 2.1). The co-first author of this section, Dr. Zhe Zhang previously attmepted different purification conditions, but none of the conditions provided better chromatographic behavior than human CFTR purified in LMNG:CHS exchanged to digitonin. However, given its high relevance in the fatal disease cystic fibrosis, we decided to try resolve the structure given its imperfect behavior. First, we froze grids with untreated wild type CFTR. The structure only resolved to ~7 Å likely due to the non-homogenous nature of the protein. CFTR likely existed in different phosphorylation states when they were purified (Fig. 2.2).

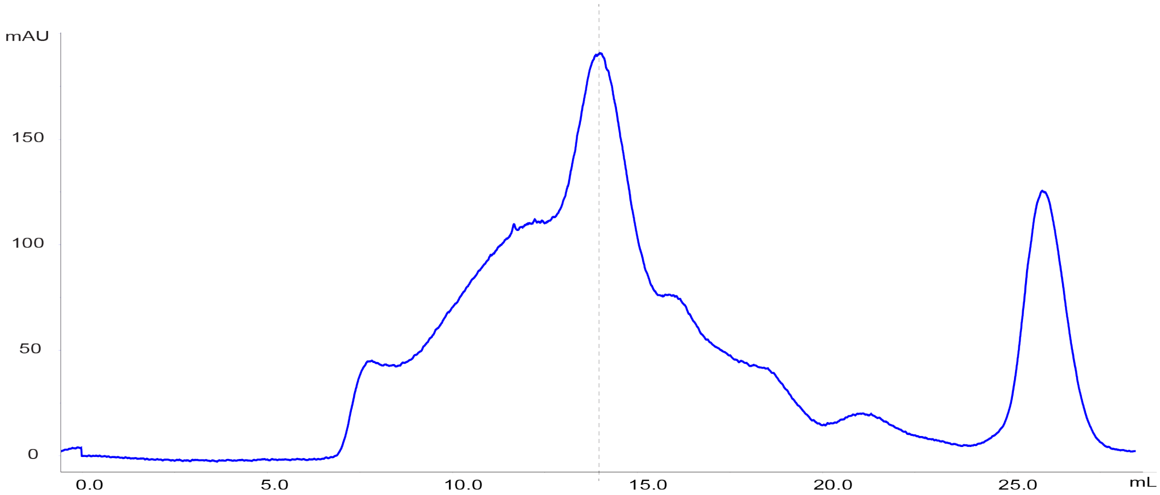


Fig. 2. 1 Gel filtration profile of human CFTR

Blue trace shows the gel filtration behavior of CFTR ran in Sepharose 6 column. The dashed line indicates the elution volume of CFTR.

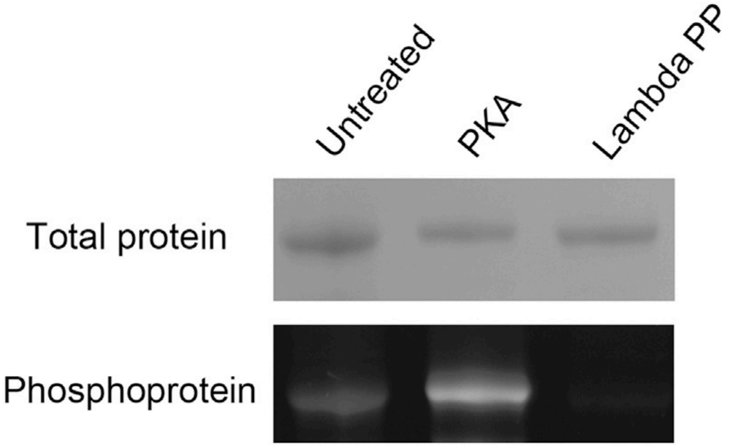


Fig. 2. 2 Analysis of the phosphorylation state of purified CFTR

Untreated, PKA-treated, and Lambda Protein Phosphatase (PP)-treated samples on SDS-PAGE were stained by Coomassie blue (upper) and the Pro-Q Diamond Phosphoprotein Gel Stain (lower).

Given the observation, to obtain a homogenous sample of the human CFTR in its closed, NBD-separated conformation, we treated human CFTR with lambda phosphatase (Sigma: P0753L) after the protein was eluted from the GFPnb-coupled Sepharose beads. A subsequent gel filtration step separated CFTR with the lambda phosphatase, and the peak fraction from the gel filtration was concentrated to 5 mg/mL for grid freezing.

Out of this sample preparation, although we can observe moderate amount of aggregation on the raw micrographs, we were able to obtain a Cryo-EM map of CFTR to 3.9 Å. The local resolution, estimated by Blocores from the Bsoft Package [57], varies from 2.4 to 6.0 Å. The TMDs were resolved better than the NBDs likely due to the local flexibility of NBDs. Therefore, we built the transmembrane helices *de novo*, whereas for the NBDs, we docked the high-resolution crystal structures in, and the side-chain positions were manually adjusted based on the EM density (Fig. 2.3). A tubular density was clearly visible between the two halves of the molecule; but it lacked side-chain features to establish the amino acid register (Fig. 2.3, 2.4). We modeled a short helix into this density, and according to its position, this helix likely corresponds to the C-proximal end of the R domain. The final structure, which contains nearly complete atomic models for the lasso, TMD1, TMD2, NBD1, NBD2, and a polyalanine model of residues 825–843, was refined against the EM maps to good statistics and excellent stereochemistry (Fig. 2.5, 2.6 and Table 2.1).

In this non-phosphorylated, closed-channel conformation, CFTR resembles a typical ABC transporter in its inward-facing conformation (Fig. 2.5). The 12 TM helices in the TMDs pack tightly in the membrane outer leaflet. In the inner leaflet, the helices segregate into two bundles extending ~30 Å into the cytoplasm. The two NBDs, each attached to a cytosolic TM helical bundle, are well separated. The R domain, mostly unstructured, inserts between the cytoplasmic extension of the TMDs and packs along NBD1 in the opening between the two halves of the molecule.

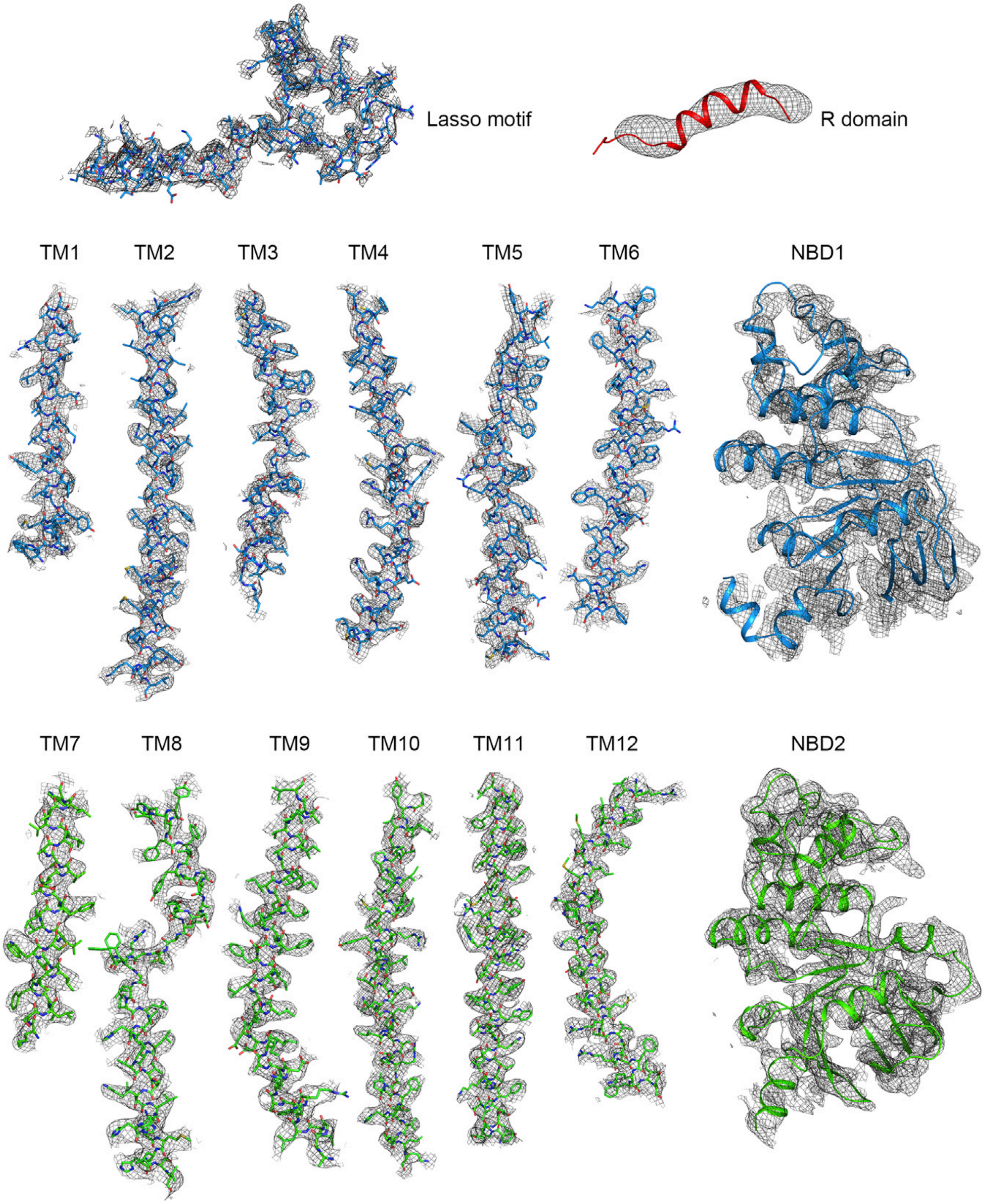


Fig. 2. 3 EM Density of different parts of the dephosphorylated CFTR structure

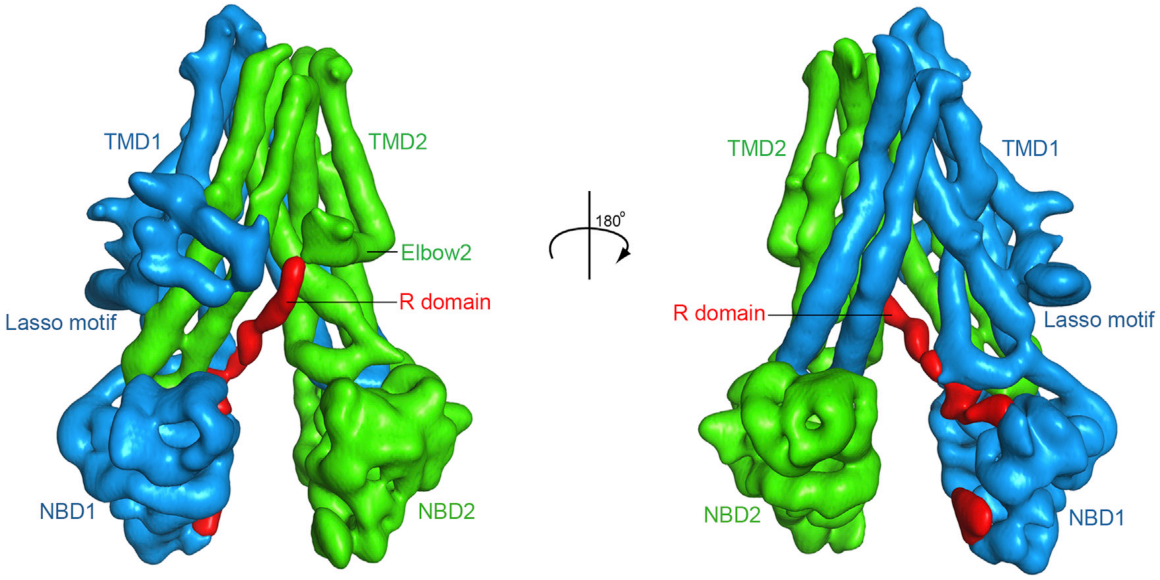


Fig. 2. 4 The Cryo-EM Map of dephosphorylated CFTR without B-factor sharpening

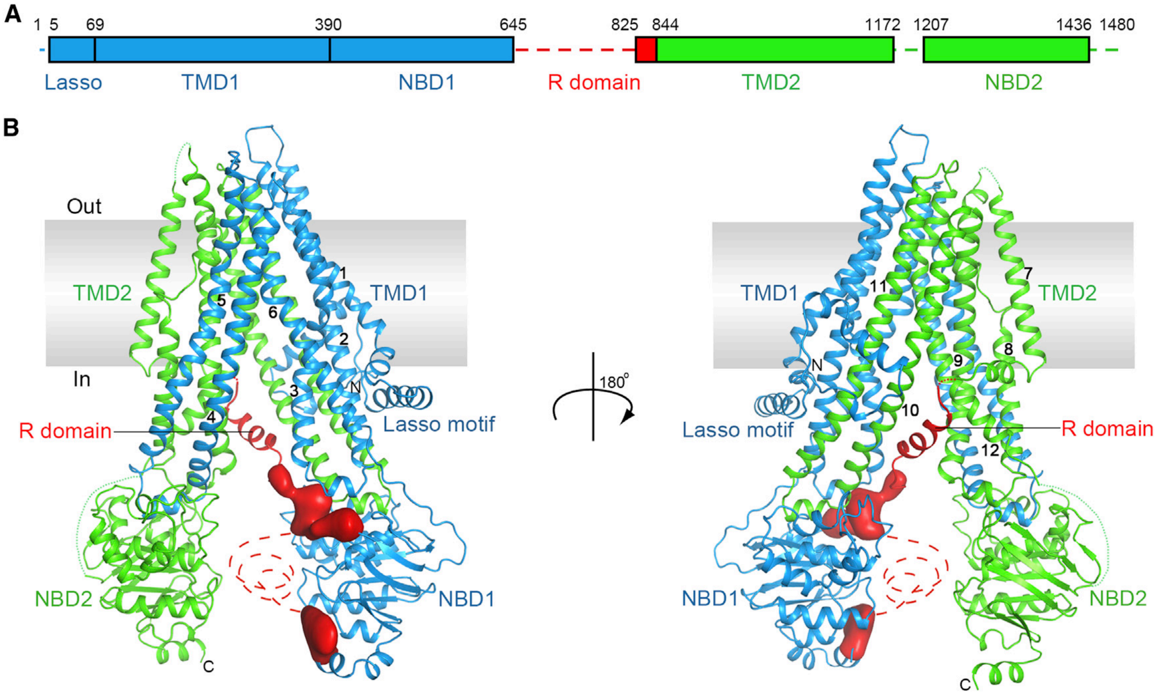


Fig. 2. 5 The overall Structure of human CFTR in the dephosphorylated, ATP-free conformation

(A) The domain structure of CFTR. (B) Ribbon diagram of two views. The EM densities shown in red correspond to unstructured regions within the R insertion of NBD1 or the R domain. Regions not resolved in the structure are shown as dashed lines for visualization purposes only, they are not meant to imply connectivity or the location of the missing regions.

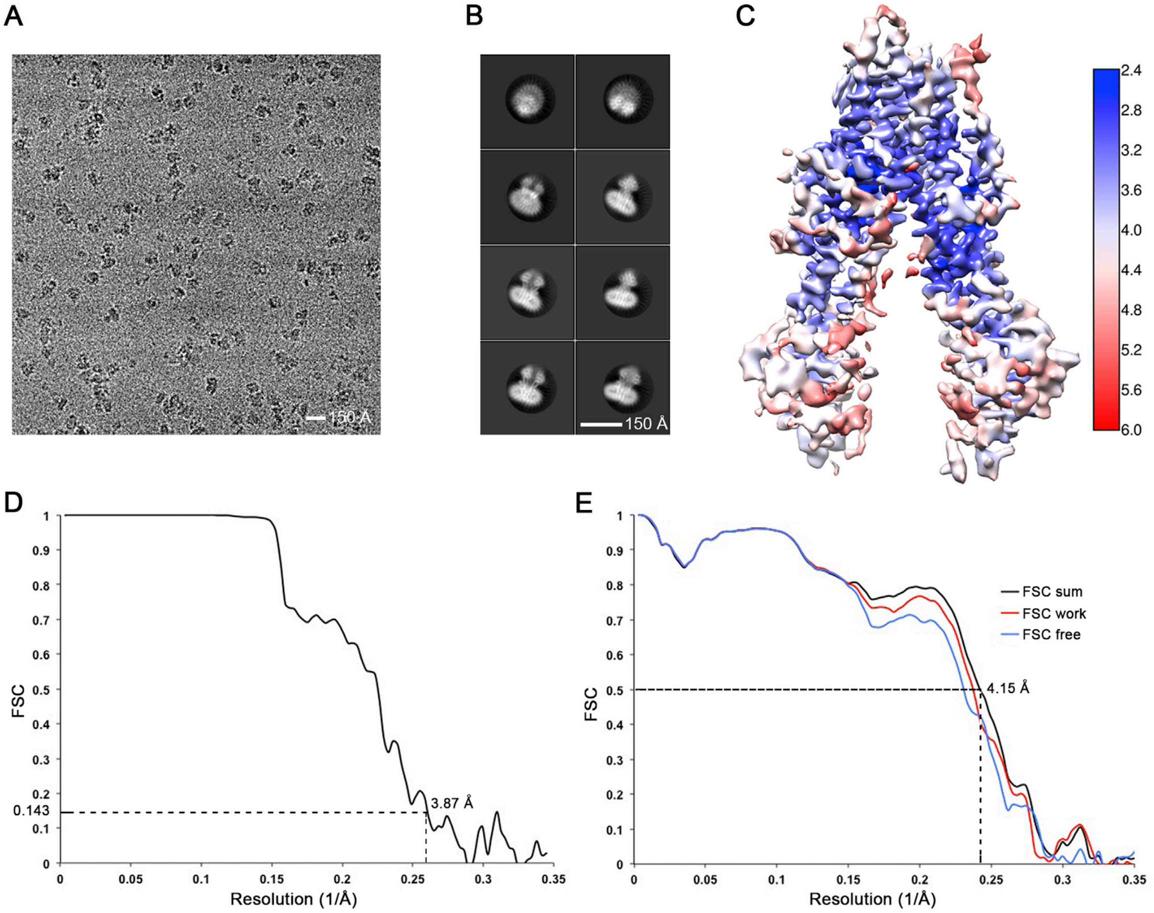


Fig. 2. 6 Cryo-EM analysis of dephosphorylated human CFTR
(A) A representative micrograph at defocus of -2.3 *μ*m. (B) Representative 2D class averages, output of Relion [58]. (C) The overall EM map colored by local resolution estimation using Blocres [57]. (D) Fourier Shell Correlation (FSC) as a function of resolution, using Frealign [59] output. (E) FSC calculated between the refined structure and three different maps. Black: the map calculated using the full dataset (sum); Red: the half map used in refinement (work); Blue: the other half map (free).

Table 2. 1 Summary of EM data and structure refinement statistics

Data collection	
Microscope	Titan Krios (FEI)
Voltage (kV)	300
Detector	K2 Summit (Gatan)
Pixel size (Å)	0.817
Defocus range (µm)	-0.8 to -2.5
Movies	3,642
Frames/movie	50
Dose rate (electrons/pixel/s)	8
Total dose (electrons/Å^2)	84
Number of particles	415,915
Model composition	
Non-hydrogen atoms	9,232
Protein residues	1,158
Refinement	
Resolution (Å)	3.87
Average Fourier Shell Correlation	0.8085
R_{work}	0.2476
R_{free}	0.2599
RMS deviations	
Bond lengths (Å)	0.0072
Bond angles (°)	0.9422
Validation	
Molprobity score	0.93
Clashscore, all atoms	0.37
Favored rotamers (%)	94.9
Ramachandran plot (%)	
Favored	95.7
Allowed	4.3
Outliers	0.0

2.2.2 Structural comparison of human and zebrafish CFTR

This structure of human CFTR is very similar to that of the dephosphorylated zebrafish ortholog, evident by the overall root-mean-square deviation (RMSD) of 1.9 Å for 1,062 Ca positions (Fig. 2.7A). Several prominent features of CFTR, including the N-terminal interfacial lasso motif (Fig. 2.7B), the discontinuity of helix TM8 (Fig. 2.7C), and the asymmetric opening of the NBDs (Fig. 2.7D), are conserved in both human and zebrafish orthologs. The positions of functionally important residues, such as the R347/D924 (TM6/TM8) salt-bridge (Fig. 2.7E), the TM6 gating residues F337 and T338 (Fig. 2.7E), and many pore-lining residues (Fig. 2.8) are very similar in both structures. These structural similarities between human and zebrafish CFTR, given that they possess only 55% sequence identity, have significance. Our observation of closely similar structures in CFTR from two different species offers confidence that both cryo-EM structures represent a functionally informative conformation. Thus, major conclusions reached on the basis of the zebrafish ortholog [39], including the construction of the ion-conduction pathway, the separation of the NBDs, the positioning of the R domain, and the structural interpretation of many cystic fibrosis-causing mutations, are now reinforced by the human CFTR structure. It is quite remarkable that the distances between the two NBDs are so similar in zebrafish and human CFTR

(Fig. 2.7D). Many ABC transporters appear to be very flexible in the inward-facing conformation. For example, several crystal structures of the mouse P-glycoprotein display a distribution of distances separating the two NBDs [60]. Similarly, EM studies of MsbA and P-glycoprotein revealed numerous conformations of both transporters [61, 62]. In stark contrast, we find practically identical conformations and NBD separations in two structures of CFTR from different species, suggesting that the conformational flexibility observed in other ABC transporters is substantially diminished in CFTR. We propose this is due to the presence of the R domain, which in its unphosphorylated state is seen to contact both halves of the molecule (Fig. 2.5).

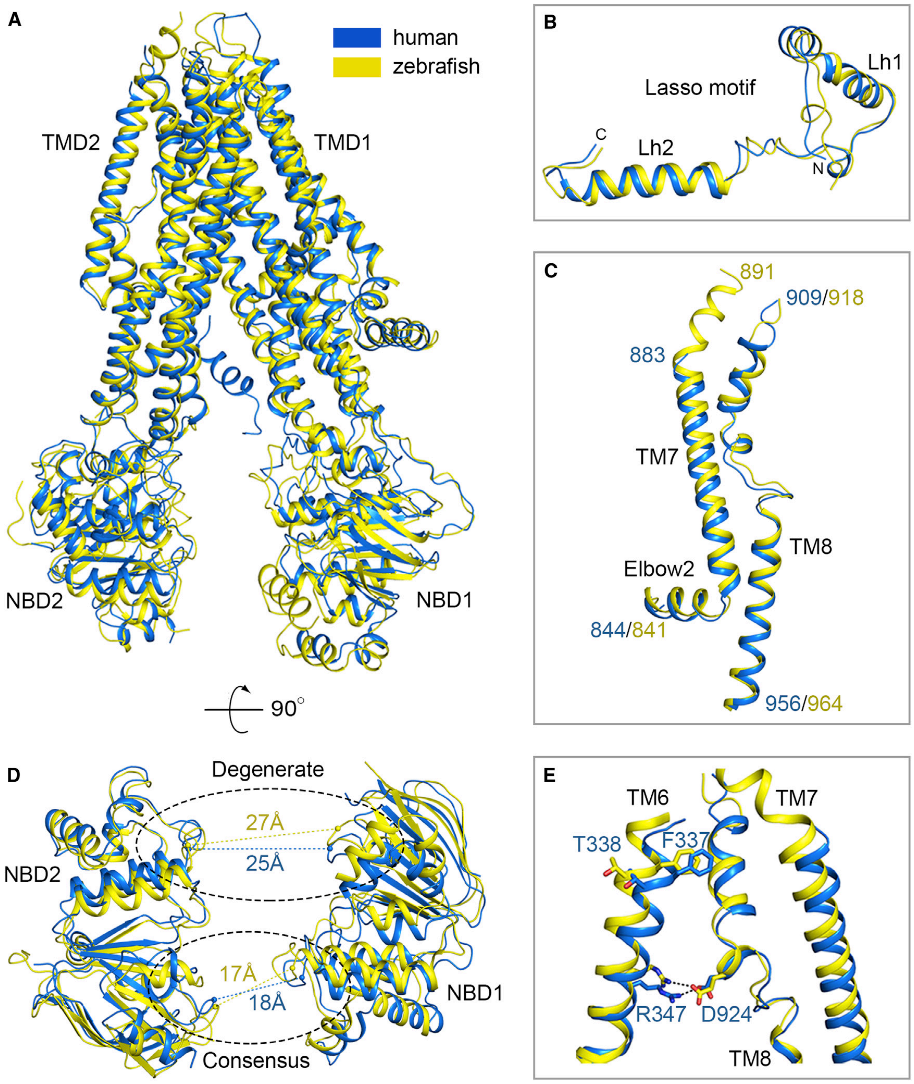


Fig. 2. 7 Structural Comparison of Human and Zebrafish CFTR

(A) Superposition of the overall structures. Human CFTR is shown in blue, and zebrafish CFTR is in yellow. (B) Superposition of the lasso motif. (C) The TM helix 7 in zebrafish CFTR is ten residues longer than that of human CFTR. The residue numbers of human (blue) and zebrafish (yellow) CFTR are indicated. (D) The asymmetric opening of the NBD interface is conserved in both human and zebrafish CFTR. The distances between the Ca atoms of the conserved glycine in the Walker A motif and the serine in the signature motif are indicated. (E) Superposition showing the locations of several functionally well-characterized residues.

2.2.3 The ion conduction pathway of CFTR

We mapped the pore of the human CFTR in this dephosphorylated, nucletotide-free conformation with the program HOLE [63] with the size of a Cl^- ion. The ion conduction pathway where is accessible by Cl^- ions is shown in grey mesh (Fig. 2.8).

Integral membrane proteins in general contain predominantly hydrophobic residues that form transmembrane segments. Positively charged arginine and lysine residues are often found at the cytoplasmic boundaries of the membrane proteins but are rare within the membrane core or near the extracellular surface. This biased distribution of positively charged residues, the "positive-inside" rule, is important in directing the topology of membrane proteins [64, 65]. CFTR, however, contains a large number of arginines and lysines distributed throughout the TMDs. The functional roles of individual arginine and lysine residues have been studied extensively [66]. Single amino acid substitution of K95, R104, R117, K190, R248, R303, R334, K335, R352, K370, K1041, and R1048 decreases the single-channel conductance in a charge-dependent manner, suggesting that they play a role in attracting anions into the pore through electrostatic interactions [26-29, 31, 36, 67]. Of these residues, R104, R117, R334, and K335 are located in the extracellular region (Fig. 2.8, pink residues). Removing these positive charges leads to inward rectification, consistent with their role in recruiting extracellular anions [25, 29, 67]. All other functionally important arginines and lysines are located on the surface of the ion conduction pathway (Fig. 2.8, magenta residues). The side chains of these residues, not engaged in any specific interactions with other parts of the protein in the closed channel, are available to bind anions. In contrast, numerous positively charged residues do not affect anion permeation, including R80, K114, R242, R251, K254, K294, R297, K298, K329, R933, K946, R975, R1030, R1102, R1128, R1162, and R1165 [27, 29, 36, 68]; all of these residues are located far from the pore (Fig. 2.8, yellow residues).

Do these positively charged residues contribute to ion selectivity? CFTR is a relatively non-selective anion channel, permitting permeation of many monovalent anions including thiocyanate, nitrate, bromide, chloride, iodide, fluoride, formate, and bicarbonate [66]. Mutagenesis studies have not identified any particular residue involved in determining the observed lyotropic anion selectivity sequence [66]. Therefore, it is likely that the selection of anions over cations in CFTR comes from the overall positive charge of the pore [39] rather than individual residues or a narrow selectivity filter as seen in K channels [69]. Large divalent anions, such as phosphates and sulfates, are not conducted by CFTR. Whether this is due to a narrow region of the ion pathway must await the structure of an open CFTR channel.

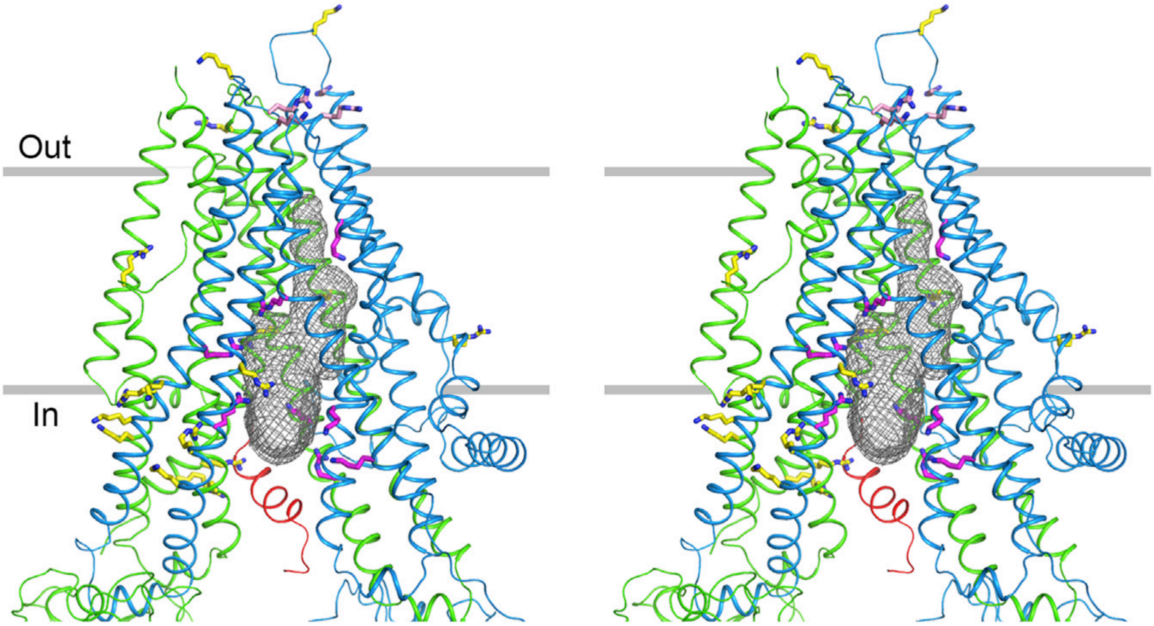


Fig. 2. 8 Stereo view of the ion-conduction pathway
The pore is shown as a gray mesh. Positively charged residues involved in recruiting extracellular ions are shown in pink, including R104, R117, R334, and K335 [25, 29, 67]. Other residues involved in ion conductance are shown in magenta, including K95, K190, R248, R303, R352, K370, K1041, and R1048 [26-29, 36]. Positively charged residues not involved in ion conduction are shown in yellow, including R80, K114, R242, R251, K254, K294, R297, K298, K329, R933, K946, R975, R1030, R1102, R1128, R1162, and R1165 [27, 29, 36, 68].

2.2.4 Positioning of R domain

It has been established that opening of a CFTR channel is strictly coupled to the formation of a closed NBD dimer [21, 70]. The location of the unphosphorylated R domain observed in both human (Fig. 2.5) and zebrafish CFTR structures [39] is consistent with its inhibitory role [71] by preventing NBD dimerization and thereby channel opening [39]. In human CFTR, we also resolve an additional helical structure wedged into the intracellular space between the two halves of the molecule, interacting with TM helices 9, 10, and 12 (Fig. 2.9A). Because the density of this region is positioned so as to connect to the elbow helix in TMD2 (Fig. 2.4), it most likely corresponds to residues 825–843 of the R domain (R825–843).

Although there is no phosphorylation site within residues 825–843, the overall charge of this region is very negative: it contains seven aspartate and glutamate residues, which account for 37% of its amino acid composition. Many polar residues in TMD2 are in van der Waals's contact with R825–843 (Fig. 2.9A). Among these potential interacting residues, substitutions of K978 with cysteine, serine, or proline were shown to increase CFTR channel activity independent of ATP [72]. It would seem possible that removing the positive charge at position 978 lowers the affinity for R825–843, thus facilitating R domain release from its inhibitory position.

To further illustrate the fact that R825–843 must relocate to allow the channel to open, we compared the local structure of CFTR with that of Sav1866, a nucleotide-bound ABC exporter in the outward-facing conformation [73]. As shown in Fig. 2.9B, the intracellular cavity between the two helical bundles narrows upon NBD dimerization. This narrowing would be incompatible with the helical R825–843 inserted as observed in the closed CFTR channel structure.

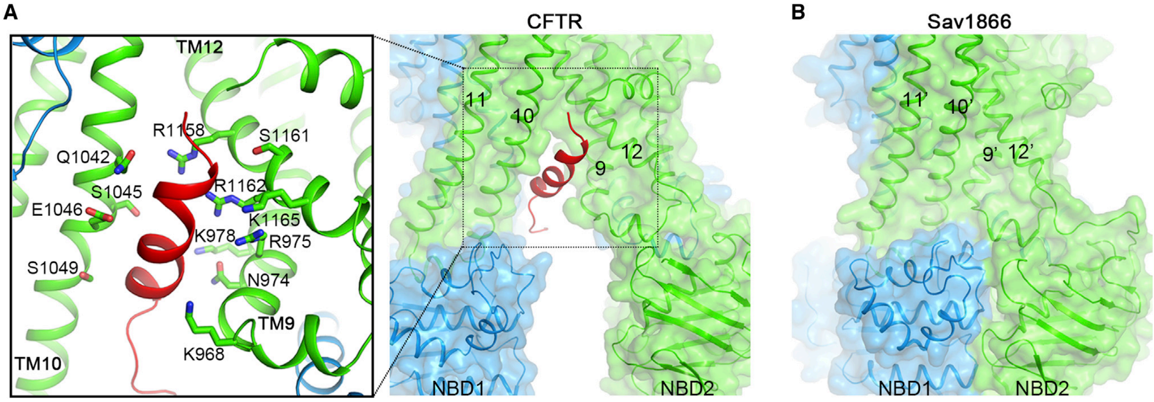


Fig. 2. 9 The dephosphorylated R Domain inhibits channel opening
(A) Residues 825–843 of the R domain (shown in red) insert into an opening between TM9 and TM10. Also shown in a close-up view are residues in van der Waals' contact with R825–843, including K978. (B) Structure of Sav1866 (PDB: 2ONJ) showing that helices corresponding to CFTR TM9 and TM10 are in close contact upon NBD dimerization.

2.2.5 Proposal of the regulatory mechanism of CFTR

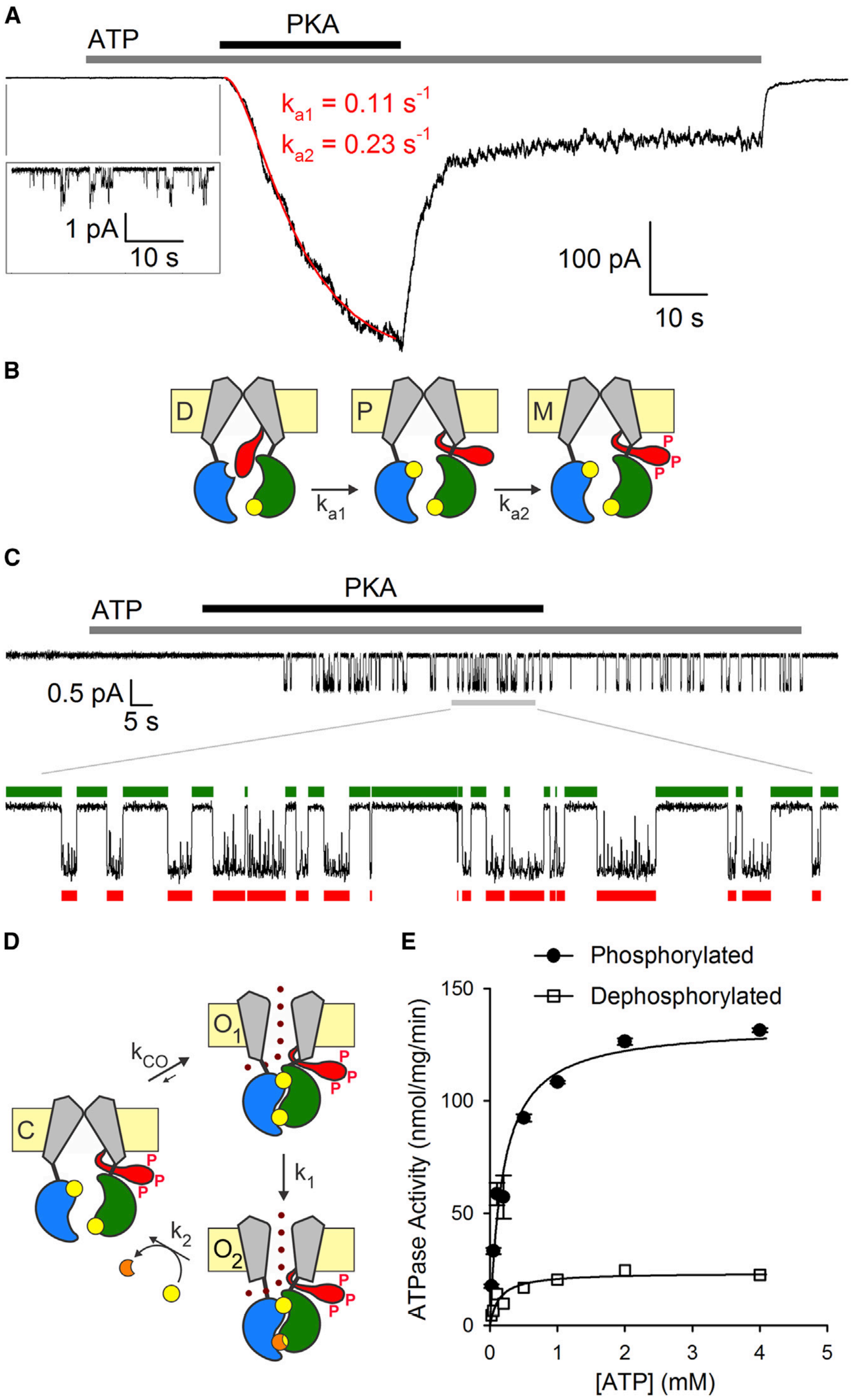
Easy access of PKA catalytic subunit (~30 kDa) to the R domain's multiple PKA consensus-site serines (Table 2.2) during channel activation by phosphorylation likely also requires the R domain to first be released from between the NBDs and TM extensions. Support for this comes from the sigmoid time course of current activation on sudden addition of a maximally activating concentration of PKA to membrane patches containing hundreds of human CFTR channels (Fig. 2.10A). Fits of such current activations (in 11 patches) to a two-step process (e.g., like that in Fig. 2.10B) yielded a slower (mean $k_{a1} = 0.07 \pm 0.01\ s^{-1}$) and a faster (mean $k_{a2} = 0.26 \pm 0.05\ s^{-1}$) rate for the two sequential activation steps. In contrast, application of a maximal PKA concentration to channels that are already partially activated increased current along an exponential time course, at a rate similar to that of the faster step, k_{a2} [74]. We therefore interpret the slower of the sequential steps to reflect infrequent spontaneous disengagement and release of the R domain and the faster step to reflect subsequent rapid phosphorylation of the R domain once released (strictly, because we are monitoring current, the faster rate k_{a2} must also encompass the, even faster, channel opening step). Accordingly, the corresponding CFTR channel states are cartooned in Fig. 2.10B as dephosphorylated ("D"), phosphorylatable ("P"), and maximally activated ("M"). The greatly magnified current segment inset in Fig. 2.10A shows that individual channels occasionally, if rarely, open and close without having been phosphorylated by PKA; this current of unphosphorylated CFTR channels exposed to ATP was previously noted [75] to amount to 1.2% ± 0.5% (n = 13) of that after the same channels were maximally phosphorylated. Because even spontaneous opening of CFTR channels is accompanied by NBD dimerization [70], which in turn requires disengagement of the R domain, we propose that the rare openings of unphosphorylated channels in Fig. 2.10A are enabled by comparably infrequent, reversible, R domain release events; those same rare releases of the R domain presumably also underlie the relatively weak ATPase activity of dephosphorylated CFTR (Fig. 2.10E). This low ATPase rate and rare channel openings in the absence of PKA can be explained by the steady-state distribution of the CFTR population among D, P, and M states (Fig. 2.10B) being shifted strongly leftward to only D and P states—with the slow D-to-P transition resulting in low fractional occupancy of state P, from which the rare ATP-driven gating and ATPase cycles can arise. In that case, most of the time, engaged R domains (state D) inhibit NBD dimerization, but occasional R domain release (to state P) supports a low level of activity. For occupancy of state P to remain low in the absence of PKA, spontaneous release of the R domain must be much slower than its spontaneous redocking. When PKA is present, relatively rapid phosphorylation of occasionally released R domains temporarily traps them, delaying their return to the inhibitory position and slowly accumulating the CFTR population in state M, hence accounting for the slow sigmoid time course of activation of channel current (Fig. 2.10A and 2.10C). Dephosphorylation of the trapped R domains by phosphatases (transition M-to-P) allows them to spontaneously redock in the channel vestibule (transition P-to-D). These spontaneous motions of the R domain likely contribute to its evident disorder observed in the cryo-EM reconstruction (Fig. 2.5).

A record from a patch containing a single CFTR channel (Fig. 2.10C) recapitulates the sequence of events observed for hundreds of channels (Fig. 2.10A): the very low probability of seeing a CFTR channel open before phosphorylation, and the delay to the first opening after adding PKA, can both be attributed to infrequent disengagement of the unphosphorylated R domain (i.e., transition D-to-P; Fig. 2.10B). Once phosphorylation of R domain serines reaches a steady state, so does channel gating.

In the latter cyclic process, two ATPs are bound in the closed channel state, C (Fig. 2.10D; equivalent to state M in Fig. 2.10B) then sandwiched at the NBD dimer interface, so signaling the channel gate to open (state O1) until the ATP at the competent, consensus, site is hydrolyzed (state O2), which leads to dimer disruption, nucleotide exchange, and channel closure (state C).

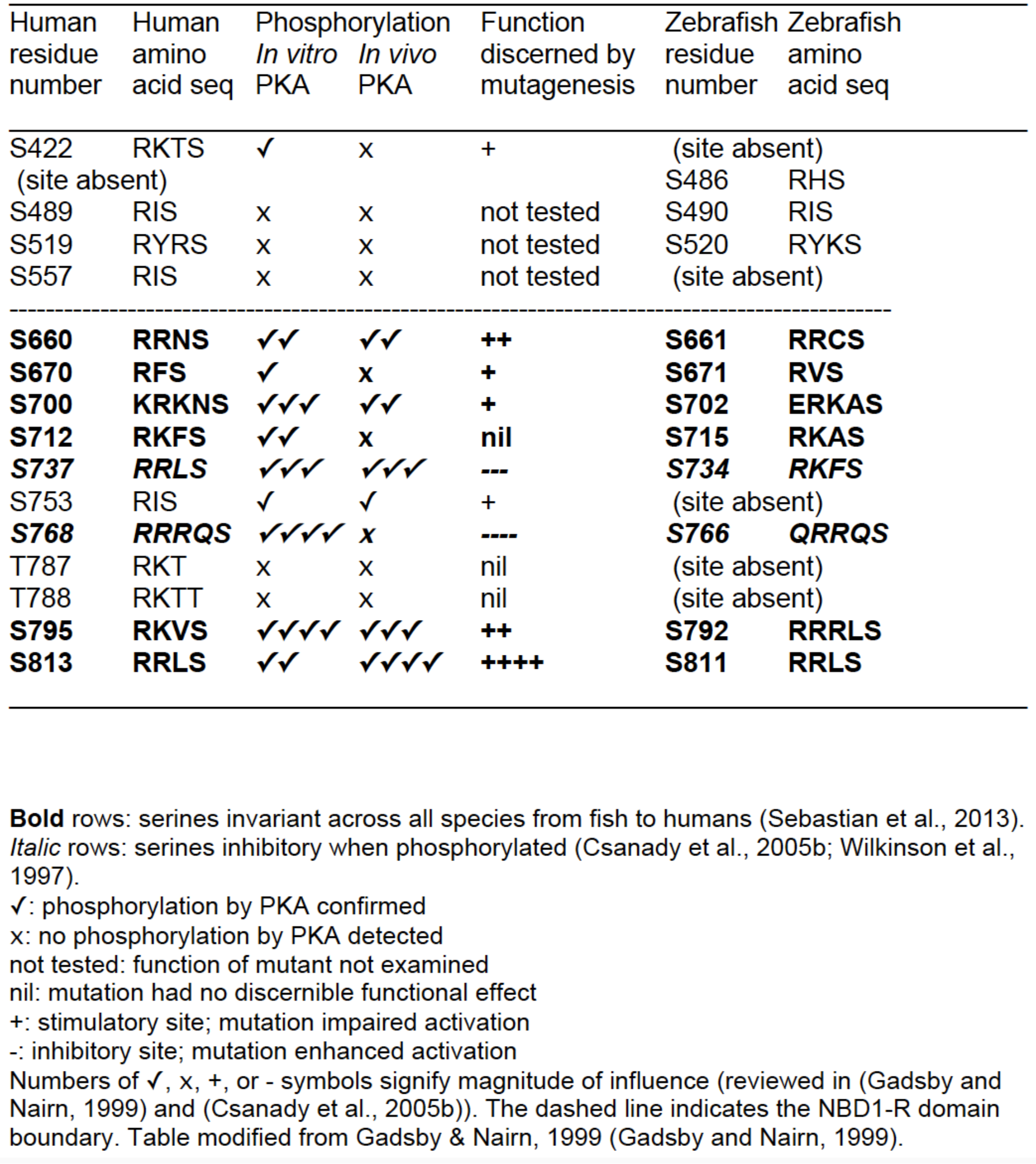
A 20-s segment of the record (gray bar, Fig. 2.10C) is shown magnified, with individual closed periods and open bursts marked. Each green bar measures time the channel spent in closed state C (interburst duration, τ_{ib}), and each red bar measures time spent in an open burst, in state O1 and then O2 (open burst duration, τ_b). Because >95% of all human CFTR channel openings are terminated by hydrolysis of a single ATP [76], on average the time for completion of the gating cycle, i.e., $\tau_{ib} + \tau_b$, which was 1.91 ± 0.5 s for nine fully phosphorylated CFTR channels (including the one in Fig. 2.10C) must be identical to that for completion of the ATPase cycle. In other words, the mean gating cycle rate, $1/(\tau_{ib} + \tau_b)$, ~0.52 s^{-1} for these channels in saturating ATP, at 25°C, should match the maximal ATPase turnover rate determined under comparable conditions. ATPase rates of dephosphorylated and fully phosphorylated (Fig. 2.2) purified human CFTR over a range of ATP concentrations at 28°C reveal strong activation (~6-fold) by PKA (Fig. 2.10E), as similarly observed for purified zebrafish CFTR [39]. The maximum ATPase rate of phosphorylated human CFTR, 134 nmol/mg/min (Fig. 2.10E), corresponds to a maximal cycle turnover rate of 0.37 s^{-1}, assuming that 100% of the purified protein is active in detergent. Given the variability of estimates of human CFTR channel gating cycle rates in 300–550 nM PKA, 2–5 mM ATP, at 21°C–26°C (0.47 s^{-1} [76], 0.45 s^{-1} [77], 0.53 s^{-1} [75], 0.72 s^{-1} [74], 0.62 s^{-1} [76]), we can conclude that up to 84% of the purified human protein retains full activity.

Fig. 2. 10 PKA phosphorylation stimulates CFTR channel gating and ATPase activity
(A) Macroscopic current in an excised patch containing hundreds of human CFTR channels exposed to 2 mM MgATP shows only very few channel openings (amplified in inset at left) before phosphorylation but is strongly stimulated along a sigmoid time course by 300 nM PKA catalytic subunit. The red line is a fit to the two-step activation scheme in (B) yielding the slower, k_{a1}, and faster, k_{a2}, rates indicated. PKA withdrawal allows endogenous phosphatases to partially dephosphorylate channels, diminishing the current in ATP. (B) Cartoon model of two-step activation process of dephosphorylated ("D") channels. Infrequent, reversible, disengagement of the R domain (red tongue) from its inhibitory position slowly (with rate k_{a1}) renders channels phosphorylatable ("P"). Relatively rapid (rate k_{a2}) phosphorylation by PKA (indicated by red P on tongue) temporarily traps the R domain in a permissive position, rendering the channels maximally activated ("M") and allowing maximal channel gating via the cycle in (D). TMDs, gray; NBD1, blue; NBD2, green; R domain, red; ATP, yellow. Although steps D-to-P and P-to-M are both presumed reversible (the former spontaneously, the latter by phosphatase-mediated dephosphorylation), fitting the activation time course (in A) with the depicted two-parameter scheme estimates "net" rates of channel progression through the sequence of states D-P-M. (C) Single-channel current recording with addition of 2 mM MgATP and 300 nM PKA as indicated. The thin gray bar beneath the record indicates the segment near the end of the exposure to PKA, magnified below, with green bars marking times, τ_{ib}, the channel spent in the interburst closed state ("C" in D), and red bars marking times, τ_b, spent in open burst states ("O1" and "O2" in D). (D) Cartoon representation of gating cycle of fully phosphorylated CFTR channel. ATP-bound closed channels (state C) open, with rate $k_{CO} = 1/\tau_{ib}$, to a prehydrolytic open state (O1), permitting flow of chloride ions (purple dots).ATP is split to ADP(orange) and Pi at competent catalytic site 2 (rate k_1), leading to dimer disruption (rate k_2) that causes channel closure. The open burst duration, τ_b, comprises time spent sequentially in both O1 and O2 states. (E) ATPase rate at 28°C of purified human CFTR as a function of ATP concentration after PKA phosphorylation (filled circles) or after dephosphorylation with lambda phosphatase (empty squares). The curves show Michaelis-Menten fits, yielding for phosphorylated CFTR, $K_m = 187 \pm 24\ \mu M$, $V_{max} = 134 \pm 4$ nmol/mg/min and for dephosphorylated CFTR, $K_m = 138 \pm 30\ \mu M$, $V_{max} = 23 \pm 1$ nmol/mg/min.



A
ATP
PKA
$k_{a1} = 0.11\ s^{-1}$
$k_{a2} = 0.23\ s^{-1}$
1 pA
10 s
100 pA
10 s
B
D
P
M
P
P
P
k_{a1}
k_{a2}
C
ATP
PKA
0.5 pA
5 s
D
C
O_1
O_2
k_{CO}
k_1
k_2
P
P
P
E
Phosphorylated
Dephosphorylated
ATPase Activity (nmol/mg/min)
150
100
50
0
0
1
2
3
4
5
[ATP] (mM)

Table 2. 2 Comparison of PKA consensus phosphorylation sites in human vs. zebrafish CFTR



Human residue number	Human amino acid seq	Phosphorylation		Function discerned by mutagenesis	Zebrafish residue number	Zebrafish amino acid seq
		In vitro PKA	*In vivo* PKA			
S422	RKTS	✓	x	+	(site absent)	
(site absent)					S486	RHS
S489	RIS	x	x	not tested	S490	RIS
S519	RYRS	x	x	not tested	S520	RYKS
S557	RIS	x	x	not tested	(site absent)	
S660	**RRNS**	**✓✓**	**✓✓**	**++**	**S661**	**RRCS**
S670	**RFS**	**✓**	**x**	**+**	**S671**	**RVS**
S700	**KRKNS**	**✓✓✓**	**✓✓**	**+**	**S702**	**ERKAS**
S712	**RKFS**	**✓✓**	**x**	**nil**	**S715**	**RKAS**
S737	***RRLS***	***✓✓✓***	***✓✓✓***	***---***	***S734***	***RKFS***
S753	RIS	✓	✓	+	(site absent)	
S768	***RRRQS***	***✓✓✓✓***	***x***	***----***	***S766***	***QRRQS***
T787	RKT	x	x	nil	(site absent)	
T788	RKTT	x	x	nil	(site absent)	
S795	**RKVS**	**✓✓✓✓**	**✓✓✓**	**++**	**S792**	**RRRLS**
S813	**RRLS**	**✓✓**	**✓✓✓✓**	**++++**	**S811**	**RRLS**

Bold rows: serines invariant across all species from fish to humans (Sebastian et al., 2013). *Italic* rows: serines inhibitory when phosphorylated (Csanady et al., 2005b; Wilkinson et al., 1997).

✓: phosphorylation by PKA confirmed

x: no phosphorylation by PKA detected

not tested: function of mutant not examined

nil: mutation had no discernible functional effect

+: stimulatory site; mutation impaired activation

-: inhibitory site; mutation enhanced activation

Numbers of ✓, x, +, or - symbols signify magnitude of influence (reviewed in (Gadsby and Nairn, 1999) and (Csanady et al., 2005b)). The dashed line indicates the NBD1-R domain boundary. Table modified from Gadsby & Nairn, 1999 (Gadsby and Nairn, 1999).

2.2.6 Conservational analysis of CFTR along the revolution

The ABC transporter family includes thousands of proteins found in all living species, out of which only CFTR is known to function as an ion channel. Almost all other ABC transporters are active pumps that translocate substrates against their electrochemical gradients [1]. We ask if there are any special structural features that render CFTR a channel instead of a pump. First, we analyzed sequence conservation among 11 functionally characterized CFTR orthologs [14, 39, 78-86]. The pairwise sequence identity between human CFTR and each ortholog used in this analysis varies from 55% (zebrafish) to 92% (pig). Highly conserved residues are found throughout the pore, the cytosolic extensions of the TM helices, and the NBDs (Fig. 2.11A). The degree of sequence conservation correlates well with the mutations that cause cystic fibrosis (https://cftr2.org). Thus, the majority of the 53 missense mutations (Fig. 2.11A, yellow spheres) occur within highly conserved positions, underscoring the importance of these regions in maintaining the structural and functional integrity of CFTR. Next, we extended this analysis to the ABCC subfamily, which includes members that are closely related in sequence but diverse in function. In addition to CFTR, the ABCC subfamily includes nine multidrug-resistance proteins (MRP1-9) that are ATP-driven efflux pumps and two sulfonylurea receptors (SUR1 and SUR2) with no known transport substrates [87]. The SUR proteins regulate the activity of potassium channels in the context of complexes known as KATP channels [88, 89]. When mapped onto the structure of CFTR, conserved residues in this subfamily are mainly found in the cytosolic region of the protein, such as the NBD catalytic cores, the surfaces of the intracellular vestibule, and the NBD/TMD interfaces (Fig. 2.11B). Sequences in the TM region are highly variable, consistent with the distinct functions of CFTR, MRP, and SUR. These results also suggest that ATP binding to the different ABCC proteins would induce a similar dimerization of the NBDs and coming together of the intracellular extensions of the TMDs to create interactions like those depicted in Fig. 2.9B.

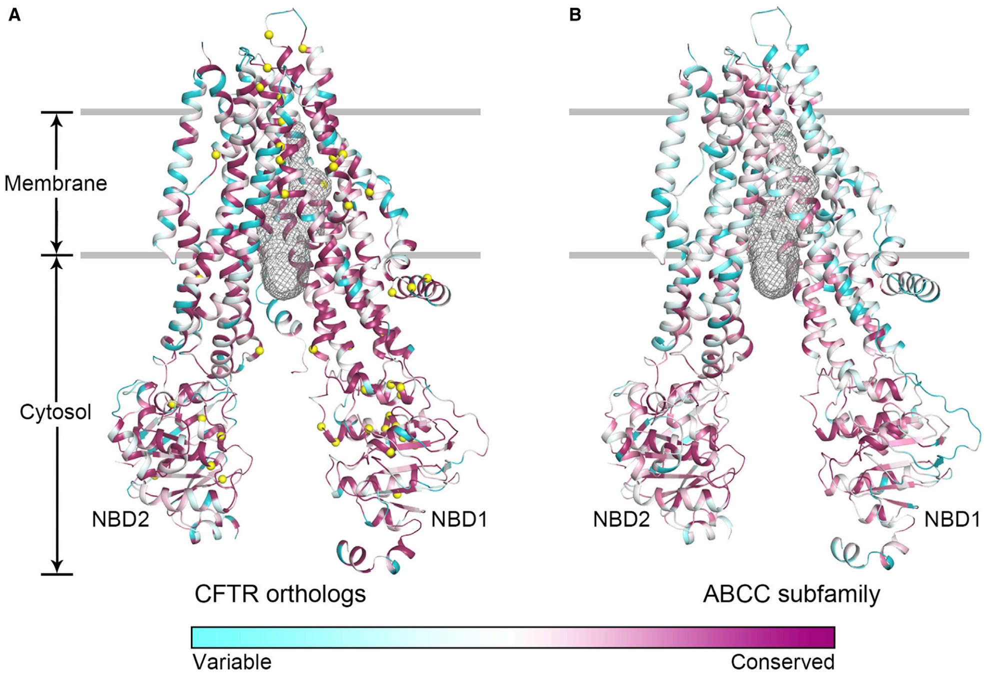


Fig. 2. 11 Sequence conservation analysis
(A) Sequence conservation analysis of 11 CFTR orthologs. The pairwise sequence identity of human CFTR and its orthologs are: human (100%), pig (92%), rabbit (92%), ferret (92%), sheep (91%), chicken (80%), mouse (78%), Xenopus (77%), shark (71%), killifish (58%), and zebrafish (55%). The locations of the 53 missense cystic fibrosis-causing mutations are indicated by yellow spheres. (B) Sequence conservation of the 12 human ABCC proteins mapped onto the structure of CFTR. The R825–843 segment is omitted as it is only found in CFTR.

2.2.7 Structural comparison between CFTR and MRP1

To understand the structural differences between a channel and pump in the same ABCC subfamily, we directly compared the structures of CFTR and MRP1 [90] (PDB: 5UJA). One half of CFTR, comprising the lasso motif, TM helices 1, 2, 3, 6, 10, 11, and NBD1, superimposes very well with the corresponding regions of MRP1 (Fig. 2.12A). The structure of the other half of CFTR is also similar to that of MRP1, except for TM helices 7 and 8 (Fig. 2.12B). In MRP1 both TM7 and TM8 are continuous helices, typical of other ABC transporters. In CFTR, however, TM8 breaks into three short helices and TM7 is displaced from its usual position (Fig. 2.12B and Fig. 2.12C). A mutation within the helical break, L927P, causes cystic fibrosis, possibly by altering the flexibility of TM8. Helical breaks inside the low dielectric core of the membrane can create potential ion binding sites [69, 91] and flexible hinges to facilitate gating. As TM8 lines the pore, it could play a critical role in ion conduction and gating. This unique feature of CFTR is likely the structural basis for its channel function.

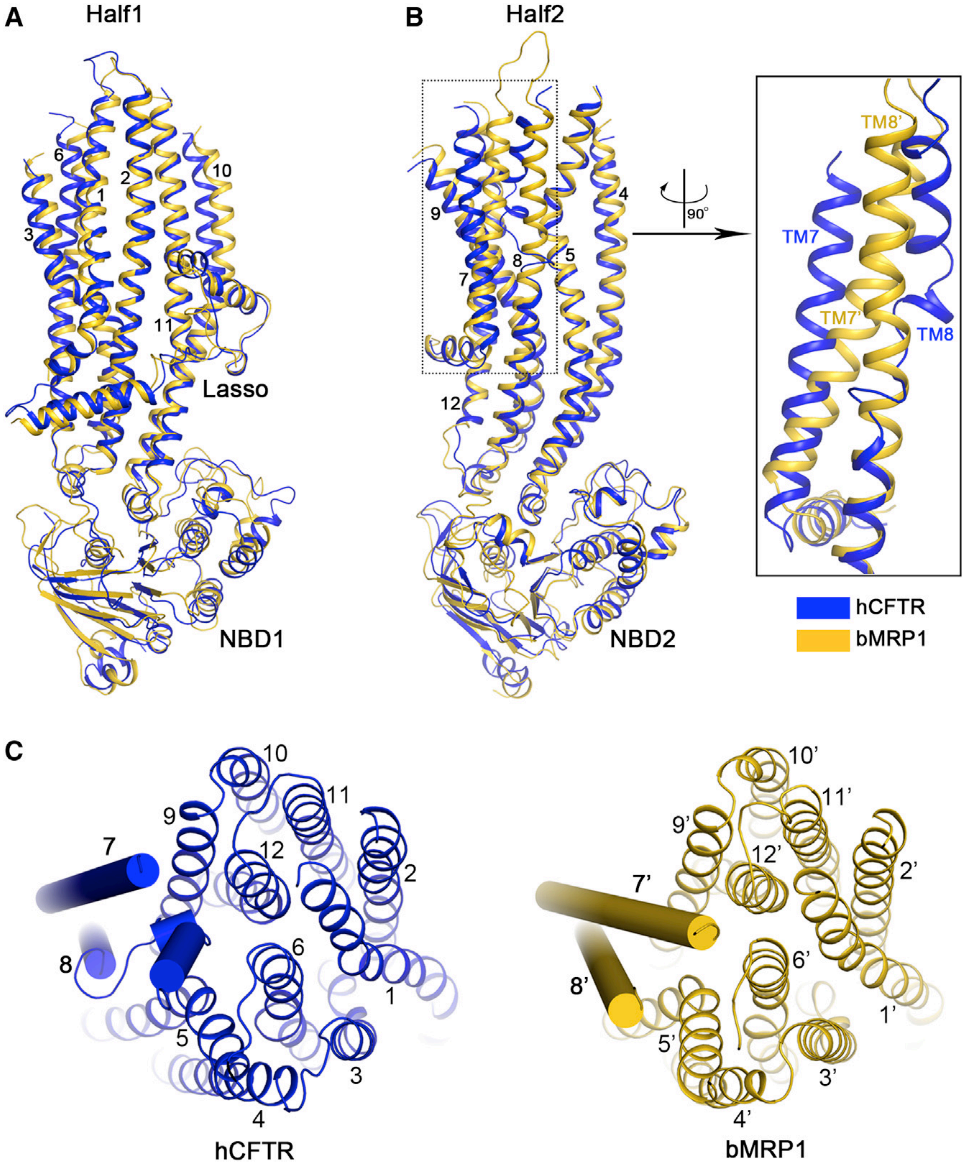


Fig. 2. 12 Structural comparison of CFTR and MRP1

Human CFTR (hCFTR) is shown in blue and bovine MRP1 (bMRP1, PDB: 5UJA) in yellow. Because bMRP1 contains an additional TMD0 domain (not shown), helices are numbered based on the CFTR structure. (A) Superimposition of half1, which comprises the lasso motif, 6 TM helices, and NBD1. (B) Superimposition of the other half. Also shown is a zoom-in view of TM7 and TM8. (C) The 12 TM helices viewed from the extracellular side of the membrane. TM7 and TM8 are shown in cylinder and other helices in ribbon.

2.3 Discussion

In summary, we presented here the cryo-EM structure of human CFTR determined in the dephosphorylated, ATP free conformation. The structure was analyzed to further establish the architecture of the ion pore, the functions of many positively charged residues, and to identify the structural basis for CFTR's channel activity. We showed how the observed location of the R domain informs understanding of CFTR's slow activation by phosphorylation, as well as of infrequent cycling by unphosphorylated CFTR. We also analyzed the relationship between ATP hydrolysis and channel gating, illuminating how these two events correlate in a CFTR duty cycle.

2.4 Material and methods

2.4.1 Expression and purification of recombinant CFTR protein

Human CFTR was expressed and purified with a process similar to that described for zebrafish CFTR [39]. Specifically, human CFTR gene with a C-terminal GFP tag was cloned into a pEG BacMam expression vector. P1, P2 and P3 baculovirus was made in Sf9 cells. Human CFTR protein was expressed in HEK293S GnTI$^-$ cells by infecting with 10% (w/w) P3 baculovirus [92]. To purify the protein, cells were first solubilized in lysis buffer for 2 hr at 4°C.

Lysis buffer	
Tris-HCl pH 7.5	20 mM
$MgCl_2$	2 mM
NaCl	200 mM
Glycerol	20% (v/v)
2,2-didecylpropane-1,3-bis-b-D-maltopyranoside (LMNG)	1% (w/v)
Cholesteryl hemisuccinate (CHS)	0.2% (w/v)
Protease inhibitors	1 mg/ml Leupeptin, 1 mg/ml Pepstatin, 1 mg/ml Aprotonin, 100 mg/ml Trypsin inhibitor, 1 mM Benzamidine, and 1 mM PMSF
DNase I	3 mg/ml

The supernatant was collected after centrifugation at 75,000 g for 1 hr. GFP nanobody (GFPnb) coupled CNBR-activated Sepharose beads (GE Healthcare) were used to affinity purify the protein. GFPnb beads were washed with buffer A.

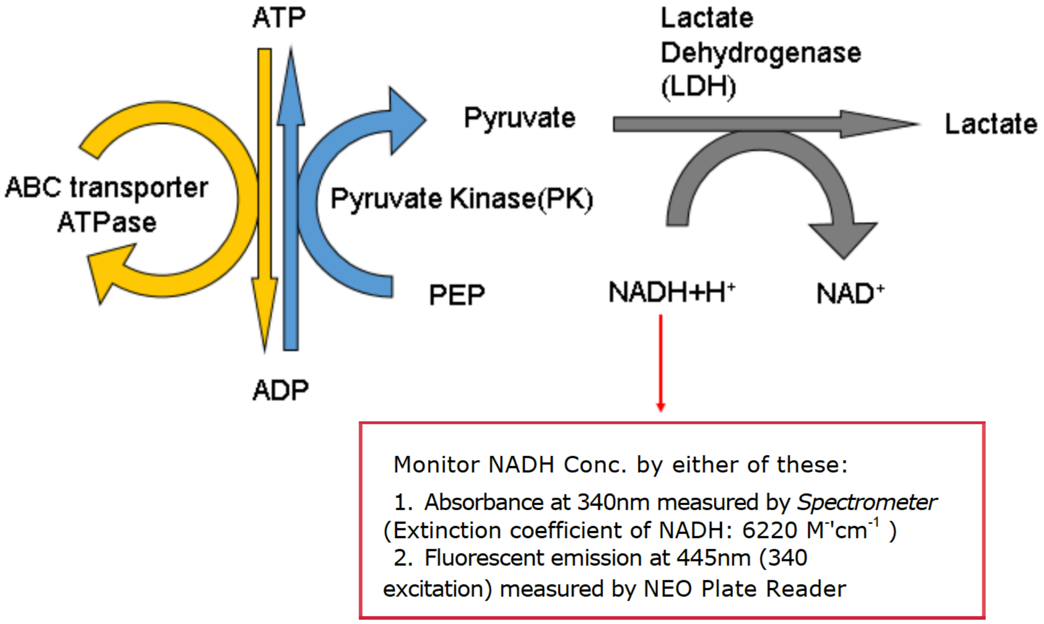
Buffer A	
Tris-HCl pH 7.5	20 mM
$MgCl_2$	2 mM
NaCl	200 mM

DTT	2 mM
Digitonin	0.06%

The protein was eluted with buffer A after incubation with PreScission protease (10:1 w/w ratio) overnight and was further purified by size exclusion chromatography using a Superose 610/300 column (GE Healthcare).

2.4.2 ATPase activity assay

The ATPase activity was measured using an ATP regeneration-NADH consumption coupled system [93], described as below.



Monitor NADH Conc. by either of these:

1. Absorbance at 340nm measured by *Spectrometer* (Extinction coefficient of NADH: 6220 $M^{-1}cm^{-1}$)
2. Fluorescent emission at 445nm (340 excitation) measured by NEO Plate Reader

Pyruvate kinase (PK; Sigma 10128155001), lactate dehydrogenase (LDH; Sigma 10127230001), NADH (Sigma: 10107735001), and phosphoenolpyruvate (PEP; Sigma: 10108294001) were purchased from Roche. Purified CFTR was treated with lambda PP (NEB) or PKA (NEB) for 1 hr followed by gel filtration to remove excess ATP and enzyme. The reaction was carried out in a reaction buffer

Reaction Buffer	5mL
HEPES pH 8.0	50 mM
KCl	150 mM
Digitonin	0.06%
DTT	2 mM
$MgCl_2$	2 mM
PK (10 mg/mL)	30 μL
LDH (5 mg/mL)	32 μL
PEP powder	11 mg
NADH (50mM stock)	15 μL

Because of the light sensitivity of NADH, the reaction buffer was kept in the dark before the assay was initiated. Reaction was carried out in 30 μL aliquots in triplicate with a final CFTR concentration of 0.2 μM and different concentrations of $MgCl_2$/ATP. $MgCl_2$/ATP was prepared in 100 mM stock concentration and the concentration was calibrated according to the extinction coefficient of ATP (15.4 x 10^3 (absorbance units M^{-1})). To initiate the reaction, 3 μL of ATP was added to 30 μL aliquots (Final ATP concentrations are 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 4 mM). NADH fluorescence depletion was recorded continuously for 45 min by an Infinite M1000 Microplate Reader (TECAN) with excitation wavelength of 340 nm and emission wavelength of 445 nm. To convert NADH depletion to ATP depletion, a standard curve of NADH was made with concentrations of 25, 50, 75, 100, 125, 150, 175, 200 μM NADH.

2.4.3 Isolation and injection of *Xenopus laevis* oocytes

Xenopus laevis oocytes were removed from anesthesized frogs by ovarian dissection following an IACUC-approved protocol, and the follicular layer removed by collagenase treatment. Oocytes injected with cRNA were incubated at 18°C in a modified Ringer's solution supplemented with 1.8 mM $CaCl_2$ and 50 mg/ml gentamycin.

2.4.4 Electrophysiology

Human CFTR cDNA (pGEMHE-CFTR) was transcribed in vitro using T7 polymerase (Ambion, mMessage T7), and 0.1-10 ng cRNA were injected into Xenopus laevis oocytes. Current recordings were made at 25°C in inside-out patches excised from the oocytes 1-3 days after injection.

Patch pipette solution contained (in mM): 136 NMDG-Cl, 2 $MgCl_2$, 5 HEPES, pH 7.4 with NMDG. The continuously flowing bath solution could be exchanged with a time constant < 100 ms and contained (in mM): 134 NMDG-Cl, 2 $MgCl_2$, 5 HEPES, 0.5 EGTA, pH 7.1 with NMDG. MgATP (2 mM) was added from a 400-mM aqueous stock solution (pH 7.1 with NMDG). 300 nM catalytic subunit of PKA (Sigma) was applied to activate CFTR channels. Macroscopic and unitary currents were recorded at a membrane potential of -80 mV (Axopatch 200B, Molecular devices), filtered at 2 kHz, and digitized at 10 kHz.

Single-channel patches were identified from long (typically 10 min) recordings without superimposed channel openings. To reconstruct open bursts and interbursts, of durations τ_b and τ_{ib} respectively, currents from patches containing no superimposed channel openings were baseline subtracted to remove slow drifts, refiltered at 100 Hz, and idealized by half-amplitude threshold crossing. After imposition of a fixed dead time of 4 ms, the resulting events lists were subjected to kinetic analysis as described [71]. Open bursts were isolated by suppressing brief closures shorter than a cutoff estimated using the method of Magleby and Pallotta [94], as described [71].

Macroscopic current activation time courses were fit to a two-step two-parameter activation scheme, $D \rightarrow P \rightarrow M$, using non-linear least-squares methods (pCLAMP 9, Molecular Devices), yielding net sequential activation rates, k_{a1} and k_{a2} (Fig. 2.10B).

2.4.5 EM sample preparation and data collection

The purified human CFTR sample was concentrated to 5.5 mg/ml and supplemented with 3 mM fluorinated Fos-Choline-8. 3 μL of the protein sample was applied to each Quantifoil R1.2/1.3 400 mesh Au holey carbon grid (Quantifoil), blotted with Vitrobot Mark IV (FEI) for 3 s at blot force 0 after 15 s wait time, and frozen in liquid ethane. A 300 kV Titan Krios (FEI) with a Gatan K2 Summit direct electron detector (Gatan) was used to image the grids. Images were recorded by Serial EM in super-resolution mode with a pixel size of 0.4085 Å/pixel. For data collection, each image was exposed for 7 s in 50 subframes with a dose rate of 8 electrons per pixel per second. A defocus range from -0.8 μm to -2.5 μm was used.

2.4.6 Imaging Processing and 3D reconstruction

Dark subtracted images were first normalized by gain reference and binned by 2x, resulting in a pixel size of 0.817 Å/pixel. Drift correction was performed using the program Unblur [95]. The contrast transfer function (CTF) was estimated using CTFFIND4 [96]. To generate templates for automatic picking, around 5000 particles were manually picked and classified by 2D classification in RELION [58]. After automatic picking and manual micrograph inspection, 415,915 particles were extracted for subsequent 2D and 3D classification. Motion correction of individual particles was performed using the program alignparts_lmbfgs [97]. Using the structure of zebrafish CFTR (PDB: 5UAR) low pass filtered to 60 Å as the initial model, 3D classification was carried out in RELION [58]. The best class, containing 33% of the particles, gave rise to an 8.26 Å map. The refinement was performed in FREALIGN [59] using this best class as the initial model and all particles post motion correction to generate the final map for model construction.

2.4.7 Model Construction, refinement and validation

To get better side chain densities for model building, we sharpened one half-map using BFACTOR.EXE (author: Nikolaus Grigorieff) with a resolution limit of 3.9 Å and a b-factor value of -250 Å^2. To build TMD1 and TMD2 (residues 5-389 and 844-1172), we first generated a model based on the structure of zebrafish CFTR (PDB: 5UAR) using the program Modeler [98], then fitted and rebuilt the model into the cryo-EM map in COOT [99]. For the NBDs, the crystal structures of human NBD1 (PDB: 2PZG) and human NBD2 (PDB: 3GD7) were used as initial models. For regions with good densities, the residues were adjusted based on the map. For regions with weak densities, the residues were kept as they were in the models. The overall densities for the R domain were too weak to build; however, the C-terminal region of the R domain, which connects to elbow helix 2 of TMD2, had continuous densities. Based on the secondary structure prediction by Phyre2 [100], this region contains an α-helix (residues 830-841). Therefore, we modeled a poly-alanine helix into the tubular density. The final structure includes residues 5-389 of TMD1; 390-402 and 439-645 of NBD1; 844-883 and 909-1172 of TMD2; 1207-1436 of NBD2; and 19 alanines that likely correspond to residues 825-843 of the R domain. The model was refined in real space using PHENIX [101] and also in reciprocal space using Refmac [102, 103]. Structural factors were calculated from a half-map (working) using the program Sfall [104]. The R work and R free values were calculated with a mask containing the model plus a 2 Å margin. Fourier shell correlations (FSCs) were calculated between the two half maps, the model against the working map, the other (free) half map, and full (sum) map [105, 106]. Local resolutions were estimated using Blocres [57]. MolProbity [107, 108] was used to validate the geometries of the model. Structure figures were genereated using Pymol (http://www.pymol.org), Chimera [109], and HOLE [63].

2.4.8 Sequence conservation analysis

Sequence alignments were carried out in ClustalX [110] and the conservation analysis was performed using the ConSurf Server [111].

CHAPTER 3: Cryo-EM structure of phosphorylated, ATP-bound zCFTR

3.1 Summary

The cystic fibrosis transmembrane conductance regulator (CFTR) is an anion channel evolved from an ATP-binding cassette transporter. CFTR channel gating is strictly coupled to phosphorylation and ATP hydrolysis. In the previous chapter, we described essentially identical structures of zebrafish and human CFTR in the dephosphorylated, ATP-free form. In this chapter, we present the structure of zebrafish CFTR (zCFTR) in the phosphorylated, ATP-bound conformation, determined by cryo-electron microscopy to 3.4 Å resolution. Comparison of the two conformations shows major structural rearrangements leading to channel opening. The phosphorylated regulatory domain is disengaged from its inhibitory position; the nucleotide-binding domains (NBDs) form a "head-to-tail" dimer upon binding ATP; and the cytoplasmic pathway, found closed off in other ATP-binding cassette transporters, is cracked open, consistent with CFTR's unique channel function. Unexpectedly, the extracellular mouth of the ion pore remains closed, indicating that local movements of the trans-membrane helices can control ion access to the pore even in the NBD-dimerized conformation.

This chapter (including this summary) is adopted from my work as 2nd author with Dr. Zhe Zhang and Prof. Jue Chen, published in *Cell.* 2017 [112].

3.2 Results

3.2.1 Structure determination

Previous work in the field has shown that non-hydrolysable ATP analogs or mutations that eliminate ATP hydrolysis prolong the lifetime of CFTR's open channel state, presumably by stabilizing the NBD dimer [21, 113, 114]. To obtain the structure of zCFTR in an ATP-bound, NBD-dimerized conformation, we replaced the general base residue (E1372 in zCFTR, E1371 in human CFTR) in the consensus site with a glutamine. This mutation abolishes ATP hydrolysis and slows the pore-closing rate by ~1,000-fold [21]. The zCFTR E1372Q mutant was overexpressed in human embryonic kidney cells. Purified protein was incubated with PKA to phosphorylate the R domain. For wild-type CFTR, this procedure increases the ATPase activity by 6- to 8-fold [39, 56]. Upon removal of PKA by the gel filtration column, the phosphorylated sample was mixed with 8 mM ATP-Mg^{2+} and frozen in vitreous ice for imaging.

The quality of the dataset, which contained 777,102 particles, was first assessed by reference-free two dimensional (2D) classification (Fig. 3.1A). Secondary elements are clearly visible in many 2D classes, and the majority of the particles appear to be in the NBD-dimerized conformation (Fig. 3.1A). The final three-dimensional (3D) reconstruction, calculated from all particles, has an estimated resolution of 3.4 Å (Fig. 3.2; Table 3.1), higher than that of the 3.7 Å structure determined in the dephosphorylated, ATP-free form [39] (PDB: 5UAR). The EM map shows strong side-chain density for most residues in the TMDs and the NBDs (Fig. 3.1B and 3.3). Two ATP-Mg^{2+} molecules are well-defined at the binding sites (Fig. 3.1C and 3.1D). In this phosphorylated, ATP-bound structure, the TMDs and NBDs pack closely together to form an elongated molecule (Fig. 3.1E). Unlike the bacterial multidrug transporter Sav1866 [73, 115], the TM helices in CFTR do not split into two wings, creating openings to the membrane; instead, they pack tightly together to shield the ion-conduction pathway from the lipid bilayer (Fig. 3.1E and 3.4). As indicated by biochemical and electrophysiological data [21, 116, 117], the two NBDs form a head-to-tail dimer with two ATP-Mg^{2+} molecules bound along the interface. No density corresponding to the R domain (residues 646–840) or the R insertion (residues 406–436) is visible.

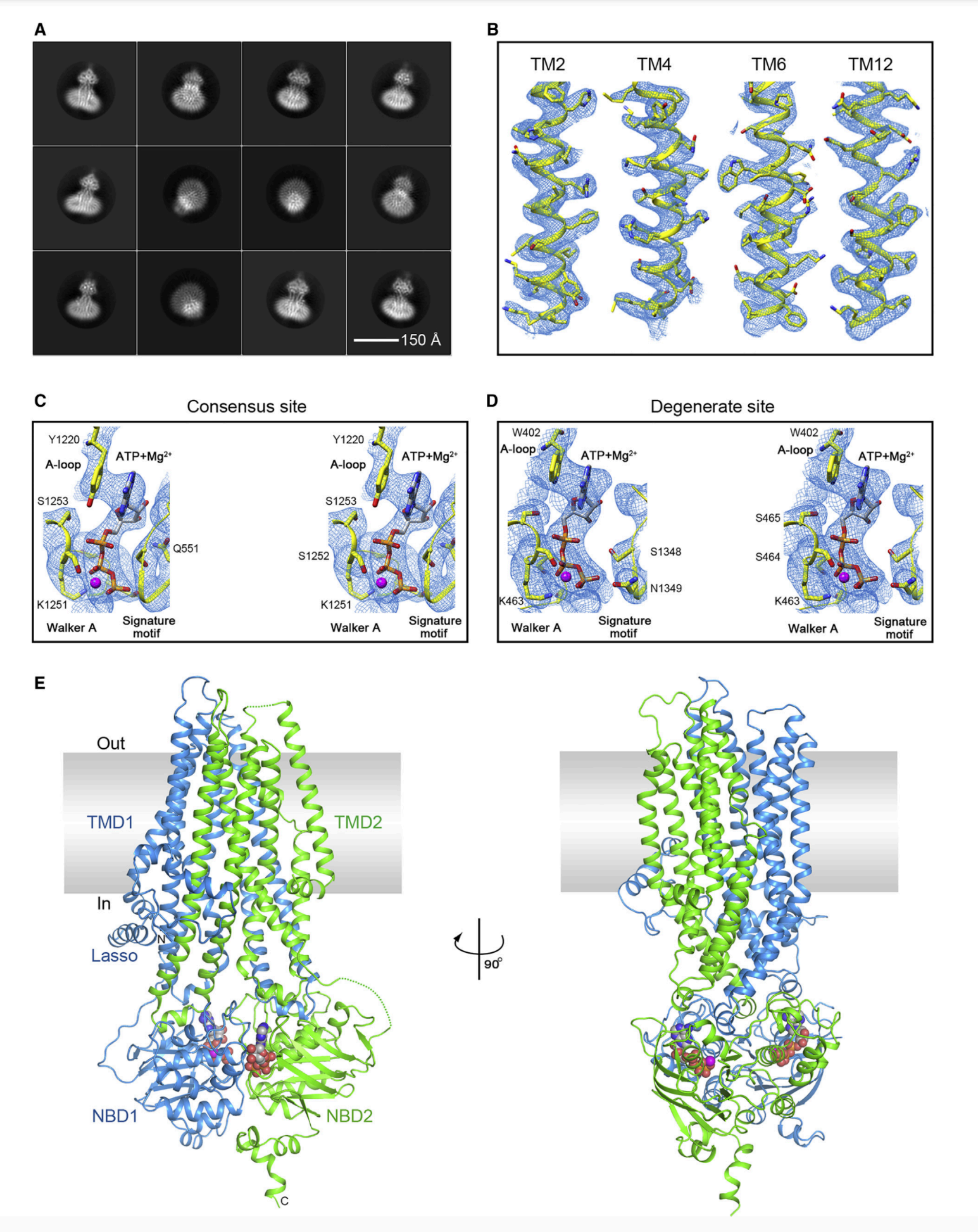


Fig. 3. 1 Cryo-EM structure of zebrafish CFTR in the phosphorylated, ATP-bound conformation

(A) Representative 2D classes from Relion [58]. (B) Representative cryo-EM densities in the transmembrane region. (C and D) Stereo view of EM densities at the consensus (C) and degenerate (D) sites. ATP is shown in sticks with carbon atoms colored silver, nitrogen atoms colored blue, phosphorus atoms colored orange, and oxygen atoms colored red. Mg2+ as a magenta sphere; protein in ribbon. (E) Two views of the overall structure. TMD1 and NBD1 are colored blue, TMD2 and NBD2 are colored green, and ATP and Mg2+ are shown as spheres colored the same as those in (C) and (D).

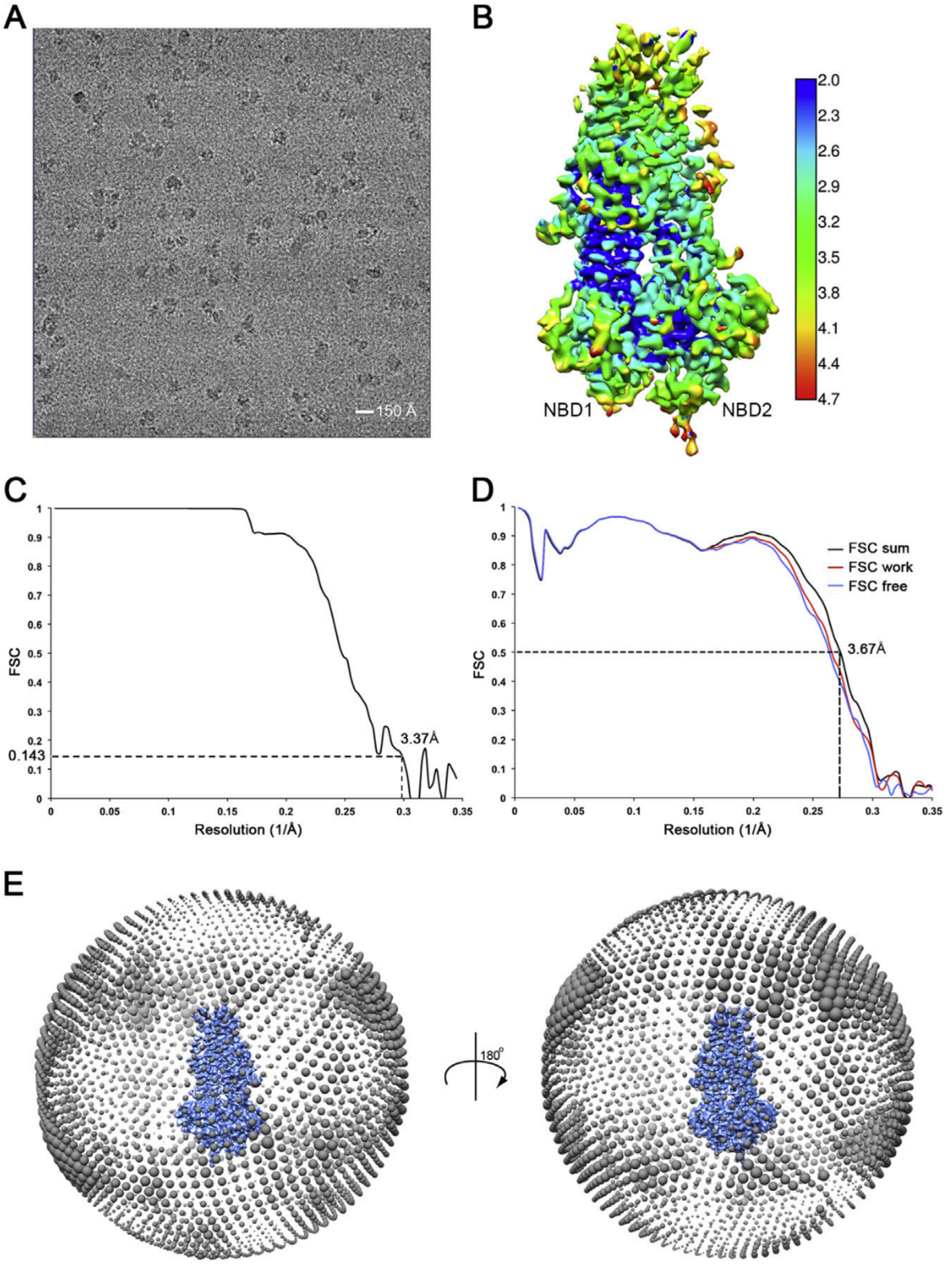


Fig. 3. 2 Assessments of the Cryo-EM Data of NBD-dimerized zCFTR
(A) A representative summed image at the 2.4 mm defocus level. (B) Local resolution distribution estimated by Blocres [57]. (C) Fourier Shell Correlation (FSC) curves between two half-datasets calculated by Frealign [59]. (D) FSC curves between the refined structure and the map calculated from the full dataset (sum, black), the half-map used in refinement (work, red), and the other half-map (free, blue). (E) Angular distributions of all the particles.

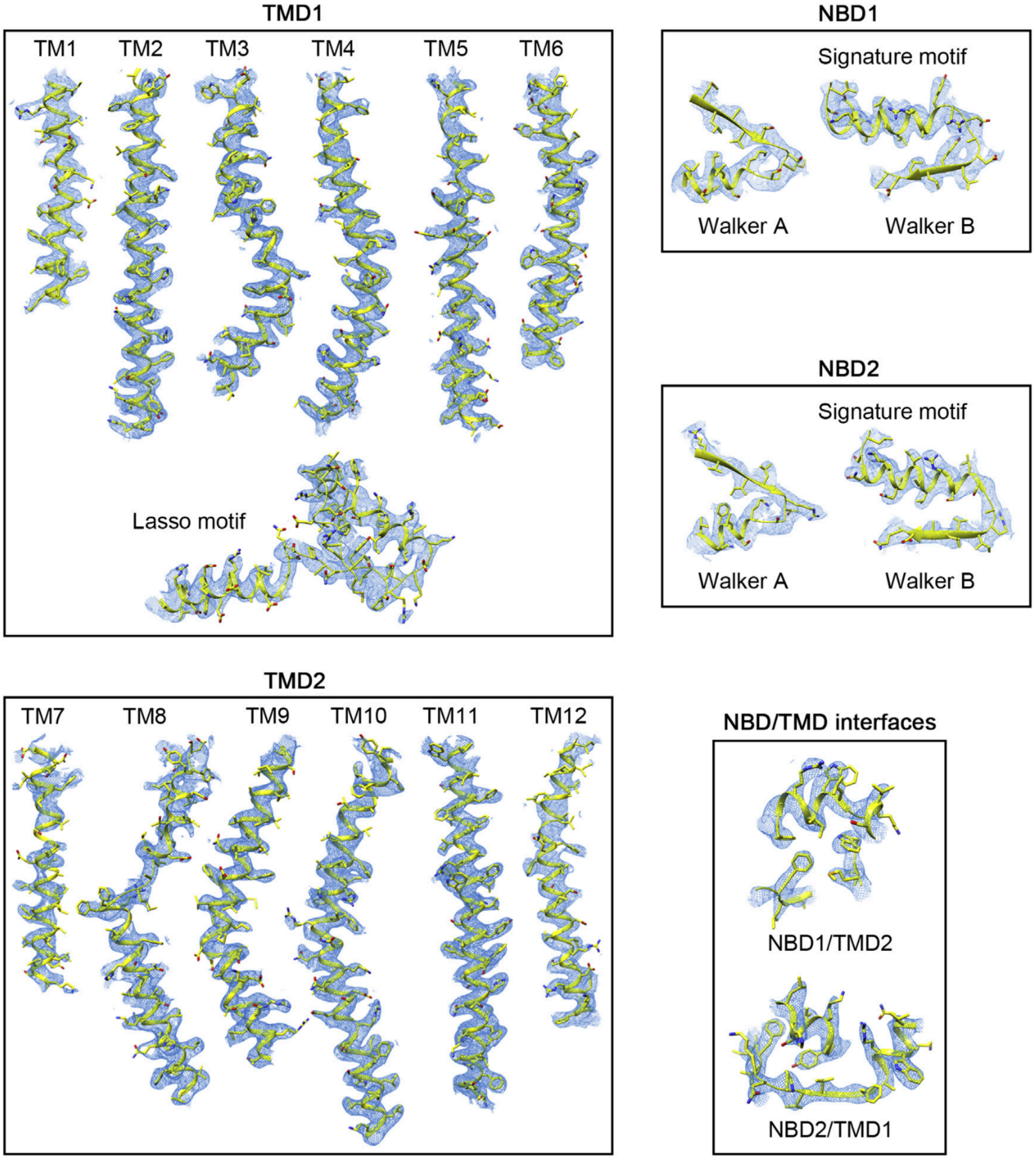


Fig. 3. 3 Cryo-EM Densities of Different Regions of NBD-dimerized zCFTR
B-factor sharpened EM densities for TMDs, NBDs, and the NBD/TMD interfaces.

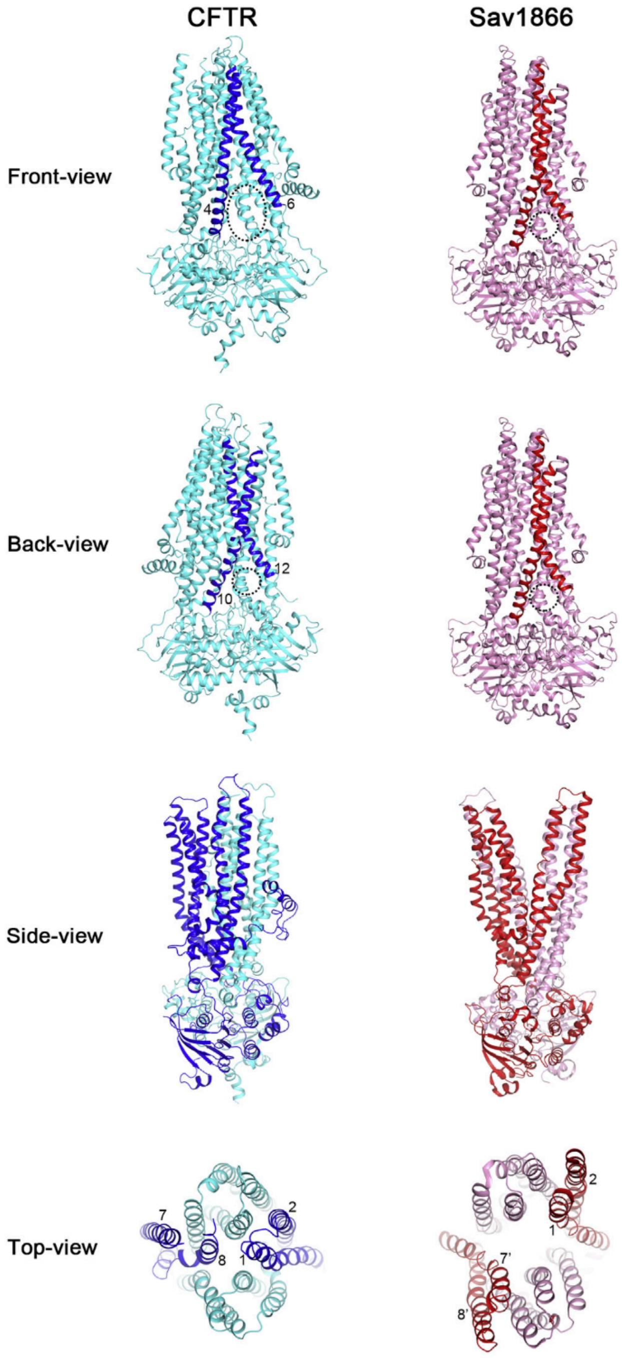


Fig. 3. 4 Structural Comparison of CFTR and Sav1866
The dotted circles illustrate the different degrees of separation between TM 4/6 and TM10/12

Table 3. 1 Summary of EM data and structure refinement statistics of NBD-dimerized zCFTR

Data collection	
Microscope	Titan Krios (FEI)
Voltage (kV)	300
Detector	K2 Summit (Gatan)
Pixel size (Å)	0.82
Defocus range (μm)	0.8 to 2.5
Movies	7,434
Frames/movie	50
Dose rate (electrons/pixel/s)	8
Total dose (electrons/Å^2)	84
Number of particles	777,102
Model composition	
Non-hydrogen atoms	9,393
Protein residues	1,173
ATP	2
Mg^{2+}	2
Refinement	
Resolution (Å)	3.37
Average Fourier Shell Correlation	0.7626
R_{work}	0.2490
R_{free}	0.2575
RMS deviations	
Bond lengths (Å)	0.0071
Bond angles (°)	1.0715
Validation	
Molprobity score	1.24
Clashscore, all atoms	1.47
Favored rotamers (%)	95.2
Ramachandran plot (%)	
Favored	94.9
Allowed	5.1
Outliers	0.0

3.2.2 Major conformational changes upon phosphorylation and ATP binding

The structure of the phosphorylated, ATP-bound CFTR is clearly distinct from that of the dephosphorylated, ATP-free conformation [39, 56] (Fig. 3.5). The R domain, which is no longer inserted in between the two halves of CFTR, has become completely disordered. Disengagement of the R domain permits NBD dimerization, which drives structural rearrangements of the TMDs (Fig. 3.5A and 5B). As a result, the ion conduction pathway has become narrower near the cytosolic side of the membrane and wider in the membrane outer leaflet (Fig. 3.5A). Access to the extracellular medium, however, remains blocked (Fig. 3.5A). The overall chemical environment of the ion conduction pathway has not changed: the pore is lined with positively charged residues in both conformations. The result of the transition between the two conformations can be described mainly as rigid-body displacements of two halves: one half consisting of the lasso motif, TM helices 1–3, 6, 10–11, and NBD1 and the other half consisting of TM helices 4–5, 7–9, 12, and NBD2 (Fig. 3.5B and 3.5C). Due to phosphorylation and ATP binding, both halves have rotated toward the molecular center. The largest displacement, observed at the C-terminal ends of the NBDs, is ~15 Å (Fig. 3.5B). Superimposed onto the global rigid body conformational change, we observe a local conformational change involving the extracellular segments of TM8 and TM12 (residues 918– 931 and 1136–1147) (Fig. 3.5C). The two helical segments move together by changing their angle of orientation to the neighboring helices by ~50^{o} (Fig. 3.5C). The TM8 segment moves away from TM5 and TM7, forming new contacts with TM1 across the pore. Similarly, TM12 also swings toward the other side of the pore to interact with TM1 and TM2 (Fig. 3.5C and 3.6). As a consequence of this local conformational change, the ion pathway becomes occluded near the extracellular end. Had the local conformational change not occurred, a purely rigid-body movement of the two halves would have resulted in an open pore, as shown in the modeled structure (Fig. 3.5D). It is evident that the conformation of the TM8 and TM12 helical segments can control ion access to the pore even when the R domain is phosphorylated, and the NBDs are dimerized.

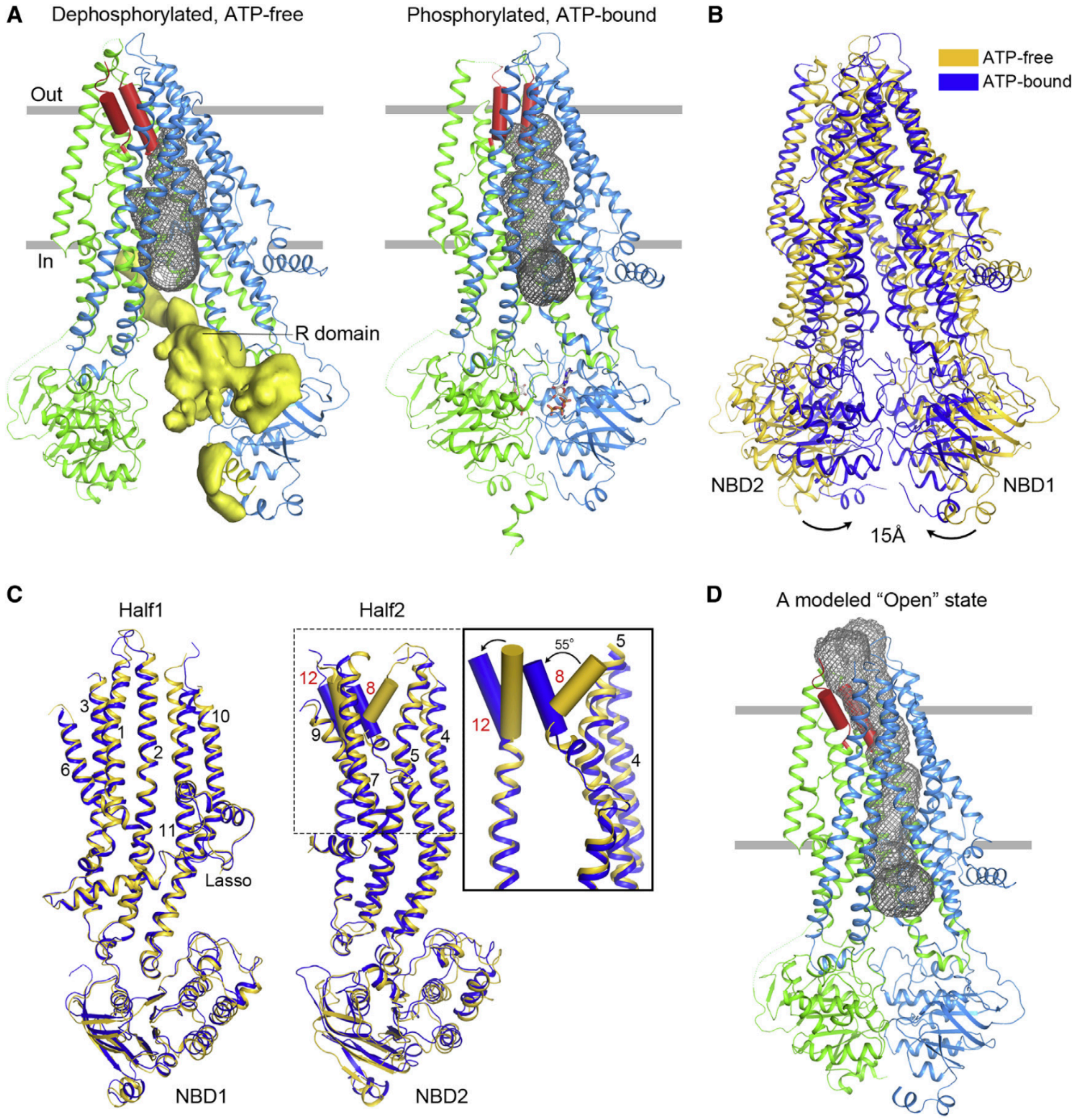


Fig. 3. 5 Conformational changes of zCFTR upon phosphorylation and ATP-binding
(A) The overall structures. The TMDs and NBDs are shown as ribbon and the R domain as a surface. The outer leaflet segments of TM8 (residues 918–931) and TM12 (residues 1136–1147) are highlighted as red cylinders. The gray mesh represents the pore, defined based on the size of a chloride ion (>1.7 Å radius). (B) Superposition of the two structures. The dephosphorylated, ATP-free structure [39] (PDB: 5UAR) is colored gold and the phosphorylated, ATP-bound structure (this study) is colored blue. (C) Superpositions of the two halves individually, together with a zoom-in view. The outer leaflet segments of TM8 and TM12 are shown in cylinders. (D) A modeled NBD-dimerized, open-pore structure based on simple rigid-body movements of the two halves without the local structural changes in TM8 and TM12.

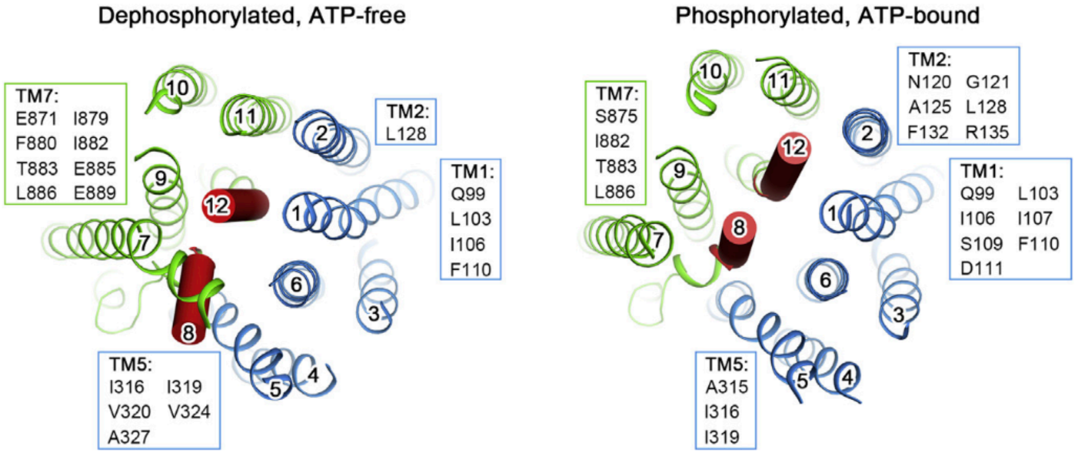


Fig. 3. 6 Local conformational changes of TM8 and TM12

Extracellular view of the transmembrane helices of the dephosphorylated, ATP-free and the phosphorylated, ATP-bound conformations. The residues in TM1, TM2, TM5, and TM7 that are within van der Waals contact distance (<4.5 Å) with the extracellular ends of TM8 (residues 918–931) and TM12 (residues 1136–1147) are listed.

3.2.3 Asymmetric NBD association

The structure of the NBD dimer exhibits the head-to-tail configuration predicted for all ABC transporters (Fig. 3.7A). Statistical analysis of multiple sequence alignments has shown that positions corresponding to R555 and T1246 change in a correlated manner across species, and mutational data have revealed them to be energetically coupled [21, 117]. Indeed, we observe that in ATP-bound CFTR the side chains of these two residues are 3.2 Å apart, compatible with a hydrogen bond interaction (Fig. 3.7B). Residues G550 and G1350, which correspond to the human cystic fibrosis-causing mutations G551D and G1349D, are located at equivalent positions at the dimer interface (Fig. 3.7C). Both residues are within van der Waals contact distances with the Walker A motif of the opposite NBD and the D loop, a highly conserved motif mediating communication between the two ATPase sites [118, 119]. The structure corroborates these disease-associated mutations because aspartate substitution at either position is likely to hinder NBD dimerization, at least in part due to electrostatic repulsion between the acidic side chain and the nucleotides phosphates, thereby causing the gating defects observed in electrophysiology studies [120-122]. Although both ATPase sites contain bound ATP-Mg^{2+}, the degenerate site is not as tightly dimerized as the consensus site (Fig. 3.7A and 3.7D–3.7F), yet still tight enough to preclude access to the signature motif, as observed [123]. When superimposed on the Walker A motifs, the signature motif in the degenerate site is ~3.5 Å further away than that of the consensus site (Fig. 3.7D). At the consensus site, the ATP molecule makes numerous contacts with both NBDs, similar to what has been observed in prokaryotic ABC transporters [115, 124] (Fig. 3.7E). In the degenerate site, ATP is primarily bound to the Walker A/B motifs of NBD1, forming 19 electrostatic and van der Waals contacts. The signature motif in NBD2 contributes only four contacts with ATP (Fig. 3.7F). Several studies have shown that the degenerate site retains ATP throughout multiple gating cycles [125, 126]. One explanation is that the two NBDs stay in close contact at the degenerate site to occlude ATP and prevent its escape [127, 128]. However, this hypothesis contradicts the accessibility analysis showing that the two NBDs separate more than 8 Å as soon as the CFTR channel closes [123]. The other possibility is that NBD1 by itself is able to bind ATP even in the NBD-separated conformation. In this case, the retention of ATP in NBD1 would be due to the maintained bonds to the phosphate chain, resulting from the lack of hydrolysis in the degenerate site. Our new structure seems more compatible with the latter explanation, as ATP remains bound at the degenerate site with the signature motif of NBD2 contributing only 18% of the total contacts.

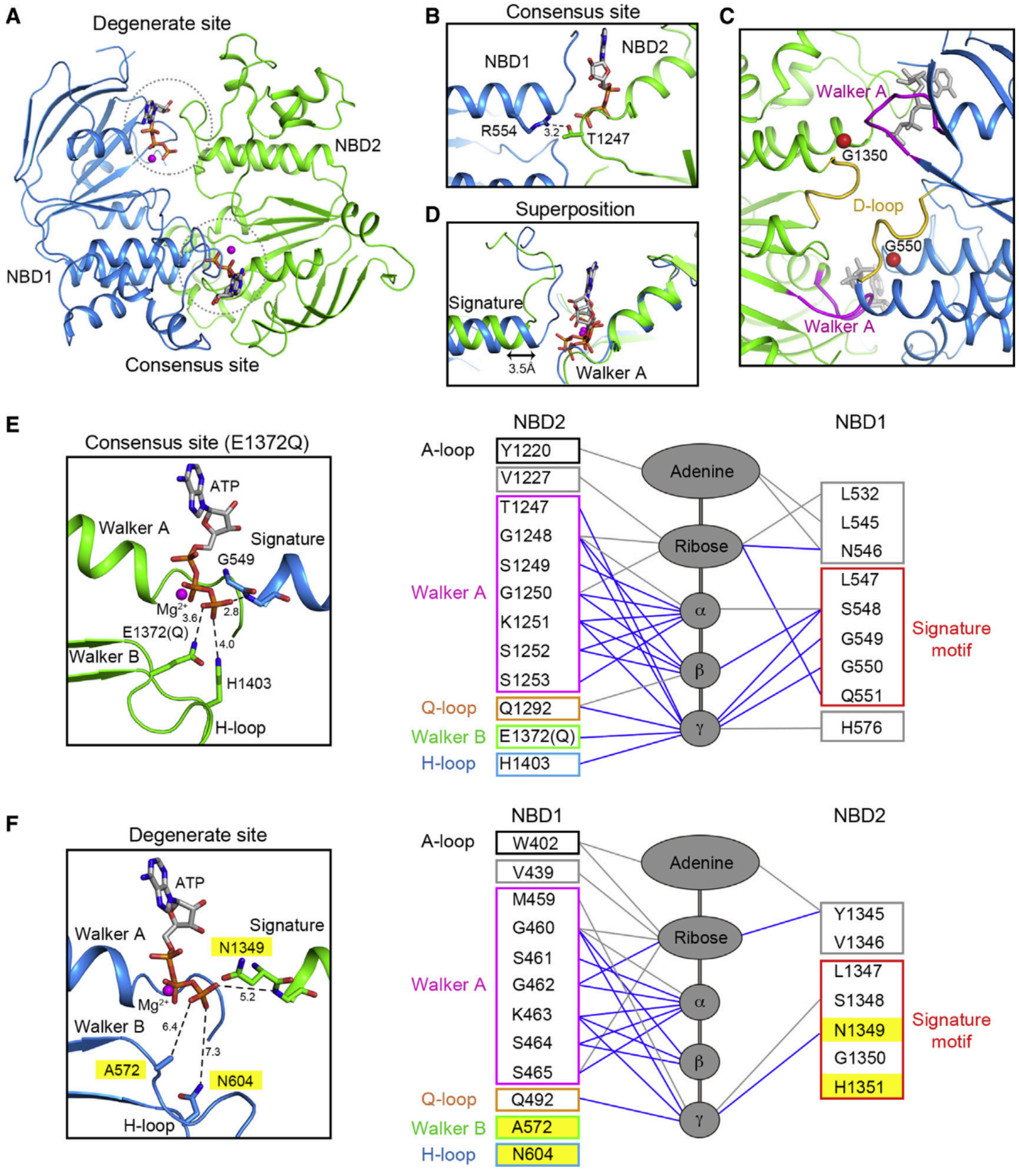


Fig. 3. 7 Molecular details of the NBD dimer

(A) The head-to-tail dimer. (B) The conserved hydrogen bond between R554 and T1247 (corresponding to R555 and T1246 in human CFTR). (C) The positions of G550 and G1350, which correspond to the human cystic fibrosis-causing mutants G551D and G1349D. The Walker A motifs are colored magenta and the D-loops are in gold. (D) The two ATPase sites superpositioned on the Walker A motifs. The signature sequence in the degenerate site (green) is further away from ATP compared to that of the consensus site (blue). (E) ATP binding at the consensus site. (F) ATP binding at the degenerate site. Non-consensus residues are highlighted in yellow. In the schematic diagrams, gray lines represent van der Waals contacts (<4.5 Å), and blue lines correspond to electrostatic interactions (<4.0 Å).

3.2.4 A “broken” intracellular gate

ABC transporters that pump substrates thermodynamically uphill are often described as having two gates, each guarding access to one side of the membrane [129, 130]. ATP binding stabilizes an outward-facing conformation in which the extracellular gate is open, and the intracellular gate is closed. ATP hydrolysis resets the transporter to its inward-facing conformation, in which access to the translocation pathway is reversed. In contrast, channels only need a single gate; when it opens ions flow continuously across the membrane bilayer. CFTR has been described as a “broken” transporter [131, 132] because it is an ion channel that has evolved from an ABC transporter. Can we identify which gate was broken to render CFTR an ion channel? When the NBDs are separated, the closed gate near the extracellular surface of CFTR resembles those of other ABC transporters [39, 56, 66, 133-135]. The intracellular gate in the NBD-dimerized conformation, however, is unique to CFTR. The intracellular gate in a typical ABC transporter, such as Sav1866, is formed by a bundle of TM helices that pack closely in the NBD-dimerized conformation [73, 115] (Fig. 3.4). In CFTR, this helical bundle is cracked open between TM helices 4 and 6 (Fig. 3.8). The ion conduction pathway is connected to the cytosol through this opening. Studies of human CFTR have shown that R248 and K370 are important in recruiting chloride ions to the pore [36]. The corresponding residues in zCFTR, K249, and K371, are both located at the cytoplasmic opening (Fig. 3.8). The same study also showed that positively charged residues along TM10 and TM12, the helices equivalent to TM4 and TM6 in the other half of CFTR, are not involved in comparable favorable electrostatic interactions with chloride ions [36]. Entirely consistent with these functional studies, we did not observe a lateral portal (like that between TM4 and TM6) between TM10 and TM12 (Fig. 3.4).

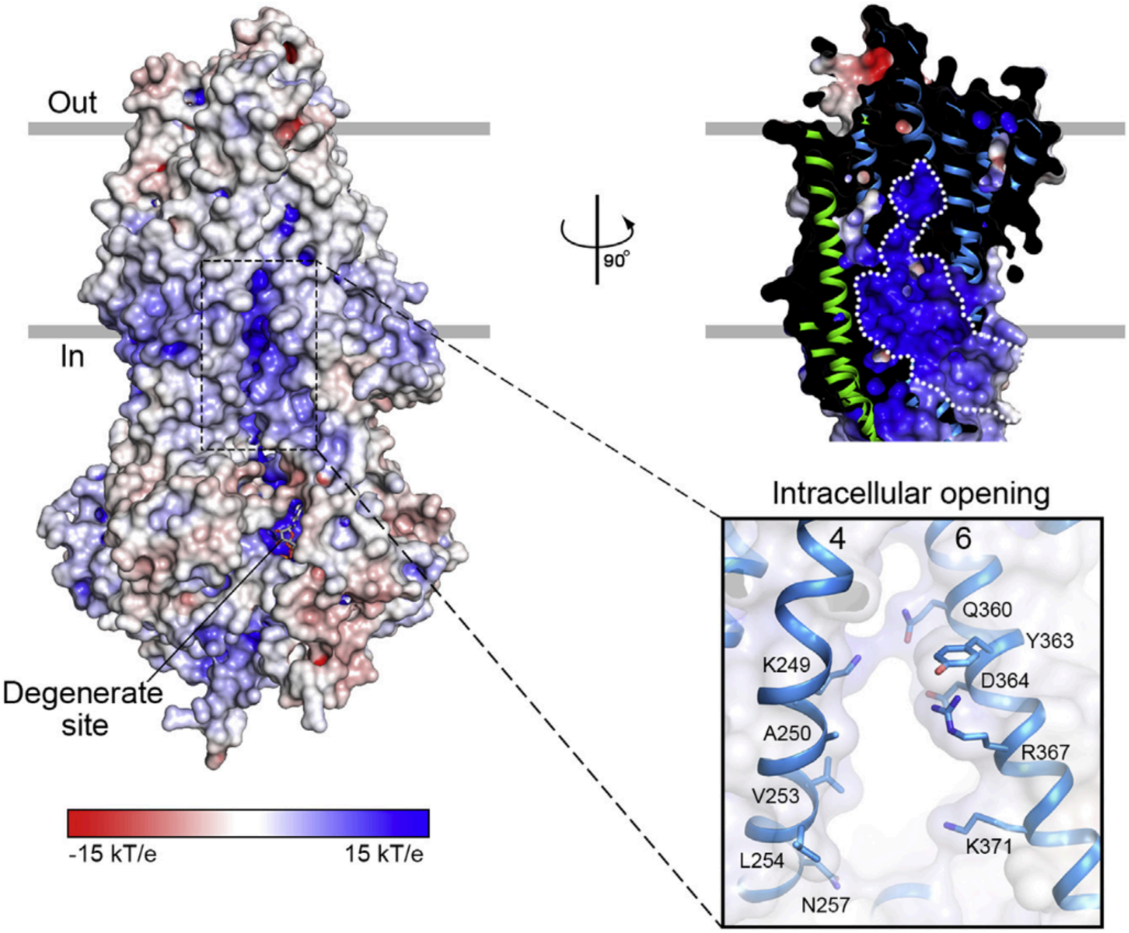


Fig. 3. 8 The ‘‘broken’’ intracellular gate
Electrostatic surface representation showing the lateral opening between TM4 and TM6. Residues lining the intracellular opening are shown with side chains in the zoom-in view. A cross section to illustrate the inner surface of the pore and the lateral opening is also shown.

3.3 Material and methods

3.3.1 Protein expression and purification

The E1372Q mutation was generated using the QuikChange Site-Directed Mutagenesis System (Stratagene). As described previously [39], zebrafish CFTR (E1372Q) with a C-terminal GFP tag was expressed in HEK293S GnTI- cells according to the published protocol (Goehring et al., 2014). ATP (1 mM) and $MgCl_2$ (1 mM) were included in all buffers during purification. The membrane was solubilized with 1% 2,2-didecylpropane-1,3-bis-b-D-maltopyranoside (LMNG) and 0.2% Cholesteryl hemisuccinate (CHS). Affinity purification and buffer exchange were carried out on GFP nanobody-coupled resin to substitute LMNG with 0.06% digitonin. After removing the GFP tag by PreScission protease, the protein was concentrated to 2 mg/ml and incubated with PKA (NEB) (40:1 molar ratio) at room temperature for 1 hr. Further purification was carried out by gel filtration chromatography using a Superose 6 10/300 column (GE Healthcare).

3.3.2 EM sample preparation, data collection, and processing

The protein sample (5.5 mg/ml) was incubated with 8 mM ATP and $MgCl_2$ on ice for 15 min. In addition, 3 mM fluorinated FosCholine-8 were added right before freezing the grids. Data were collected on a 300 kV Titan Krios (FEI) microscope with a K2 summit detector (Gatan) in super-resolution mode. The dose rate was 8 electrons/pixel/second at a physical pixel size of 0.82 Å. Exposures of 7 s per frame were dose-fractionated into 50 subframes. The images were first corrected for gain reference and binned by 2 to reach a physical pixel size of 0.82 Å. Drift correction and contrast transfer function (CTF) estimation were performed using the programs Unblur [95] and CTFFIND4 [96], respectively. After auto-picking in Relion [58], we manually removed low-quality images and false-positive particles. The program alignparts_lmbfgs [97] was used for motion correction of individual particles. 2D and 3D classifications were carried out in Relion [58]. The 3D classification using the ATP-free structure [39] (EMD-8461) as a reference resulted in classes with a similar NBD-dimerized conformation, suggesting structural heterogeneity and model bias were minimal. 3D auto-refinement in Relion produced a 7.5 Å map, which was used as an initial reference for subsequent refinements in Frealign [59]. The final map, calculated from all 777,102 particles, is estimated to be 3.4 Å using the 0.143 cutoff criteria [59].

3.3.3 Model building and refinement

The full dataset was divided randomly into two halves. One half was used for model building and refinement (the working set) and the other half used for validation (the free set). Model building was carried out in Chimera [109] and Coot [99]. Phenix [101] was used to refine the model in real space with secondary structure restraints followed by refinement in reciprocal space using Refmac [102, 103]. A mask with a 2 Å margin was used to calculate the working and free R factors. The FSC curves were calculated with a mask 3.5 times larger than the estimated volume of the model [105, 106]. MolProbity [107, 108] was used for the validation of geometries and Blocres [57] was used for local resolution estimation. The final structure contains residues 1–391 of TMD1, 392–405 and 437–641 of NBD1, 844–886 and 918–1181 of TMD2, 1203–1458 of NBD2, and two ATP and Mg^{2+}. EM data collection and structure refinement statistics were summarized in Table 3.1. Figures were generated using Pymol (http://www.pymol.org), Chimera [109], HOLE [63], and APBS [136, 137].

3.4 Discussion

The structure of zCFTR in the phosphorylated, ATP-bound conformation displays many of the features anticipated for the open channel state, such as disengagement of the R domain from the molecular center, dimerization of the NBDs, and a lateral opening for Cl^- flow to the cytosol. However, the pore remains closed near the extracellular entryway.

Are there functional data supporting the existence of a channel conformation in which the NBDs are dimerized and yet the pore is non-conductive? The answer is "Yes." Detailed kinetic analyses under three different circumstances led to the conclusion that when a CFTR channel opens in response to ATP binding, the NBDs first dimerize and then the pore opens [21, 138]. This sequence of steps implies the existence of an intermediate conformation consistent with the structure observed here. However, this conformation has generally been viewed as a fleeting, high-energy transition state that seems unlikely to have been captured by cryo-EM.

More likely, the structure we have determined corresponds to a state that occurs after ATP has bound, NBDs have dimerized, and the channel has already opened [139, 140]. In electrophysiological records, such a non-con-ducting state appears as brief closures within longer "open burst" periods and is described as an intraburst closure, or flicker closure. Unlike interburst closures that follow ATP hydrolysis, the kinetics of intraburst closure are largely independent of the degree of phosphorylation and the concentration of ATP [77, 141]. The obvious explanation for this phenomenon is that the pore flickers closed and open relatively rapidly while the NBDs are dimerized. The short lifetime of the closures implies that only a small energetic barrier separates them from the open burst state. The conformation we have observed seems to fit this functional behavior. The local conformational change involving the mobile segments of TM8 and TM12 could well explain these brief closures. We observe relatively less well-defined EM densities and higher B factors for the extracellular ends of TM8 and TM12, consistent with these regions being dynamic (Fig. 3.2 and 3.9). Functional studies have also shown that disruption of a conserved salt bridge between TM8 and TM6 (D924-R347 in human CFTR, E932-R348 in zCFTR) introduced brief interruptions of conductance [142], lending further support to the idea that TM8 movement might underlie channel flicker. However, there is one inconsistency in ascribing the conformation we observe to the flicker-closed state. In electrophysiological records, the dwell time for the open state is much longer than that of the flicker-closed state [77]. We would, therefore, have to invoke the possibility that the different environmental conditions of the structural studies (detergent instead of membrane, low temperature) have shifted the equilibrium to favor the flicker-closed state. Alternatively, it is possible that CFTR also undergoes a closure event that is too brief to be resolved in electrophysiological recordings (e.g., microsecond closures would not be resolvable as discrete non-conducting events with a 1 KHz filter). If this were the case, then the measured single channel conductance of ~10 pS [20, 143-146] would actually represent the filtered average value during a higher frequency burst. To give an explicit example, if a 40 pS conductance CFTR channel exhibited closure events with a mean dwell time of 3.0 microseconds separated by open events with a mean dwell time of 1.0 microsecond, then with a 1.0 kHz filter the channel would appear to have a conductance of 10 pS. The local structural rearrangement that would distinguish the cryo-EM structure from the modeled open structure (Fig. 3.5D) could easily occur in the microsecond time domain. This scenario would explain the cryo-EM structure that is closed near the extracellular pore entryway. In conclusion, by comparison with dephosphorylated CFTR in the absence of ATP, this new structure shows that CFTR undergoes major conformational changes after phosphorylation and upon ATP binding. The NBDs have dimerized as predicted by functional studies. The TMDs have moved essentially as rigid bodies, in concert with the NBDs. And, unexpectedly, ATP-bound CFTR also exhibits local conformational changes that are revealing new and unanticipated aspects of this important anion channel.

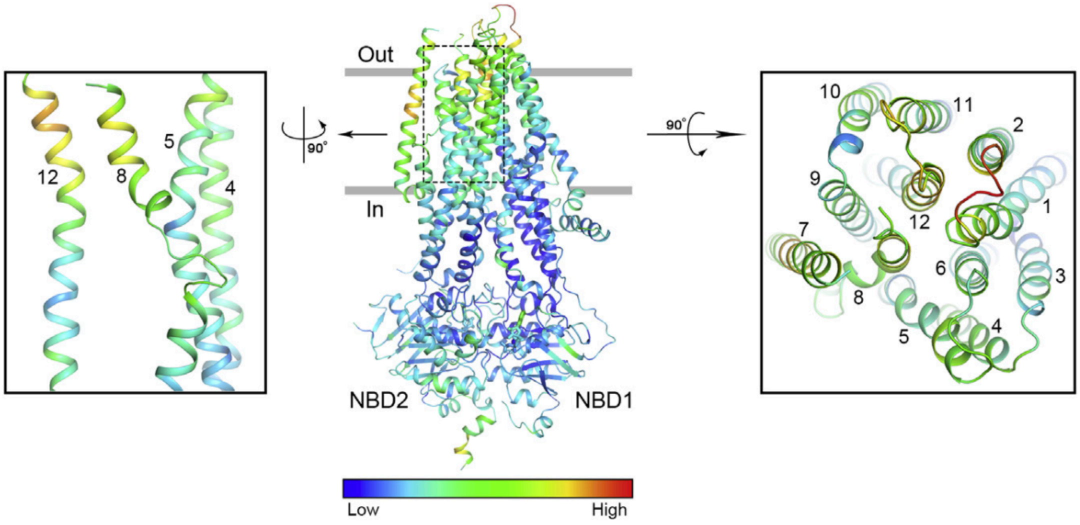


Fig. 3. 9 B-Factor distribution colored in rainbow
A zoom-in view of the extracellular ends of TM8 and TM12 is shown on the left; and a top view of the TMDs is shown on the right.

CHAPTER 4. Cryo-EM structure of phosphorylated, ATP-bound hCFTR

4.1 Summary

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.

This chapter (including this summary) is adopted from my work co-authored with Dr. Zhe Zhang and Prof. Jue Chen, published in *PNAS* 2018 [147].

4.2 Results

4.2.1 Structure determination

Based on previous work showing that eliminating ATP hydrolysis prolongs the lifetime of the open CFTR channel [21], 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-Mg^{2+} (9 mM). The final reconstruction, with an estimated overall resolution of 3.2 Å (Fig. 4.1 and Table 4.1), shows well-defined densities for the lasso motif, 12 TM helices, 2 NBDs, and 2 ATP-Mg^{2+} molecules (Fig. 4.2). Additional density, likely corresponding to part of the phosphorylated R domain, is also visible (Fig. 4.2). 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. 4.3B). The structure is refined to excellent geometry and stereochemistry statistics (Table 4.3). With this structure, we can now directly visualize conformational changes within hCFTR upon phosphorylation and ATP binding (Fig. 4.3). 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 phosphorylated R domain relocates to the peripheral surface, away from the NBD dimer interface.

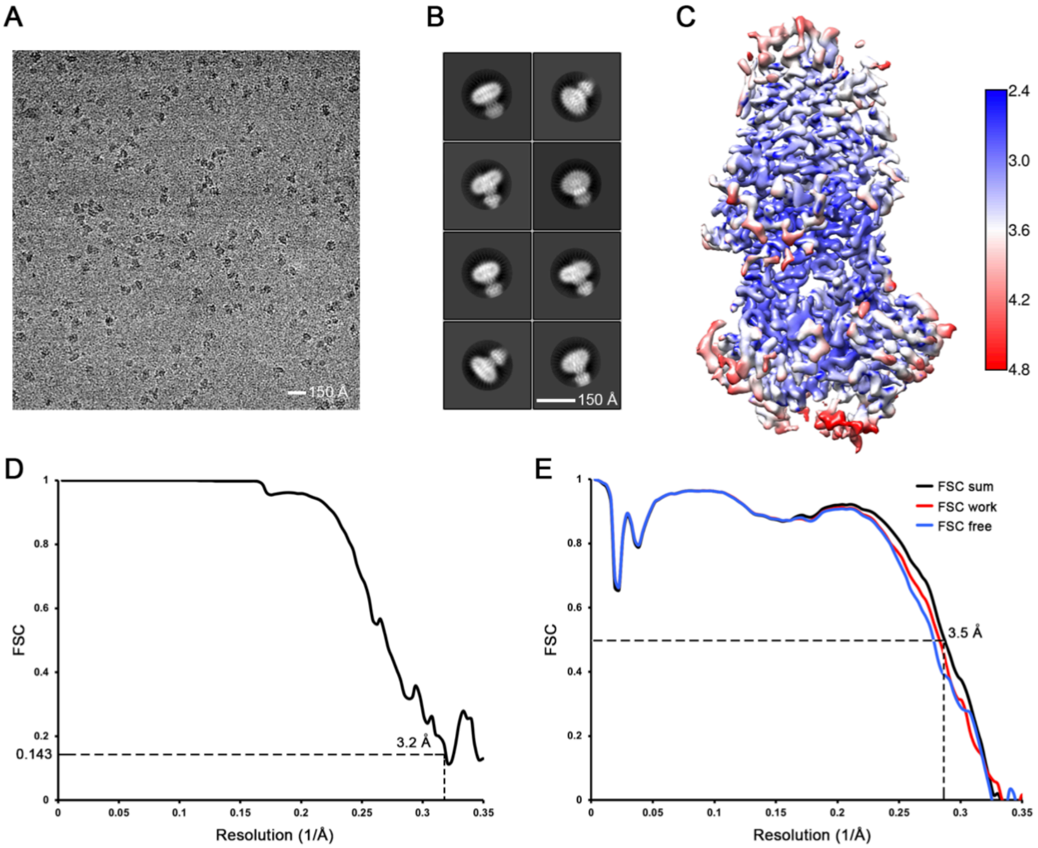


Fig. 4. 1 Cryo-EM analysis of phosphorylated, ATP-bound human CFTR
(A) A representative summed image at the 2.4 µm defocus level. (B) 2-D class average of the top 8 classes. (C) Local resolution distribution estimated by Blocres [57]. (D) Fourier Shell Correlation (FSC) curves between two half-datasets calculated by Frealign [59]. (E) FSC curves between the refined structure and the map calculated from the full dataset (sum, black), the half- map used in refinement (work, red), and the other half-map (free, blue).

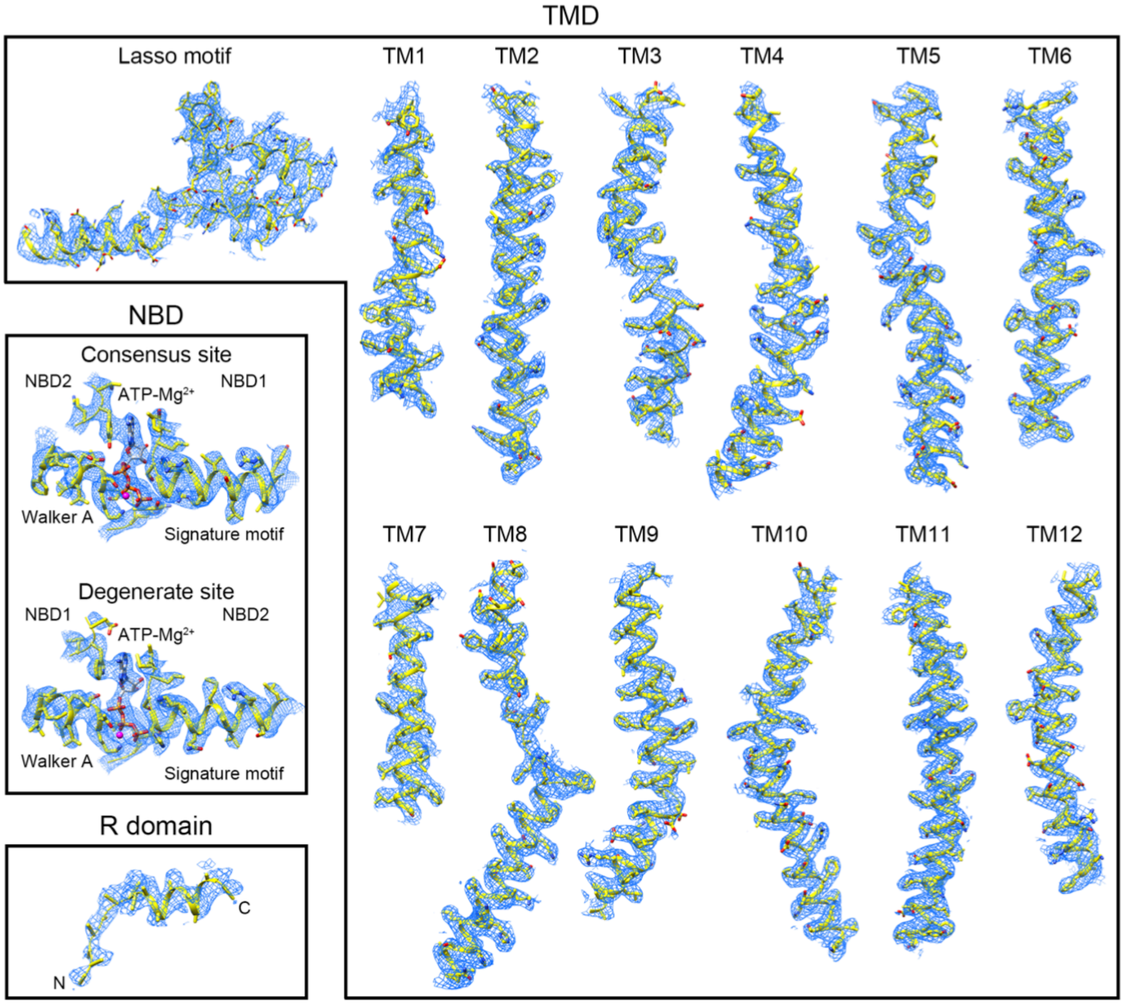


Fig. 4. 2 B-factor sharpened densities of different regions of phosphorylated, ATP-bound human CFTR

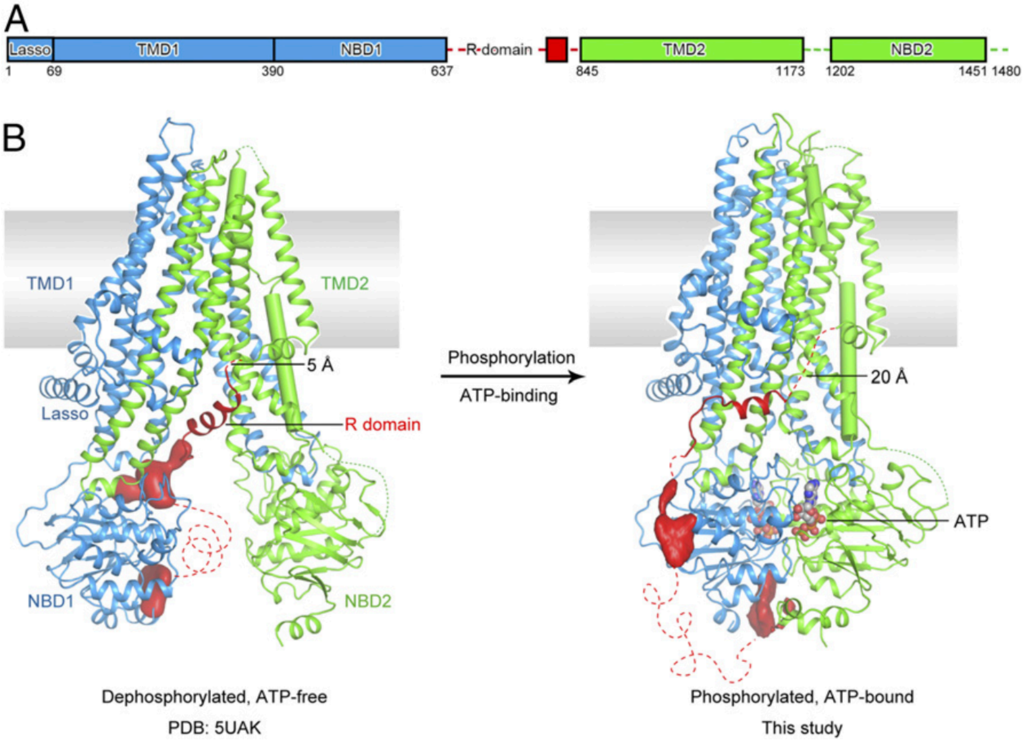


Fig. 4. 3 Two functional states of human CFTR
(A) Schematic diagram showing the domain structure of CFTR. The numbers represent the range of residues visible in the cryo-EM map, not the exact boundaries of different domains. (B) Ribbon diagrams of the dephosphorylated, ATP-free conformation (Left; PDB ID code 5UAK) and the phosphorylated, ATP-bound structure (Right). Regions not resolved in the structure are shown as dashed lines. The EM densities, shown in red, correspond to unstructured regions within the R domain. TM 8 is highlighted in cylinder. ATP is shown in ball-and-stick and colored by heteroatom. The distances between the last visible residue in the R domain and the first residue in TMD2 are indicated.

Table 4. 1 Summary of EM data and structure refinement statistics of phosphorylated, ATP-bound human CFTR

Data collection	
Microscope	Titan Krios (FEI)
Voltage (kV)	300
Detector	K2 Summit (Gatan)
Pixel size (Å)	1.03
Defocus range (μm)	0.8 to 2.5
Movies	2,093
Frames/movie	50
Dose rate (electrons/pixel/s)	8
Total dose (electrons/Å^2)	75
Number of particles	677,308
Model composition	
Non-hydrogen atoms	9,694
Protein residues	1,197
Lipids	6
ATP	2
Mg^{2+}	2
Refinement	
Resolution (Å)	3.2
Sharpening B-factor (Å^2)	-50
RMS deviations	
Bond lengths (Å)	0.013
Bond angles (°)	1.509
Validation	
Molprobity score	1.70
Clashscore, all atoms	7.46
Favored rotamers (%)	94.7
Ramachandran plot (%)	
Favored	95.9
Allowed	4.1
Outliers	0.0

4.2.2 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. 4.4A). Dimerization of the NBDs, which leads to channel opening [21], 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. 4.4B). 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. 4.4B). 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 [148]. The distance between the observed helix and the first visible residue in TMD2 (T845) is 20 Å (Fig. 4.3B), in line with all three possibilities. In the dephosphorylated hCFTR structure, a helical segment of the R domain was also observed (Fig. 4.3A). Because this helix is connected to TMD2, it is most likely to correspond to residues 825–843 (Fig. 4.3A). At a lower contour, amorphous densities are observed along the peripheral surface of NBD1 and at the C-terminal region of NBD1 (Figs. 4.3B and 4.4B). 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 [148-150]. Interaction of the R domain with the lasso motif was previously reported using isolated peptides [151], 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. Co-expression of the two halves of CFTR without the R domain results in an ATP-gated channel that opens independent of phosphorylation [71].

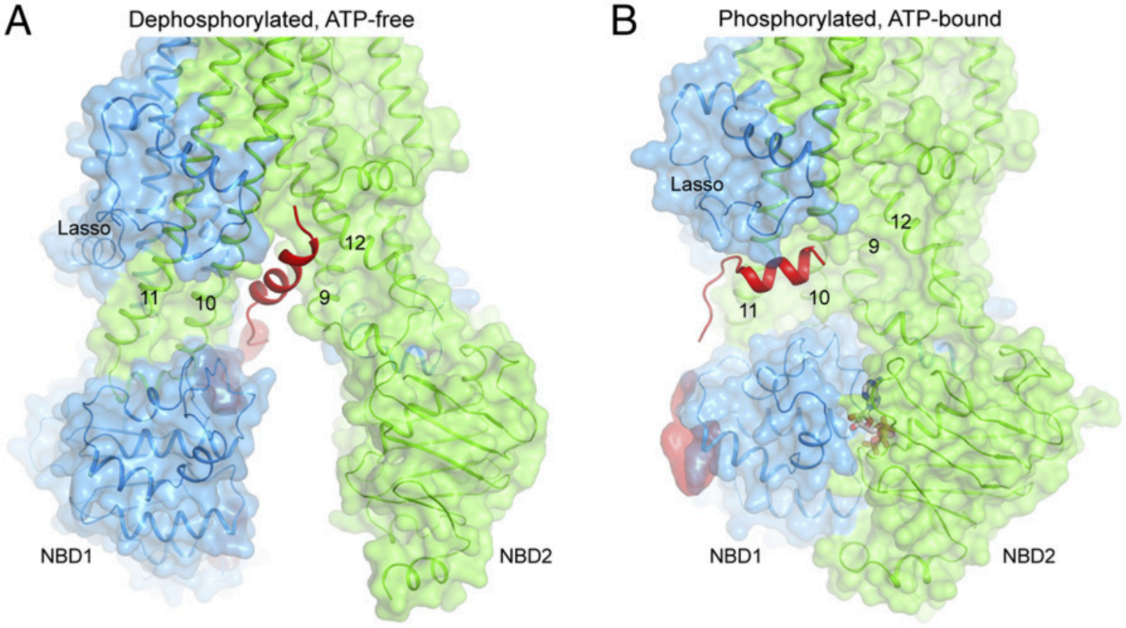


Fig. 4. 4 Conformational changes of the R domain
Close-up view of the intracellular opening in the dephosphorylated, ATP-free (A) and phosphorylated, ATP-bound (B) conformations. The R domain is shown in red. The lasso motif and TM helices 9–12 are indicated.

4.2.3 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. 4.5 A and B). Functional studies indicate that this passageway is the path of ion flow in CFTR [56, 152]. 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. 4.5 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. 4.5B). At the extracellular surface, a small opening is observed between TM 1 and TM 6 (Fig. 4.5 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 [35, 152]. Three positively charged residues, R104, R334, and K335, cluster on the extracellular side of the opening (Fig. 4.5 E and F). Modification of these residues influences anion conduction [25, 29, 67], 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.

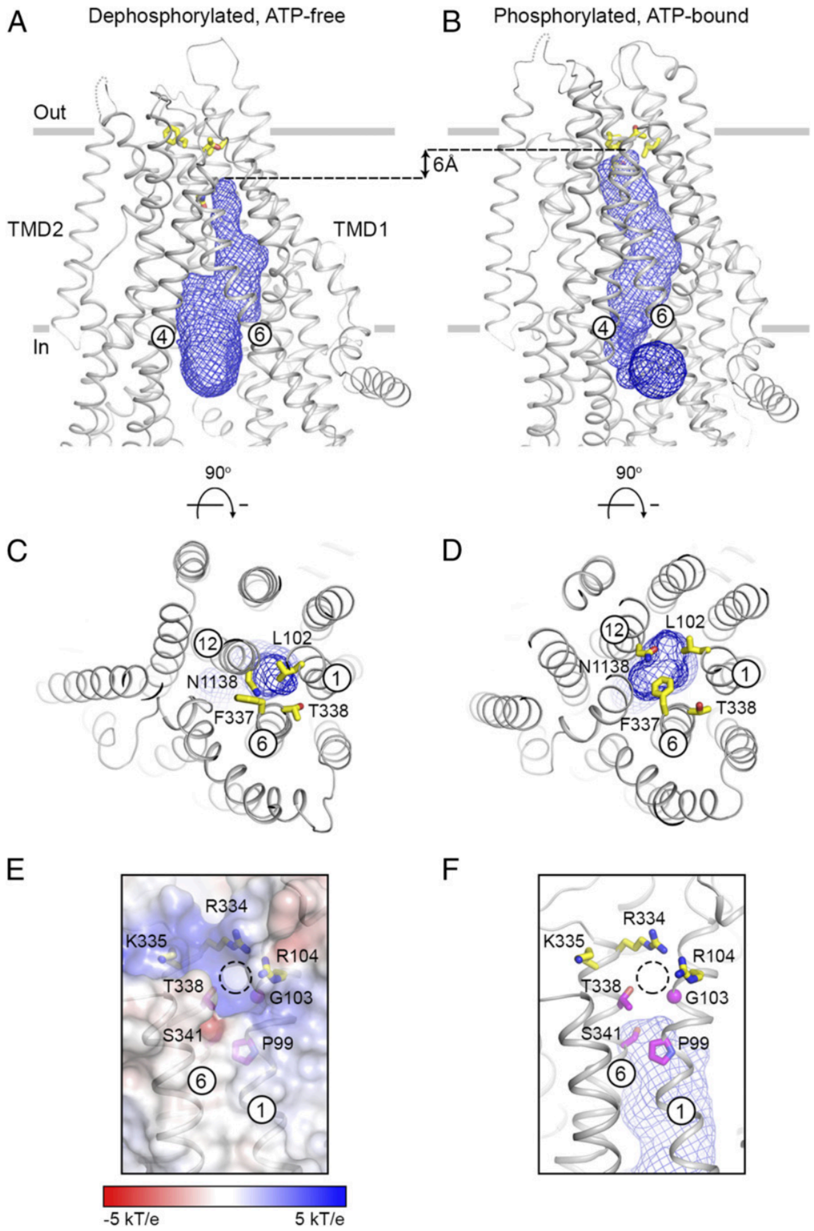


Fig. 4. 5 Conformational changes of the ion permeation pathway

Two orthogonal views of the dephosphorylated, ATP-free (A and C) and the phosphorylated, ATP-bound (B and D) structures. The pore, shown as a blue mesh, is defined by a probe with the size of a dehydrated chloride ion (1.7-Å radius). It is connected to the cytosol through an opening between TMs 4 and 6. The side chains of residues L102, F337, T338, and N1138 are shown as yellow sticks. (E and F) The extracellular opening (shown as a dashed circle). CFTR is shown in ribbon along with the electrostatic surface (E) or the pore as blue mesh (F). Residues located around the external region of the opening are shown in stick models. In purple are potential gating residues; in yellow are positively charged residues influencing ion conduction.

4.2.4 Structural comparison between zebrafish and human CFTR

The structure of the phosphorylated, ATP-bound hCFTR share many features with that of zCFTR [112] (Fig. 4.6). 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. 4.6B). 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 Cα 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. 4.6B, red region). The small extracellular opening observed in hCFTR (Fig. 4.3 E and F) is completely sealed in zCFTR. A second major structural difference is found in the NBD dimer (Fig. 4.6C). 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 [125, 126]. The NBD dimer in zCFTR is asymmetric: The degenerate site is more open than that of the consensus site (Fig. 4.6C). 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 [153]. Thus, the asymmetric NBD dimer observed 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 [154]. Whereas the Cα 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. 4.6D). 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 the pore [27, 155, 156]. However, in a different study, the R352Q mutant was proposed to alter the structure such that an endogenous cysteine becomes accessible to thiol reagents [67]. 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 [22]. In the structure of phosphorylated hCFTR, this salt bridge is indeed observed (Fig. 4.6E), 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.

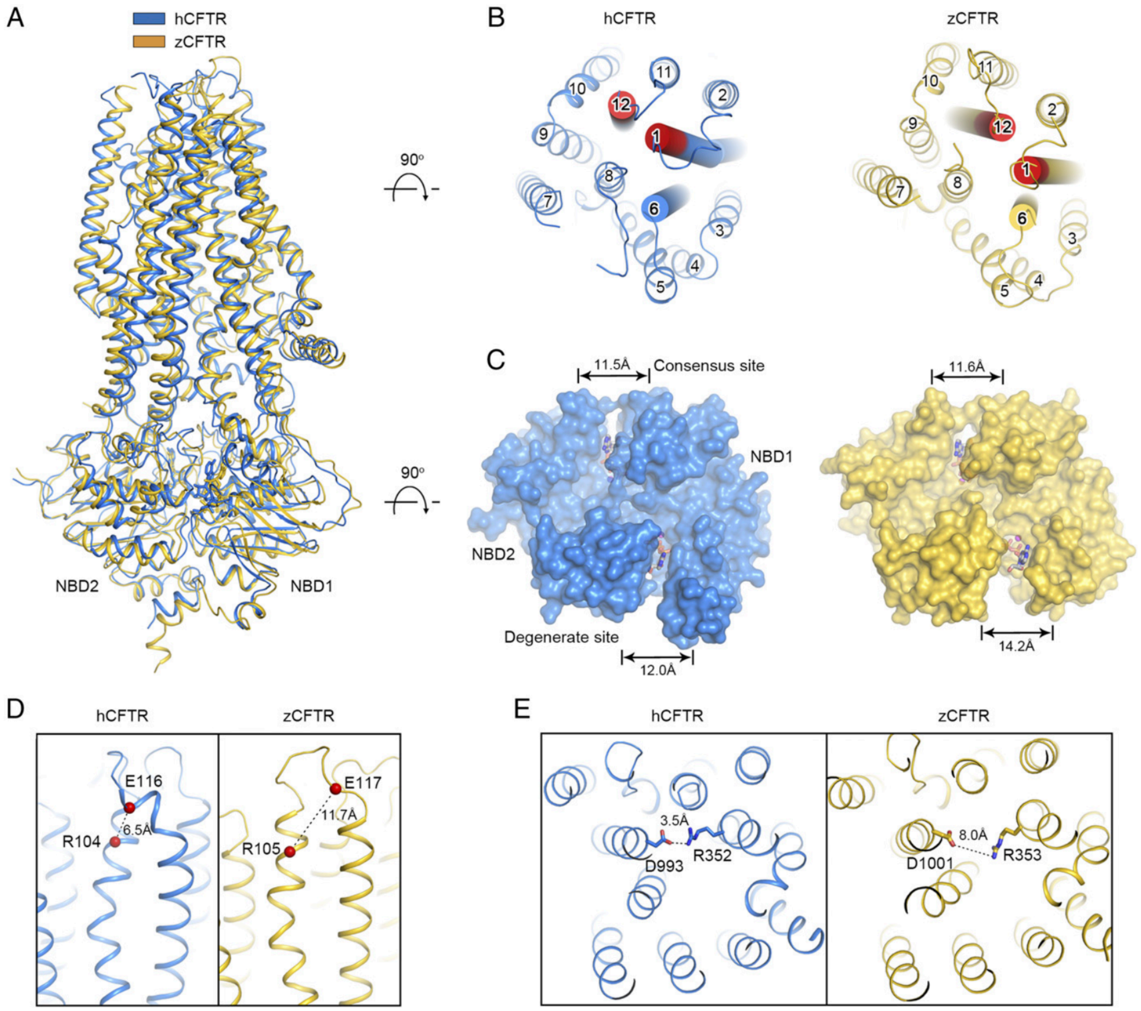


Fig. 4. 6 Structural comparison of phosphorylated human and zebrafish CFTR
(A) Superposition of the overall structures. hCFTR is shown in blue, and zCFTR (PDB ID code 5W81) is in yellow. (B) An extracellular view showing the different locations of TMs 1, 6, and 12 (highlighted as cylinders) in the two 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 Cα 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.

4.2.5 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 [39]. At that time, we could not explain several mutations, including G178R, G970R, and L927P [39]. 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. 4.7A). 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. 4.7A). 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 [157, 158]. 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 [159]. Similarly, L927P may also cause cystic fibrosis by interfering with conformational changes necessary for channel opening (Fig. 4.7B). The surface expression level of the L927P mutant is 43% that of the wild-type protein, but its channel activity is only 0.1% [40]. 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. 4.7B). 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.

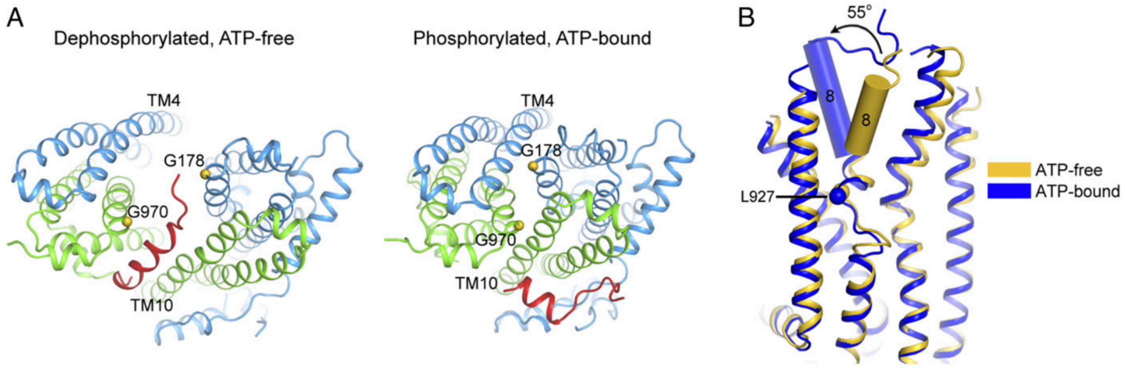


Fig. 4. 7 Disease-causing mutations that obstruct conformational changes
(A) The positions of G178 and G970. R domain is shown in red. (B) Superposition of the hCFTR in the ATP-free and ATP-bound conformations, showing the rotation of the outer leaflet segment of TM 8 and the location of L927.

4.3 Material and methods

4.3.1 Protein expression and purification

The E1371Q mutation was introduced to the wild-type construct with QuikChange Site-Directed Mutagenesis (Stratagene). The E1371Q mutant construct was expressed and purified using a protocol previously established for the wild-type CFTR [39]. Briefly, the construct was expressed in HEK293S GnTI$^-$ cells and purified using a nanobody resin against the GFP tag attached to the C-terminal of CFTR [92]. The membrane was solubilized with 1% 2,2-

didecylpropane-1,3-bis-β-D-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 $MgCl_2$, 1 mM ATP, and 0.06% digitonin.

4.3.2 EM sample preparation

The protein sample was concentrated to 6 mg/mL and supplemented with additional 8 mM ATP, 8 mM $MgCl_2$, and 3 mM fluorinated Fos-Choline-8 before applied to Quantifoil R1.2/1.3 400 mesh Au holey carbon grids (Quantifoil). Grids were frozen with Vitrobot Mark IV (FEI).

4.3.3 Data collection and processing

Data were collected on a 300-kV Titan Krios (FEI) microscope with a K2 summit detector (Gatan) in super resolution 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 CTFFIND4 [96] and Unblur [95], respectively. Subsequent data processing, including particle picking, 2D, and 3D classification, was carried out in Relion [58], except that drifts of individual particles were corrected using the program alignparts_lmbfgs [97]. The ATP-bound zebrafish CFTR [112] (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 [59] 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 (Fig. 4.1).

4.3.4 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 BFACTOR.EXE (written by Nikolaus Grigorieff) with a resolution limit of 3.2 Å and a b-factor of −50 $Å^2$. The model was built in Chimera [109] and Coot [99]. The final structure contains residues 1–390 of TMD1; 391–409, 435–637 of NBD1; 845–889, 900–1173 of TMD2; 1202–1451 of NBD2; and 17 residues of R domain (built as alanines). Phenix [101] and Refmac [102, 103] 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 Blocres [57]. The geometry of the model was validated using MolProbity [107, 108]. Figures were generated with Pymol (https://pymol. org/2/), Chimera [109], and HOLE [63].

CHAPTER 5. Structural identification of a hotspot on CFTR for potentiation

5.1 Summary

Cystic fibrosis is a fatal disease caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR). Two main categories of drugs are being developed: correctors that improve folding of CFTR and potentiators that recover the function of CFTR. In this chapter, we

report two cryo–electron microscopy structures of human CFTR in complex with potentiators: one with the U.S. Food and Drug Administration (FDA)--approved drug ivacaftor at 3.3-angstrom resolution and the other with an investigational drug, GLPG1837, at 3.2-angstrom resolution. These two drugs, although chemically dissimilar, bind to the same site within the transmembrane region. Mutagenesis suggests that in both cases, hydrogen bonds provided by the protein are important for drug recognition. The molecular details of how ivacaftor and GLPG1837 interact with CFTR may facilitate structure-based optimization of therapeutic compounds.

This chapter (including this summary) is adopted from my work co-authored with Dr. Zhe Zhang, Dr. Anat Levit, Jesper Levring, Dr. Kouki K. Touhara, Prof Brian K. Shoichet and Prof. Jue Chen, published in *Science* 2019 [160].

5.2 Results

5.2.1 Structure determination of CFTR in complex with ivacaftor

Ivacaftor was discovered by screening compounds that increase anion flux in G551D CFTR–expressing cells [41]. Subsequent studies have shown that ivacaftor increases the open probability (P_o) of both wild-type (WT) and mutant CFTRs in membrane patches, proteoliposomes, and planar lipid bilayers [161-163]. The potentiation by ivacaftor requires phosphorylation of CFTR by protein kinase A (PKA), but is independent of ATP [162]. These results suggest that ivacaftor acts directly on CFTR, rather than functioning through other regulatory mechanisms. To describe the specific molecular interactions between ivacaftor and CFTR, we determined a cryo-EM structure of ivacaftor in complex with phosphorylated E1371Q CFTR in the presence of saturating ATP-Mg^{2+} (10 mM) (Fig. 5.1A, Fig. 5.2 to 5.4, and table 5.1). The final map has an overall resolution of 3.3 Å, showing well-defined density throughout the protein, except for the R domain. Density for both ATP-Mg^{2+} molecules are visible at the NBD dimer interface (Fig. 5.7). An additional strong density is observed on the outer surface of the TMDs in the center of the lipid bilayer (Fig. 5.1C and Fig. 5.5C). This density, not observed in any of the previous CFTR structures [39, 56, 112, 147], has a shape and size consistent with the chemical structure of ivacaftor (Fig. 5.1C and Fig. 5.4). Within this density, we built a model of ivacaftor and examined multiple orientations of it in the site, using molecular docking [164, 165] followed by energy minimization of the protein-ligand complex. Complexes were prioritized by their energetic complementarity, ability to make favorable polar interactions, and subsequent refinement to the electron density maps. We previously reported the structure of the phosphorylated E1371Q construct in the presence of ATP-Mg^{2+} but in the absence of ivacaftor [147]. The ivacaftor-bound E1371Q exhibits the same protein conformation in which the ATP-bound NBDs form a closed dimer, and the two TMDs pack closely together to form an ion conduction pathway open to the cytoplasmic solution. The R domain, largely unstructured, is located along the peripheral surface of NBD1 and the cytoplasmic region of the TMDs (Fig. 5.1A). No significant protein conformational changes were observed upon binding of ivacaftor, and the overall root mean square deviation between the two structures is 0.14 Å (Fig. 5.1B).

5.2.2 Identification of a binding pocket of ivacaftor

Ivacaftor binds CFTR at the protein-lipid interface, docking into a cleft formed by transmembrane (TM) helices 4, 5, and 8 (Fig. 5.1, C and D). The binding site coincides with a hinge region in TM8, a structural feature of CFTR not found in other ABC transporters [112, 147]. The extracellular segment of TM 8 rotates around this hinge upon ATP binding [39, 56, 112]; stabilizing this rotation may explain the drug's efficacy. Whereas about 40% of the molecular surface of ivacaftor is buried against CFTR, the remaining 60% is exposed to the hydrophobic region of the membrane (Fig. 5.1, C and D).

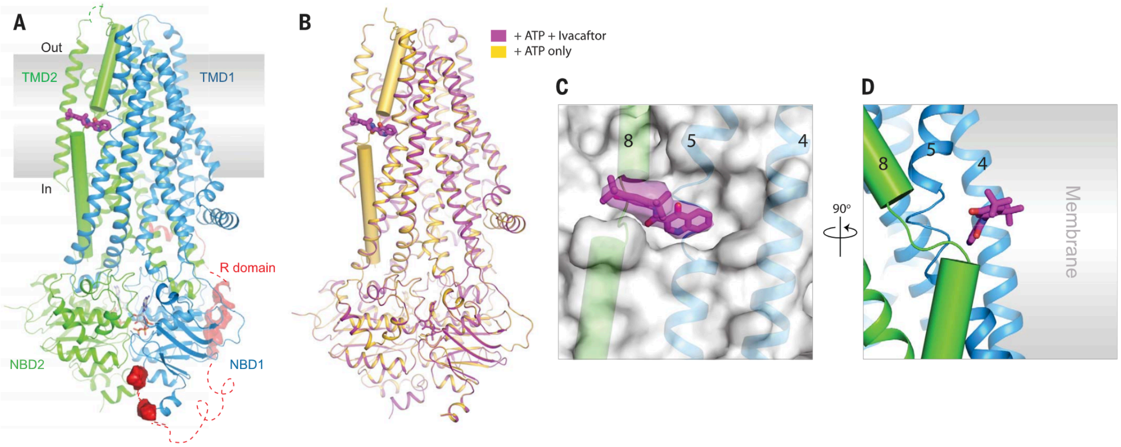


Fig. 5. 1 Ivacaftor binds CFTR inside the membrane
(A) Overall structure of the phosphorylated, ATP-bound human CFTR in complex with ivacaftor (shown in magenta). TMD1 and NBD1 are shown in blue, TMD2 and NBD2 in green, and R domain in red. TM8 is highlighted in cylinder representation. Regions not resolved in the structure are shown as dashed lines. (B) Superposition of phosphorylated, ATP-bound CFTR in the absence (yellow) and presence of ivacaftor (magenta). (C) A magnified view of the ivacaftor-binding site. CFTR is shown as a transparent surface model with TM 4, 5, and 8 indicated. Ivacaftor is shown as a stick model together with the corresponding EM density. (D) Ivacaftor binds at the protein-lipid interface, exposing half of its surface to the lipid bilayer.

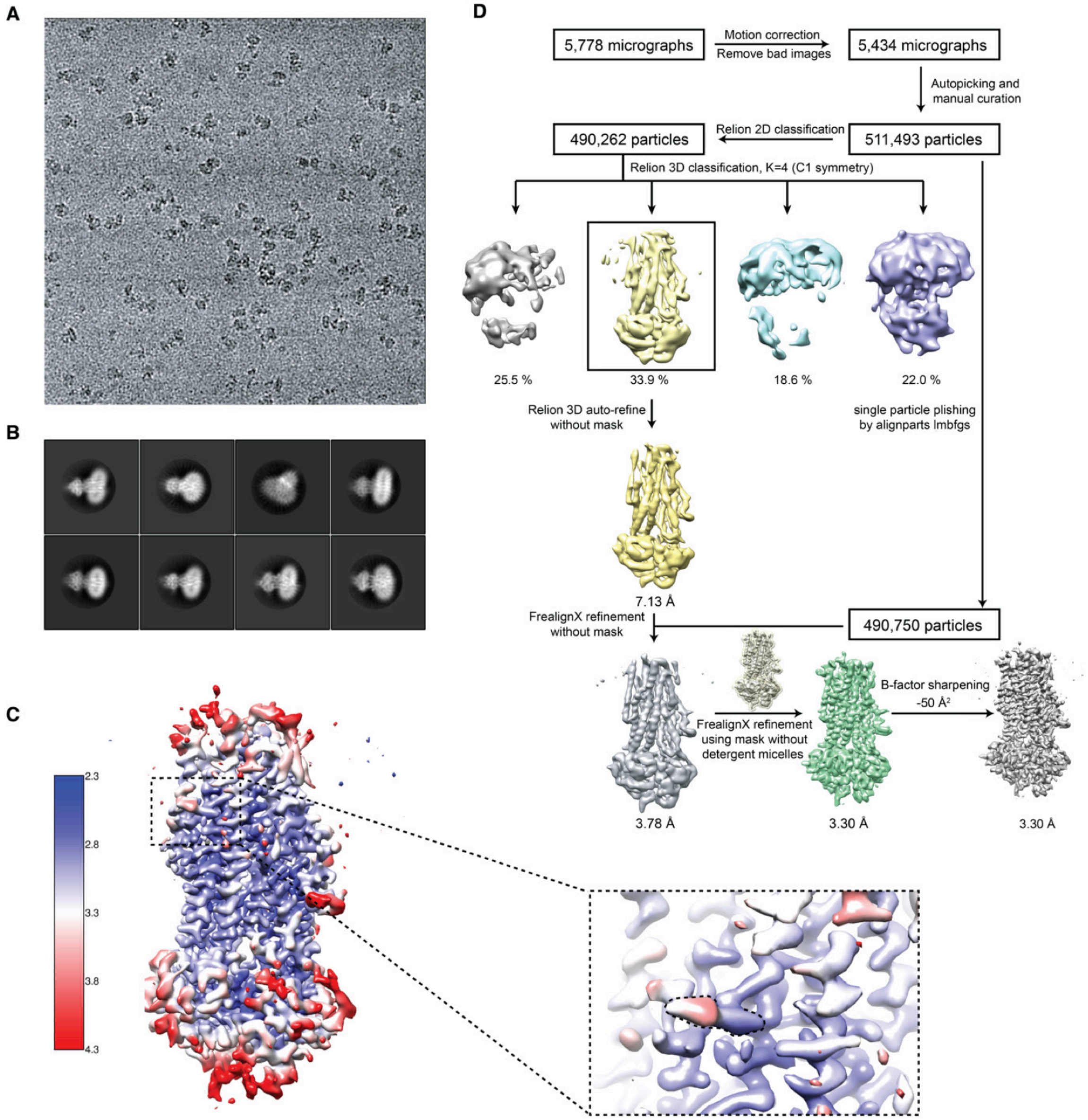


Fig. 5. 2 Cryo-EM analysis of the CFTR/ivacaftor complex
(A) A representative micrograph at defocus of -2 ***μ***m. (B) Representative 2D class averages from Relion [58]. (C) Overall EM map colored by local resolution estimation using Blocres [57]. Inset: zoom-in view of the drug binding site; density corresponding to ivacaftor is outlined with black dash line. (D) Summary of the image processing workflow.

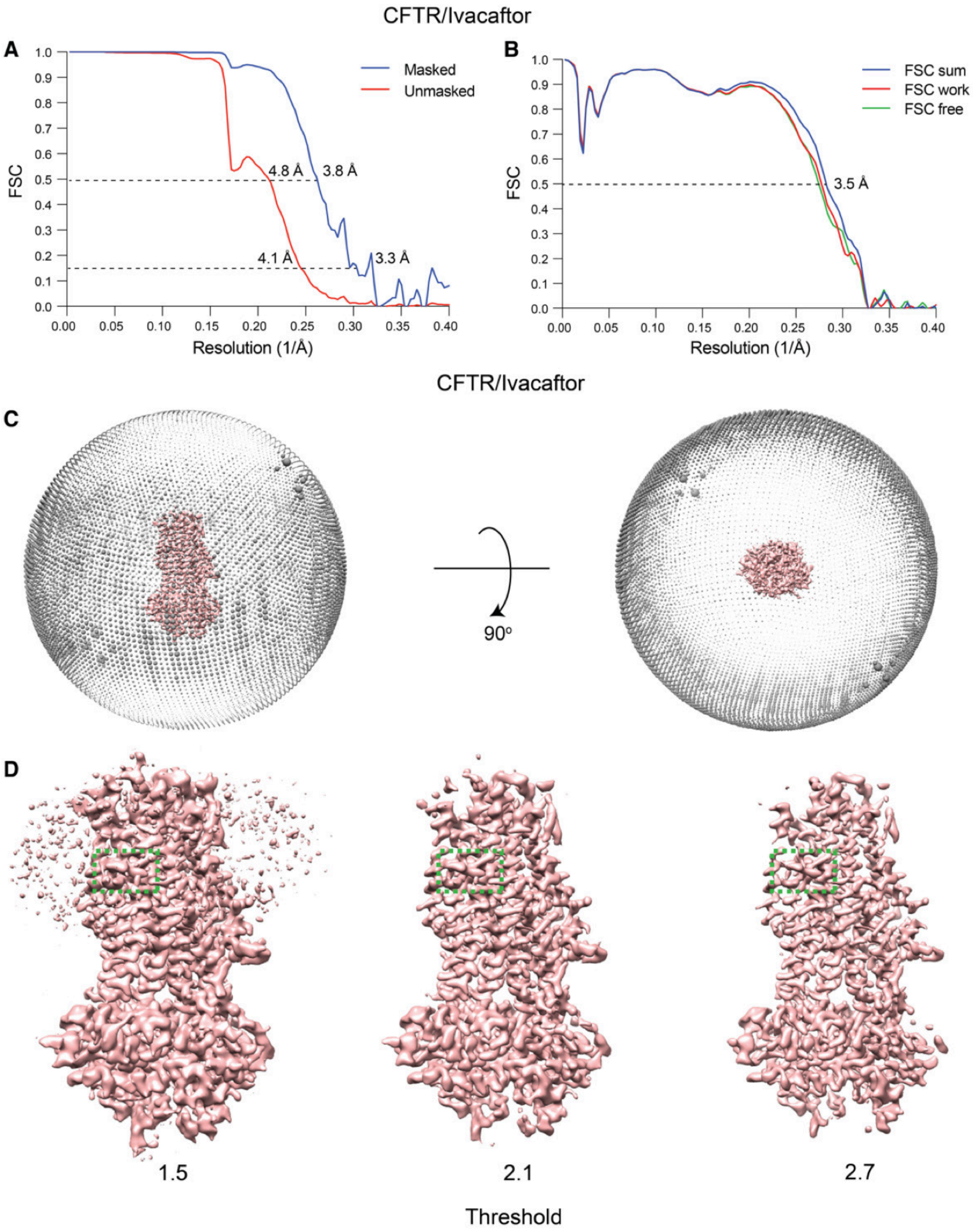


Fig. 5. 3 Data and model quality assessment for CFTR/ivacaftor complex
(A) Fourier Shell Correlation (FSC) (blue: masked; red: unmasked) as a function of resolution, using FrealignX [59] output. (B) FSC calculated between the full, half map for refinement (FSC work), half map for validation (FSC free) and the refined structure. (C) Angular distribution of the particles used in the final reconstruction. Each sphere indicates the number of particles from this angle and the size of the spheres corresponds to the number of particles. (D) CFTR/Ivacaftor map plotted with three different thresholds. At threshold 1.5, ivacaftor density is much stronger the micelle density; at 2.1, ivacaftor density is as strong as the protein density; at 2.7, ivacaftor density is still prominent as some main chain density starts to disappear.

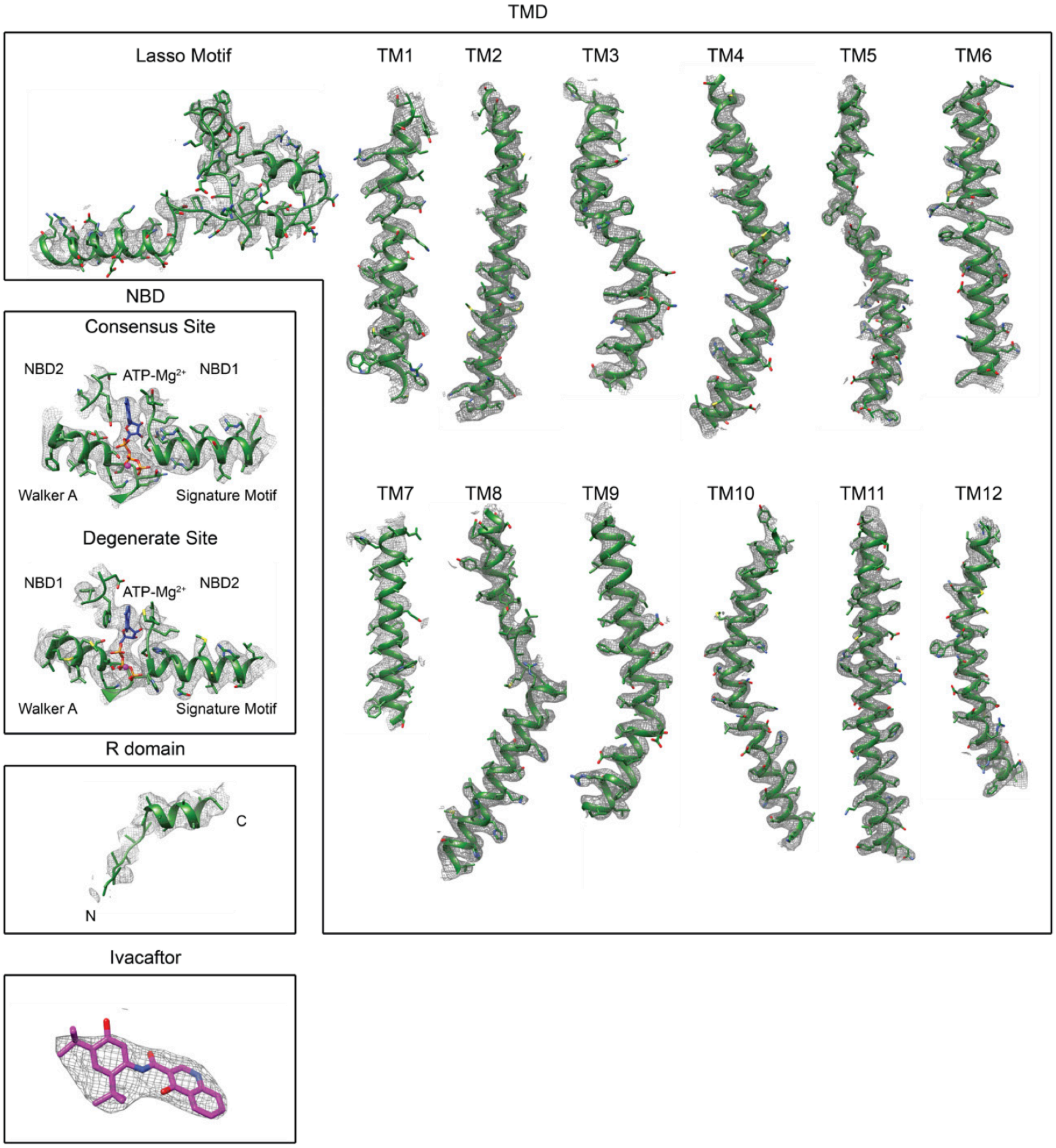


Fig. 5. 4 EM density of different parts of the CFTR/ivacaftor complex

Table 5. 1 Summary of EM data and structure refinement statistics for CFTR/ivacaftor complex

Data collection and processing	
Microscope	Titan Krios (FEI)
Voltage (kV)	300
Detector	K2 Summit (Gatan)
Pixel size (Å)	0.817
Defocus range (μm)	0.8 to 2.5
Frames/movie	50
Exposure rate (electrons/pixel/s)	8
Total dose (electrons/Å^2)	84
Data acquisition software	SerialEM
Micrographs collected/processed (no.)	5,778/5,434
Particles extracted (no.)	511,493
Particles for 3D/final reconstruction (no.)	490,262/490,750
Symmetry imposed	C1
Model composition	
Non-hydrogen atoms (no.)	9,767
Protein residues (no.)	1,196
Ivacaftor/Lipids/ ATP/Mg^{2+} (no.)	1/7/2/2
Refinement	
Resolution at 0.143 FSC threshold (Å)	
masked/unmasked	4.1/3.3
Resolution at 0.5 FSC threshold (Å)	
masked/unmasked	4.8/3.8
Local resolution range (Å)	2.3 - 4.3
Sharpening B-factor (Å^2)	-50
Model refinement package	Coot and REFMAC (reciprocal space refinement)
RMS deviations	
Bond lengths (Å)	0.0029
Bond angles (°)	1.2537
B-factors of the model	
Protein residues	207.9
ATPs	184.9
Mg^{2+}	156.2
Lipids and detergent	226.9
Ivacaftor	189.8
Validation	
Molprobity score	1.25
Clashscore, all atoms	1.41
Favored rotamers (%)	93.1
CaBLAM outliers (%)	1.11
Average FSC	0.69
Average CC	0.52
EMRinger Score	2.72
Ramachandran plot (%)	
Favored/Allowed/Outliers	95.4/4.6/0.0

5.2.3 Verfication of the ivacaftor binding site

The interactions between ivacaftor and CFTR include two hydrogen bonds, two aromatic interactions, and six hydrophobic interactions (Fig. 5.5, A and B). To evaluate how each residue contributes to ivacaftor binding, we used scintillation proximity assay (SPA) to measure the apparent affinity of ivacaftor for CFTR (Fig. 5.5, C to E). Every residue in the binding site was individually substituted by alanine. Every mutant eluted from the size-exclusion column as a monomeric peak, similar to the WT CFTR, indicating that these mutations did not alter CFTR folding (Fig. 5.6). Specific binding of ivacaftor to the WT CFTR increased as a function of ivacaftor concentration (Fig. 5.5C). Nonlinear regression analysis shows that the data fit well to a single site binding model with an equilibrium dissociation constant (K_d) of 6.6 ± 1.2 nM. In comparison, all but one mutation resulted in a reduced binding affinity for ivacaftor (Fig. 5.5, C to E, and table 5.2). Four residues appear to be most important—their alanine substitutions nearly abolishing ivacaftor binding (Fig. 5.5, C and D). Among them, S308 and F312 directly coordinate the oxoquinoline moiety through a hydrogen bond and a π- π stacking interaction, respectively (Fig. 5.5, A and B). The other two residues, R933 and Y304A, hydrogen bond to main-chain carbonyls in the TM 8 hinge, thus stabilizing the overall structure of the binding site (Fig. 5.5, A and B). Another important residue, F931, forms an edge to-face interaction with the phenol ring of ivacaftor. Mutation of F931 to alanine decreased drug affinity by about 10-fold (Fig. 5.5D and table 5.2). Mutating F305, L233, and F236, the three residues within van der Waals distance from the oxoquinoline of the drug, also decreased its affinity (Fig. 5.5E and table 5.2). By contrast, substitution of F932, which interacts with one of the lipidexposed tert-butyl groups, had no effect (Fig. 5.5E and table 5.2). These mutational effects are consistent with the structure of the CFTR–ivacaftor complex, indicating that hydrogen bonds are critical for ligand recognition in the low-dielectric environment of the membrane. The structure of CFTR–ivacaftor complex largely explains the structure activity relationship (SAR) for the series of analogs that led to this drug [41]. From the published SAR, 48 ivacaftor analogs, ranging from the initial high-throughput screening hit to optimized leads, may be readily docked into the ivacaftor site, making favorable interactions (Fig. 5.7). The docked poses superpose with the ivafactor structure, recapitulating more and more of the drug's interactions with CFTR as the molecules are optimized. Key interactions common to most of the docked complexes include the internal hydrogen bond between the conserved side-chain amide and the ubiquitous oxoquinoline oxygen, the hydrogen bond between the main-chain nitrogen of F931 and that same amide, and the interaction between the oxoquinoline nitrogen and S308. Similarly, the stacking observed between F312 and ivacaftor's oxoquinoline ring is conserved among the analogs. The phenolic hydroxyl, which appears late in the affinity maturation and is retained in ivacaftor, docks to interact with R933, as observed in the ivacaftor complex; addition of this group substantially increases affinity, at least partly reflecting its new interaction with the arginine (Fig. 5.7).

(Detailed analysis of the docking pose: to further validate the observed binding mode, our collaborator Dr. Levit and Prof. Shoichet docked a series of 48 compounds from ivacaftor's structure-activity relationship (SAR) series [41] and analyzed whether their poses are consistent with ivacaftor's pose as observed through cryo-EM and molecular docking.

The original high-throughput screen (HTS) hit, compound **1** (EC50 = 2 μM), docks to the same binding site as the one identified for ivacaftor (Fig. 5.7, A and M). Compound **1**'s quinoline ring adopts the same orientation as observed for ivacaftor, however, the quinoline nitrogen forms a hydrogen bond with the backbone of F931, instead of S308. The quinoline rings also stack with

F312, as observed for ivacaftor. Compound **1** adapts this configuration because its geometry does not allow it to recapitulate the exact position of ivacaftor in the binding pocket. If it were to do so, there will be a clash between the benzene ring and the F931 loop area. Hence, compound **1** rearranges and docks to hydrogen bond with the backbone of F931.

The SAR data suggested that the quinoline ring has two chemical features that are important for retaining activity [41]. First, an inter-molecular hydrogen bond involving the quinoline nitrogen. In the HTS hit, this hydrogen bond is formed to the backbone of F931, while in the more potent analogs (below), the quinoline nitrogen hydrogen bonds with S308. A second feature is the formation of an intra-molecular hydrogen between the quinoline carbonyl and the side chain amide nitrogen, which stabilizes a planar conformation. This is particularly apparent in the more potent analogs.

The potent compounds from the SAR series that led to ivacaftor share these intra- and intermolecular interactions. For instance, the analog **7** (EC50 = 1.3 μM) hydrogen bonds with S308 through the quinoline nitrogen, as also observed for ivacaftor (Fig. 5.7B). Like the drug, it stacks with F312. The carbonyl group of the linker side chain hydrogen bonds with the backbone of F931, as also observed for ivacaftor. This is also true for compound **8** (EC50 = 0.2 μM), where the introduction of a linear aliphatic group on the ortho position increases potency (Fig. 5.7C), likely reflecting better packing. Compound **16** (EC50 = 0.1 μM; Fig. 5.7D) hydrogen bonds with the backbone of F931 through the bridging carbonyl group, in addition to the previously installed hydrogen bond between the oxoquinoline nitrogen and S308. Both of these interactions are also observed for ivacaftor. The quinoline rings stack with F312 while the indole group stacks with F932. The SAR data suggested that the indole of **16** formed an important interaction which led to major improvement in activity, which is supported by this pose, that is also capable of forming an intra-molecular hydrogen bond. These interactions are also observed in the docked poses of compounds **27**, **28**, **31**, **35**, **42** and **44** (Fig. 5.7, E-J). Finally, the pose of compound **45** (EC50 = 3 nM) recapitulates that of ivacaftor (Fig. 5.7K), hydrogen bonding with the backbone of F931 through the bridging carbonyl group while the oxoquinoline nitrogen continues to hydrogen bond with S308. The hydroxyl group of this compound, which only appears late in this series, points towards the guanidino group on R933, forming an ion-dipole interaction.

Overall, the ready accommodation of this SAR series into the ivacaftor binding site, and the appearance of new favorable interactions as affinity rises in this series, supports the pose of ivacaftor modeled in the cryo-EM structure.

Several caveats bear airing. Whereas the compounds in the SAR series overall can be well-accommodated in the ivacaftor pose, there were several cases where weak inhibitors also fit well. For example, compound **37**, which only has 24% inhibition at 30 μM (Fig. 5.7L), docks well in the site, forming most of the hydrogen bonds and non-polar interactions observed for the active compounds. The one thing that appears out of place is an additional amide that is unsatisfied and will pay a substantial desolvation penalty; this may partly explain its loss of activity. Additional examples include compounds **21** (12% inhibition at 30 μM) and **22** (21% inhibition at 30 μM), which also dock similarly to ivacaftor, but also bury at least one polar—and so hard to desolvate-- group without a complementary group with which to interact on the receptor. Finally, we note that we do not expect a linear or even monotonic relationship between docking scores and binding affinities. At most what can be expected is a qualitative agreement between SAR and structural features; this certainly is seen here.)

These results are consistent with the ivafactor structure determined here and the SAR observed in its development [41].

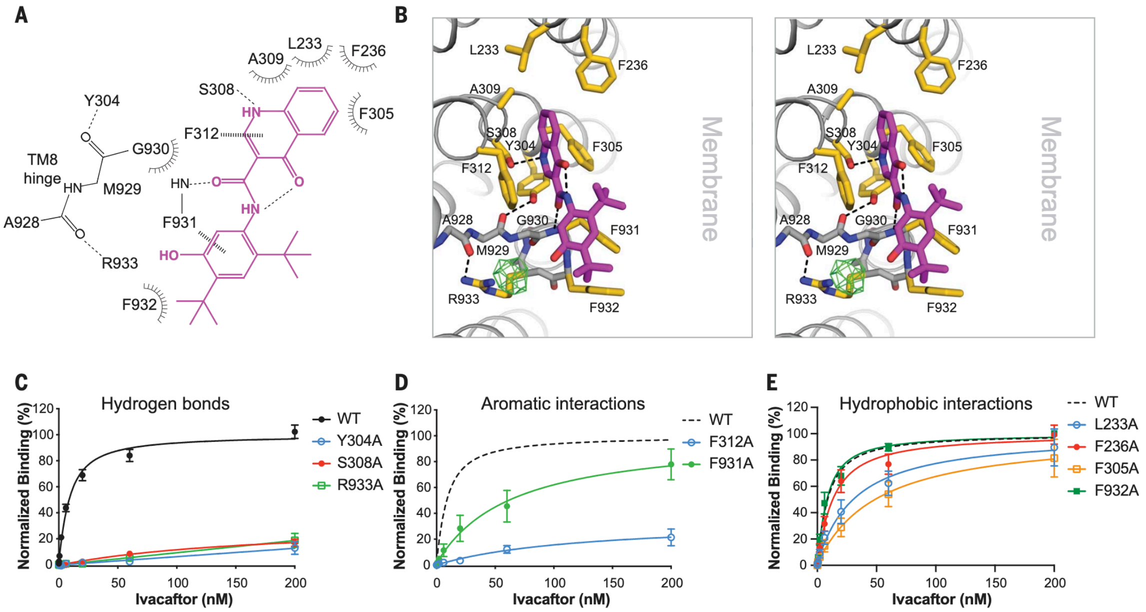


Fig. 5. 5 Contribution of individual residues to ivacaftor binding

(A) Schematic drawing of interactions formed between ivacaftor (magenta) and the CFTR-binding site. Residues within van der Waals distances (<4.5 Å) are shown. Representations: black dashed lines, hydrogen bonds; blue vertical lines, aromatic interactions; spokes, hydrophobic interactions. The hinge region in TM8 is shown as black sticks and labeled. (B) Stereo view of the ivacaftor-binding site. Residues within van der Waals distances are shown in yellow, and hydrogen bonds are depicted as black dashed lines. An unknown density between R933 and ivacaftor is shown as green mesh. (C to E) Binding affinities of WT CFTR and mutants replacing residues making (C) hydrogen bonds, (D) aromatic interactions, and (E) hydrophobic interactions with ivacaftor. Data points represent the means and SEMs of at least three measurements. The calculated K_d values are listed in Table 5.2. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; D, Asp; E, Glu; F, Phe; G, Gly; L, Leu; M, Met; Q, Gln R, Arg; S, Ser; and Y, Tyr.

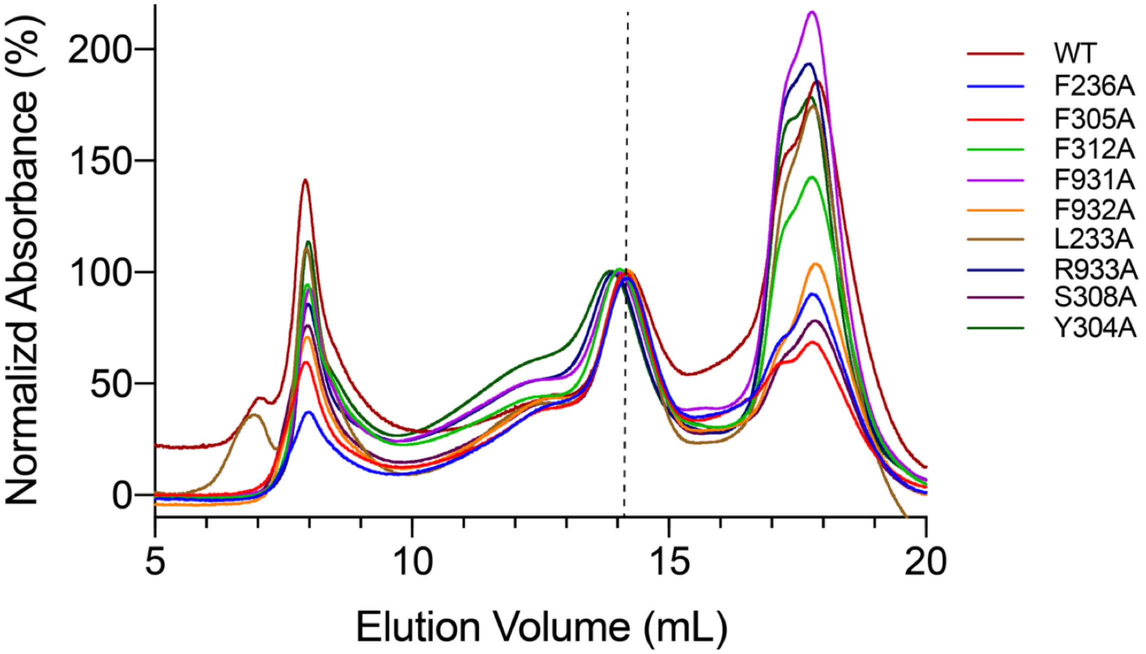


Fig. 5. 6 Fluorescence size exclusion profiles of *wt* and mutant CFTR
The dashed line indicates the position of the *wt* CFTR.

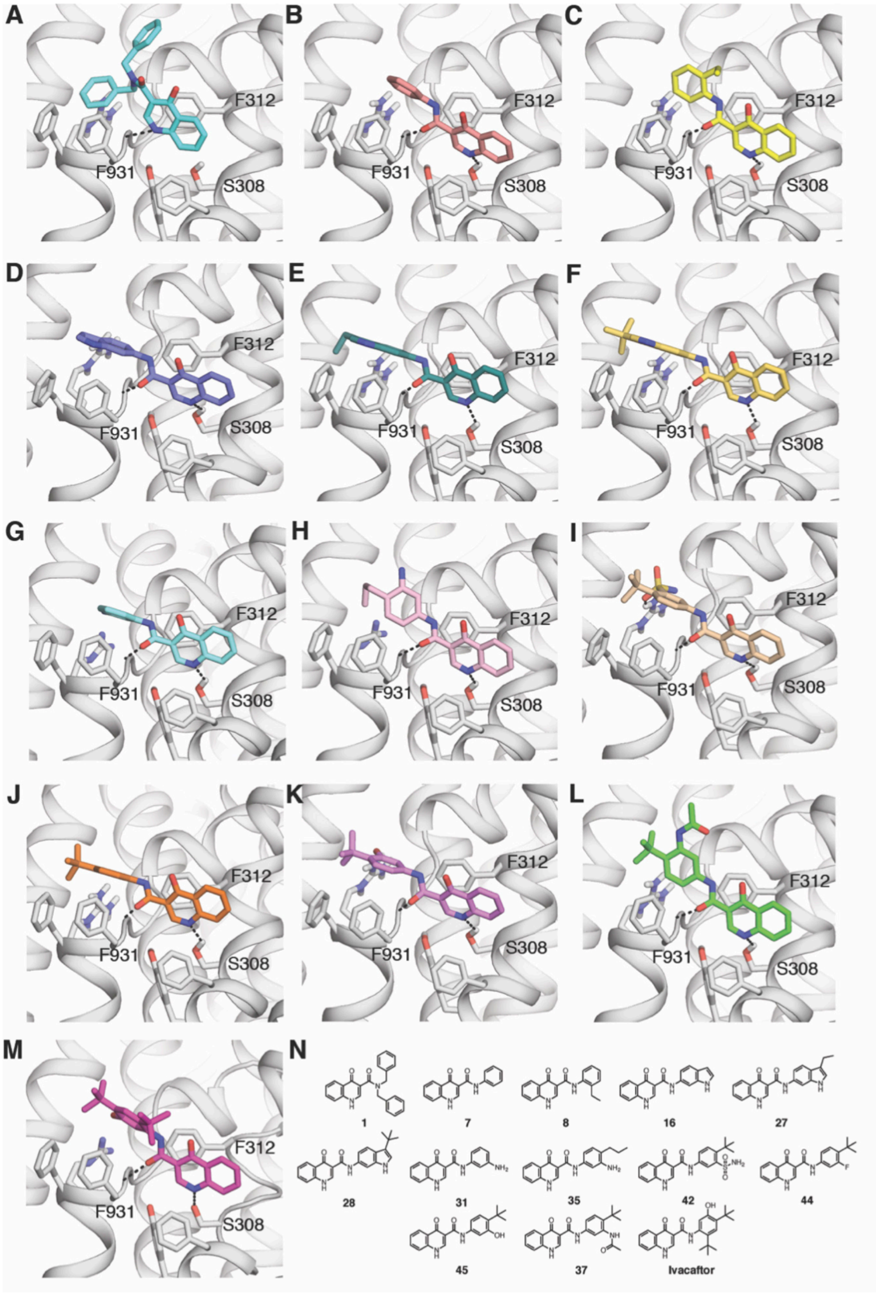


Fig. 5. 7 SAR study of Ivacaftor and its analog series from HTS hit to the final lead
Docked poses of representative active and potent analogs out of 48 analogs which were docked are shown here, as well as the experimentally determined conformation of Ivacaftor in the CFTR binding site. Compound numbers are as described in [41]. Residues within van der Waals distances are shown as sticks. Hydrogen bonds are shown as black dashed lines. (A) compound 1 (HTS hit). (B) compound 7. (C) compound 8. (D) compound 16. (E) compound 27. (F) compound 28. (G) compound 31. (H) compound 35. (I) compound 42. (J) compound 44, (K) compound 45. (L) compound 37. (M) ivacaftor. (N) 2D chemical structures of the compounds shown in panels A-M.

Table 5. 2 Dissociation constants of ivacaftor to CFTR

Mutation	K_d ± S.E. (nM)	Type of Interaction
WT	6.6 ± 1.2	N/A
Y304A	ND	H-bond with TM8 backbone
S308A	ND	H-bond
R933A	ND	H-bond with TM8 backbone
F312A	ND	Face to face π- π stacking
F931A	61 ± 29	Edge to face aromatic
F305A	48 ± 20	Hydrophobic
L233A	29 ± 11	Hydrophobic
F236A	11 ± 3	Hydrophobic
F932A	5.8 ± 1.0	Hydrophobic

5.2.4 A different potentiator GLPG1837 binds at the same site

Recently, a new potentiator, GLPG1837, has been discovered to have higher efficacy than that of ivacaftor [42]. We first studied the effects of GLPG1837 in a planar bilayer system, where detergent purified CFTR channels were reconstituted into liposomes then fused with a bilayer lipid membrane. At saturating ATP concentration, the Po of the phosphorylated WT CFTR increased from 0.23 to 0.54 upon addition of 10 mM GLPG1837 (Fig. 5.8, A and B). The Po of E1371Q was also increased by GLPG1837, from 0.64 to 0.88 (Fig. 5.8, A and B). The Po values are lower than those measured in cellular membranes [166, 167], possibly owing to differences in their lipid compositions. A competitive binding assay shows that GLGP1837 reduced the apparent affinity of ivacaftor (Fig. 5.8C); the inhibition constant (K_i) was determined to be 0.30 ± 0.08 μM (Fig. 5.8D). Next, we determined the cryo-EM structure of phosphorylated E1371Q CFTR in complex with ATP-Mg^{2+} and GLPG1837 (Fig. 5.8E, Figs. 5.9 to 5.11, and table 5.3). The overall structure, at 3.2 Å resolution, is essentially indistinguishable from those of drug-free and ivacaftor-bound forms. Although GLPG1837 clearly binds in the same pocket, its orientation and shape differ from that of the ivacaftor density (Fig. 5.8F). Here too, a model of the CFTR–GLPG1837 complex was built by molecular docking, followed by energy minimization and refinement of the ligand receptor complex. In the final model, the drug fits well into the electron density, with favorable polar and nonpolar interactions, and a relatively unstrained ligand geometry (Fig. 5.8F). Although the chemical structure of GLPG1837 differs from that of ivacaftor, residues interacting with GLPG1837 largely overlap with those engaging ivacaftor (Fig. 5.12, A and B). Specifically, S308 and Y304 form hydrogen bonds; and L233, F236, F305, A309, F312 form hydrophobic interactions with the drug (Fig. 5.12, A and B). Four polar groups on GLPG1837 are engaged in intramolecular interactions (Fig. 5.12, A and B), also observed in the crystal structure of the compound itself [42], that shield charges, permitting this relatively polar molecule to permeate the membrane.

5.2.5 Verificaiton of the GLPG1837 binding site

To evaluate the contribution of the intermolecular hydrogen bonds, we measured the half maximal effective concentration (EC50) values of GLPG1837 for the WT, Y304A, and S308A CFTR (Fig. 5.12, C to E). The EC50 value of the WT CFTR was determined to be 0.12 ± 0.03 μM (Fig. 5.12C), similar to the reported value of 0.23 ± 0.12 μM [167]. Replacing Y304 or S308 with an alanine increased the EC50 values by 57- and 34-fold, respectively (Fig. 5.12, D and E), underscoring the importance of the structurally observed hydrogen bonds in GLPG1837 recognition.

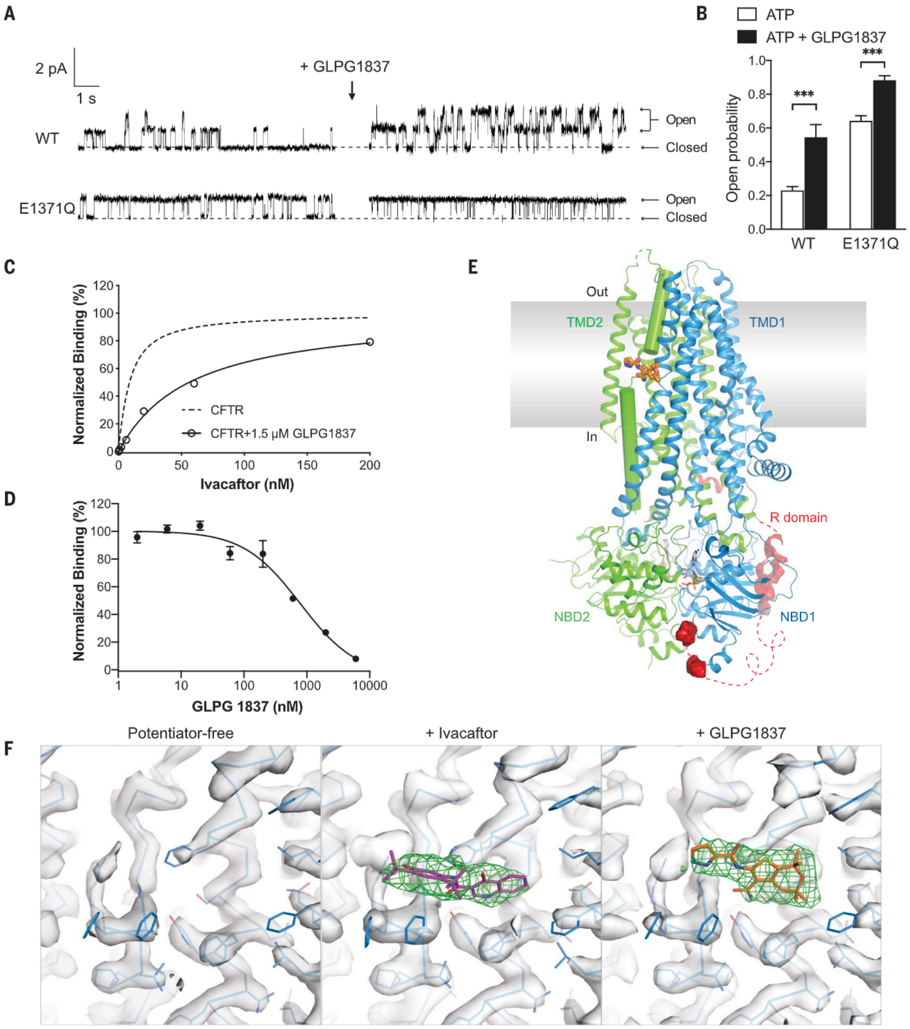


Fig. 5. 8 GLPG1837 binds to the same site as ivacaftor

(A) Representative recordings of WT (upper trace) and E1371Q (lower trace) CFTR reconstituted in synthetic lipid bilayers. CFTR was phosphorylated with PKA prior to fusion with bilayers. Recordings were performed on individual membranes with 2 mM ATP before (left) and after (right) addition of 10 **μ**M GLPG1837. (B) Open probabilities of WT and E1371Q CFTR before (open bar) and after (filled bar) addition of 10 **μ**M GLPG1837. Data points represent the means and SEMs of three to nine membranes. $P_o = 0.23 \pm 0.02$, for WT; $P_o = 0.54 \pm 0.08$, for WT + GLPG1837; $P_o = 0.64 \pm 0.03$, for E1371Q; $P_o = 0.88 \pm 0.03$, for E1371Q + GLPG1837. ***$p < 0.001$ by Student's t test. (C) The presence of 1.5 **μ**M GLPG1837 shifts the apparent K_d of ivacaftor from 6.6 ± 1.2 nM (dashed line) to 54 ± 4 nM (solid line, n = 9). (D) Competition binding assay. Ivacaftor was kept at a constant concentration of 8 nM, and its binding to WT CFTR is plotted as a function of GLPG1837 concentration. Ki = 0.30 ± 0.08 **μ**M. (E) Ribbon diagram of the phosphorylated, ATP-bound CFTR in complex with GLPG1837 (stick representation, orange). (F) EM density at the potentiator-binding site in the potentiator-free (left), ivacaftor-bound (center), and GLPG1837bound (right) reconstructions. Densities corresponding to CFTR are shown in gray, whereas densities only observed in the potentiator-bound maps are shown as green meshes. All maps are contoured at $9\boldsymbol{\sigma}$.

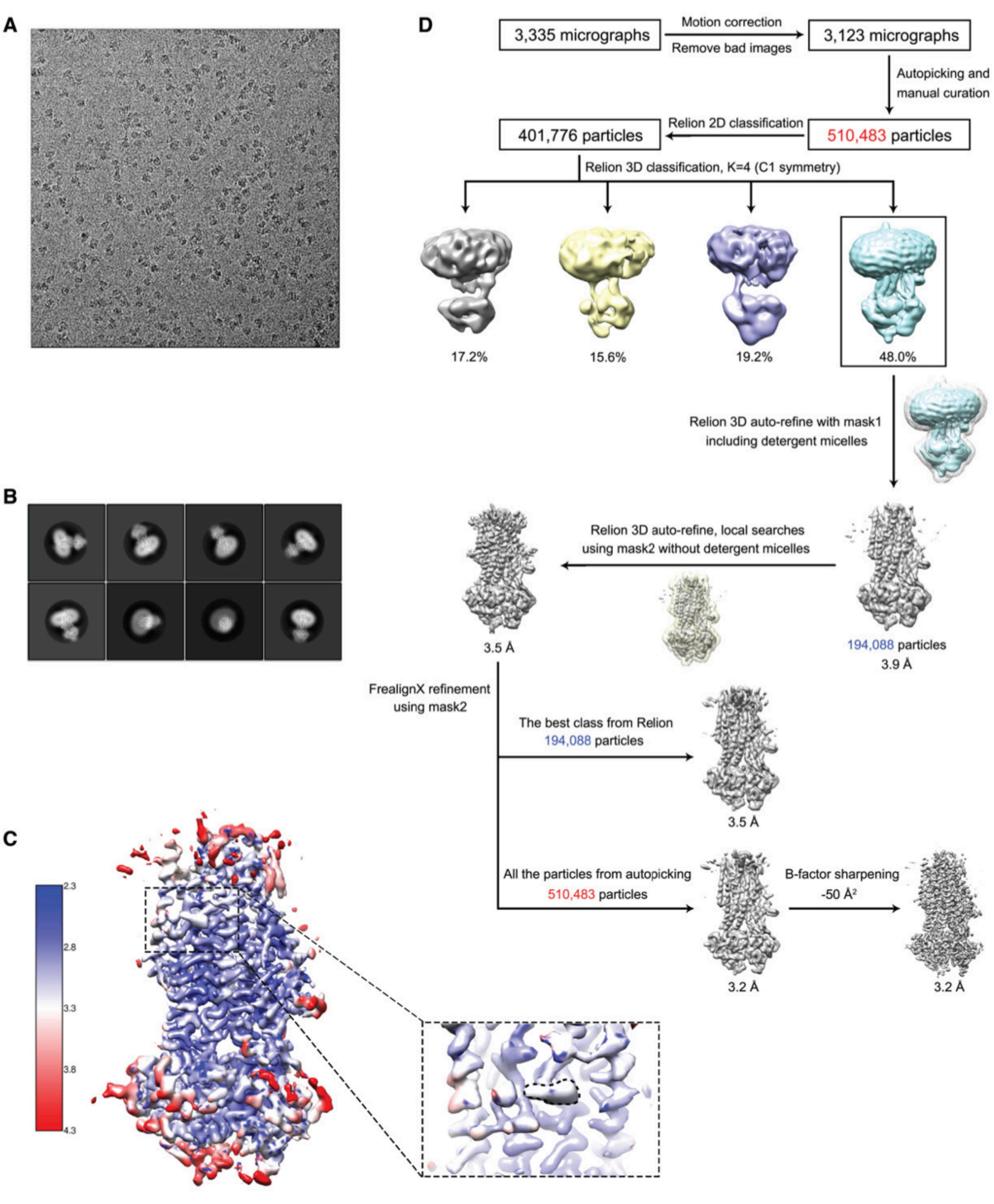


Fig. 5. 9 Cryo-EM analysis of the CFTR/GLPG1837 complex
(A) A representative micrograph at defocus of -1.5 μm. (B) Representative 2D class averages from Relion [58]. (C) Overall EM map colored by local resolution estimation using Blocres [57]. Inset: zoom-in view of the drug binding site for GLPG1837; GLPG1837 is outlined with black dash line. (D) Summary of image processing for the CFTR/GLPG1837 data.

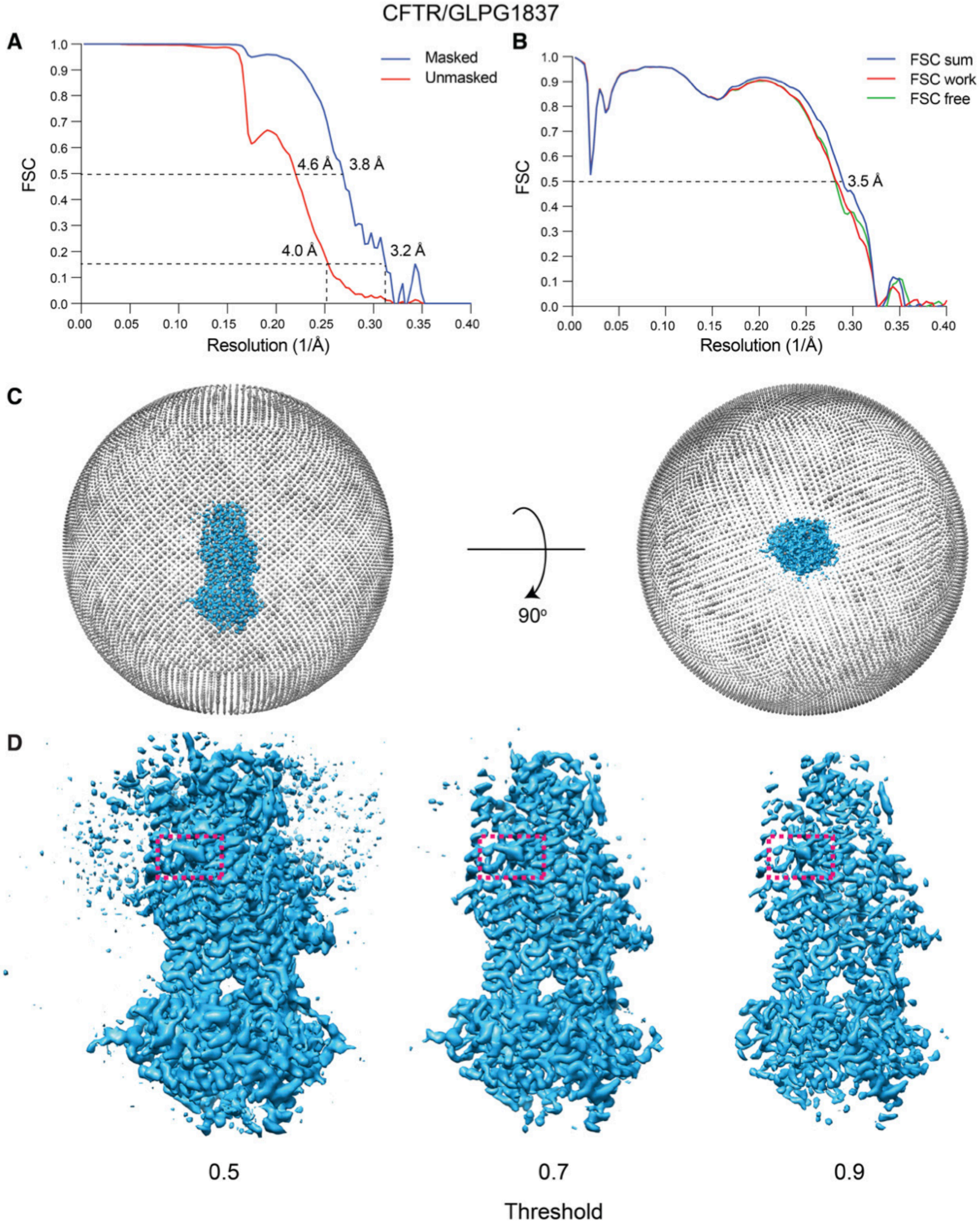


Fig. 5. 10 Data and model quality assessment for CFTR/GLPG1837 complex
(A) Fourier Shell Correlation (FSC) (blue: masked; red: unmasked) as a function of resolution, using FrealignX [59] output. (B) FSC calculated between the full, half map for refinement (FSC work), half map for validation (FSC free) and the refined structure. (C) Angular distribution of the particles used in the final reconstruction. Each sphere indicates the number of particles from this angle and the size of the spheres corresponds to the number of particles. (D) CFTR/GLPG1837 map shown in three different thresholds in Chimera. At threshold 0.5, GLPG1837 density is much stronger the micelle density; at 0.7, GLPG1837 density is as strong as the protein density; at 0.9, GLPG1837 density is still prominent as some main chain density starts to disappear.

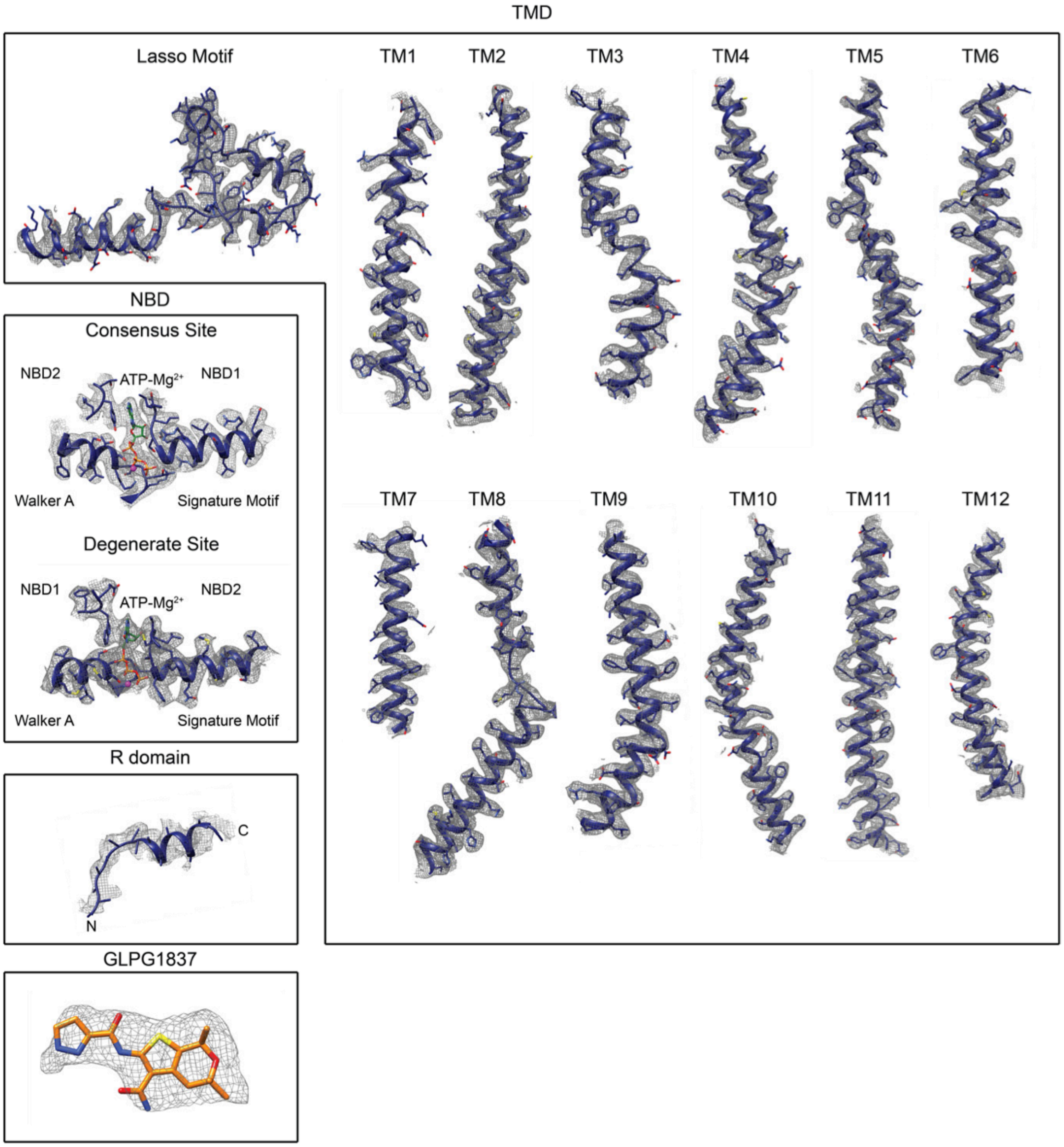


Fig. 5. 11 EM density of different parts of the CFTR/GLPG1837 complex

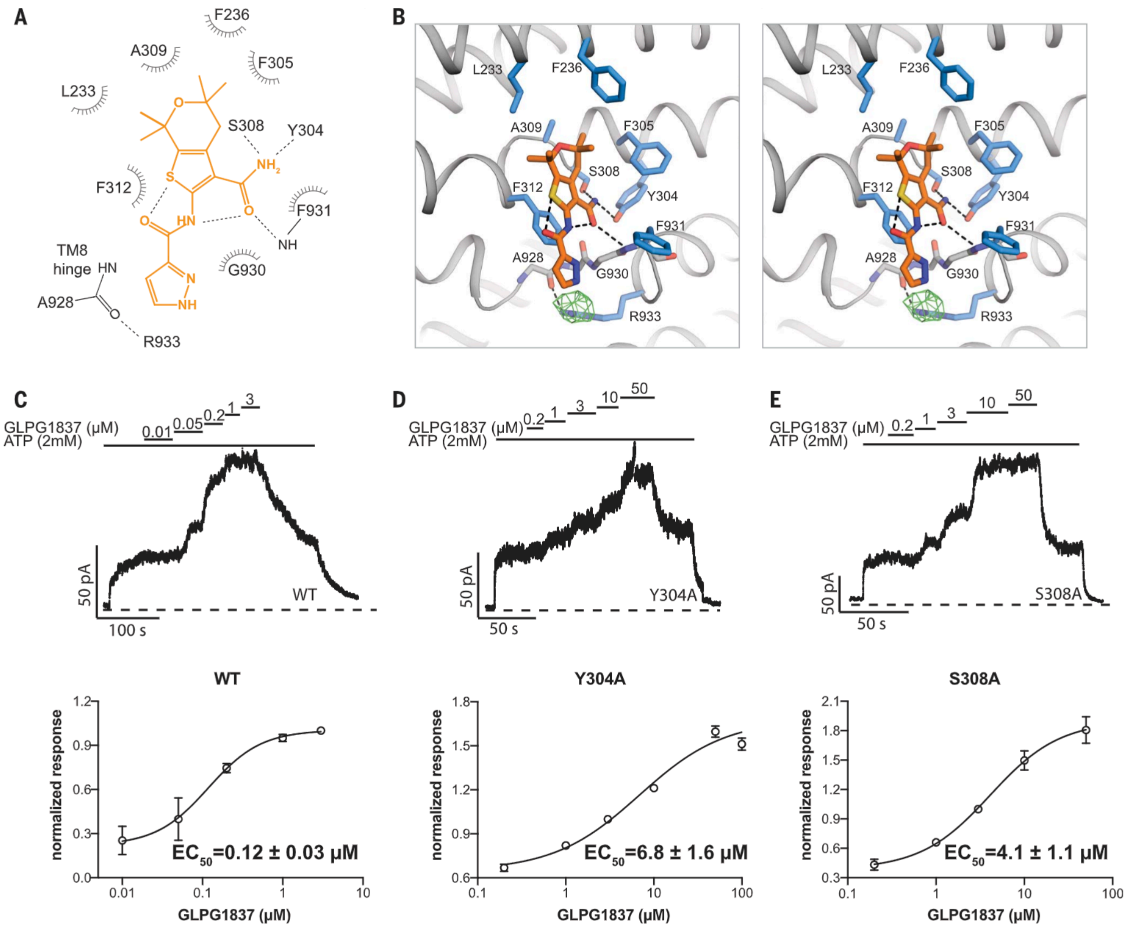


Fig. 5. 12 Molecular details of GLPG1837 binding

(A) Schematic drawing of the interactions between GLPG1837 (orange) and CFTR. Hydrogen bonds are represented by black dashed lines, and hydrophobic interactions are shown by the spokes. (B) Stereo view of the GLPG1837 binding site. Residues within van der Waals distances (<4.5 Å) are shown as blue sticks, and hydrogen bonds are depicted as black dashed lines. An unknown density between R933 and GLPG1837 is shown in green mesh. (C to E) Upper panel: Representative macroscopic current traces of WT and mutant CFTR in response to GLPG1837 perfusion. Different concentrations of GLPG1837 applied after channel activation are marked above the trace. Lower panel: Dose-response curves with estimated EC50 values. CFTR-containing membrane patches were fully phosphorylated by PKA in the presence of saturating amount of ATP before GLPG1837 titration. The total current at 3 **μ**M GLPG1837 was used to normalize the current potentiated by different concentrations of GLPG1837. The dose responses were fitted with the Hill equation. Each data point represents values determined from five to nine patches.

Table 5. 3 Summary of EM data and structure refinement statistics for CFTR/GLPG1837 complex

Data collection and processing	
Microscope	Titan Krios (FEI)
Voltage (kV)	300
Detector	K2 Summit (Gatan)
Pixel size (Å)	1.03
Defocus range (μm)	0.8 to 2.5
Frames/movie	50
Exposure rate (electrons/pixel/s)	8
Total dose (electrons/Å^2)	75
Data acquisition software	SerialEM
Micrographs collected/processed (no.)	3,335/3,123
Particles extracted (no.)	510,483
Particles for 3D/final reconstruction (no.)	401,776/510,483
Symmetry imposed	C1
Model composition	
Non-hydrogen atoms (no.)	9,711
Protein residues (no.)	1,196
GLPG1837/Lipids/ATP/Mg^{2+} (no.)	1/6/2/2
Refinement	
Resolution at 0.143 FSC threshold (Å) masked/unmasked	4.0/3.2
Resolution at 0.5 FSC threshold (Å) masked/unmasked	4.6/3.8
Local resolution range (Å)	2.3 - 4.3
Sharpening B-factor (Å^2)	-50
Model refinement package	Coot and REFMAC (reciprocal space refinement)
RMS deviations	
Bond lengths (Å)	0.0063
Bond angles (°)	1.2607
B-factors of the model	
Protein residues	168.5
ATPs	149.6
Mg^{2+}	132.6
Lipids	146.5
GLPG1837	162.3
Validation	
Molprobity score	1.67
Clashscore, all atoms	6.19
Favored rotamers (%)	94.4
CaBLAM outliers (%)	0.94
Average FSC	0.70
Average CC	0.53
EMRinger score	3.04
Ramachandran plot (%)	
Favored/Allowed/Outliers	94.4/6.6/0.0

5.3 Material and methods

5.3.1 Cell culture

Sf9 cells were cultured in Sf-900 II SFM medium (GIBCO) supplemented with 5% FBS and 1% Antibiotic-Antimycotic. HEK293S GnTI$^-$ cells were cultured in Freestyle 293 (GIBCO) supplemented with 2% FBS and 1% Antibiotic-Antimycotic. Chinese hamster ovary (CHO) cells were cultured in DMEM-F12 (ATCC) supplemented with 10% FBS and 1% Antibiotic-Antimycotic.

5.3.2 Mutagenesis

All mutations were introduced using QuikChange Site-Directed Mutagenesis System (Stratagene) or a mutagenesis methodology develop by Liu and Naismith [168].

5.3.3 Protein expression and purification

All CFTR constructs were expressed and purified as described [39, 92]. Briefly, bacmids carrying CFTR constructs were generated in *E. Coli* DH10Bac cells (Invitrogen). Recombinant baculoviruses were produced and amplified in Sf9 cells. Proteins were expressed in HEK293S GnTI$^-$ cells infected with 10% baculovirus at a density of $3x10^6$ cells/ml. Cells were induced with 10 mM sodium butyrate 12 hours after infection and cultured at 30°C for another 48 hours before harvesting. For protein purification, cell membranes were solubilized in buffer containing 1.25% 2,2-didecylpropane-1,3-bis-β-D-maltopyranoside (LMNG) and 0.25% Cholesteryl hemisuccinate (CHS). Protein was purified via its C-terminal green fluorescence protein (GFP) tag using GFP nanobody coupled Sepharose Beads (GE Healthcare) and eluted by removing the GFP tag with the PreScission Protease. The sample was phosphorylated using protein kinase A (NEB) and then further purified on size exclusion chromatography.

5.3.4 EM data acquisition and processing

The phosphorylated E1371Q sample (5.5 mg/mL in 0.06% digitonin) was incubated with 10 mM ATP and $MgCl_2$ plus either 200 μM ivacaftor or GLPG1837 on ice for 15 min. About 3 mM fluorinated Fos-choline-8 was added to the sample right before freezing on to Quantifoil R1.2/1.3 400 mesh Au grids using Vitrobot Mark IV (FEI). EM images were collected on a 300 kV Titian Krios (FEI) with a K2 Summit detector (Gatan) in super- resolution mode using Serial EM. The defocus ranged from -0.8 to -2.5 µm and the dose rate was 8 e^-/pixel/sec.

The images of the CFTR/Ivacaftor dataset were first corrected for gain reference and binned by 2 to obtain a physical pixel size of 0.817 Å. Whole frame drift correction was corrected using Unblur [95] and contrast transfer function (CTF) estimation was performed using CTFFIND4 [96]. After manual inspection to remove images with bad ice, 5,434 out of 5,778 collected images were selected for further processing (Fig. 5.5). Particles were automatically picked by Relion [58]. Each image was manually inspected to remove false positives (aggregates, detergent micelles, and ice contaminates), which accounted for approximately 10% of auto-picked particles. A total number of 511,493 particles were used for 2D classification, and subsequently 490,262 particles were used for 3D classification in RELION2 [169]. The phosphorylated, ATP-bound human CFTR map (EMD-9230) was used as the initial reference. The best class, containing 34% of the total particles, was auto refined to 7.13 Å (with no mask) in RELION2 and used as the initial model for subsequent refinement in FrealignX [59]. The particle stack used in FrealignX was generated from the auto-picked particles and motion corrected using alignparts_lmbfgs [97]. This process removed particles at the edge of the images and reduced the total number of particles from 511,493 to 490,750 (Fig. 5.2D). After testing many refinement parameters, the best reconstruction was obtained using 6.5 Å as the high-resolution limit for particle alignment. The map used for model building (Fig. 5.2C) was low-pass filtered to 3.3 Å and sharpened with a B-factor of -50 Å^2 using the BFACTOR.EXE program (written by Nikolaus Grigorieff).

The CFTR/GLPG1837 dataset was processed similarly, except for the following differences (Fig. 5.9). The binned images have a physical pixel size of 1.03 Å. Beam- induced sample motion was corrected using MotionCor2 [170]. CTF estimation was performed using Gctf [171]. The particle auto-picking was done by Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang). After 3D classification, the best class, containing 194,088 particles, were used for 3D auto-refine. The refinement using a loose mask including the entire micelles density (mask1) yielded a 3.9 Å map. Application of a tighter mask (mask2) for local refinement increased the resolution to 3.5 Å. Further refinement in FrealignX using particles in the best class (194,088) produced a map with similar to that of RELION2 (3.5 Å). Refining all particles (510,483) improved the resolution to 3.2 Å, likely due to the internal weighting mechanism implemented in FrealignX [59]. The best reconstruction was obtained using 6 Å as the high-resolution limit for particle alignment. The final map was low-pass filtered to 3.2 Å and sharpened with a B-factor of -50 Å^2 using BFACTOR.EXE (written by Nikolaus Grigorieff).

5.3.5 Model building and refinement

The model building and refinement of CFTR were carried out as described [39]. In brief, the dataset was randomly split into two halves, one half for model building and refinement (working set) and the other half for validation (free set). The model was built in Chimera [109] and Coot [99] and refined in reciprocal space with Refmac [102, 103]. MolProbity [107, 108] was used for geometries validation and Blocres [57] was used for local resolution estimation. Final structures contain residues 1–390 of TMD1; 391–409, 437–637 of NBD1; 845–889, 900–1173 of TMD2; 1202–1451 of NBD2; and 17 residues of R domain (built as alanines). Figures were generated with PyMOL and Chimera.

Molecular docking of ivacaftor and its analog series, as well as GLPG1837 into the identified binding site was performed using DOCK3.7 [164], which places pre-generated ligand conformations into the binding site by superimposing atoms of each molecule on matching spheres, representing favorable positions for individual ligand atoms. The resulting docked ligand poses were scored by pre-calculating energy grids which evaluate the different energy terms of the DOCK scoring function - the AMBER van der Waals potential (http://ambermd.org); Poisson-Boltzmann electrostatic potentials using QNIFFT [172, 173]; and ligand desolvation from the occluded volume of the target for different ligand orientations [174]. Receptor structures were protonated using Reduce [175]. Partial charges from the united-atom AMBER (http://ambermd.org) force field were used for all receptor atoms. Ligand protonation states and tautomers were calculated using JChem (v15.11.23.0, www.chemaxon.com), at pH 7.4. Each protomer was rendered into 3D using Corina (v3.6.0026, www.molecularnetworks.com) and conformationally sampled using Omega (v2.5.1.4, www.eyesopen.com) [176]. Ligand charges and initial solvation energies were calculated using AMSOL [177, 178]. The resulting ligand-receptor complexes were further optimized through energy minimization with the AMBER protein force field and the GAFF ligand force field supplemented with AM1BCC charges (http://ambermd.org). Figures were generated with PyMOL and Chimera.

5.3.6 Inside-out patch clamp recording

All CFTR constructs used for electrophysiology were cloned into the BacMam expression vector. A GFP tag was fused to the C-terminal of CFTR for visualization. Chinese hamster ovary (CHO) cells were plated at 0.4×10^6 cells per 35-mm dish (Falcon) 12 hours before the transfection. For each dish, cells were transfected with 1μg BacMam plasmids using Lipofectamine 3000 (Invitrogen) in Opti-MEM media (Gibco). After 12 hours of transfection, media was changed to DMEM: F12 (ATCC) supplemented with 10% FBS and the cells were incubated at 30°C for 2 days before recording.

Macroscopic currents were recorded in inside-out excised patches, using buffer compositions and recording parameters as described [167]. The culture media were first changed to a bath solution consists of 145 mM NaCl, 2 mM $MgCl_2$, 5 mM KCl, 1 mM $CaCl_2$, 5 mM glucose, 5 mM HEPES, and 20 mM sucrose, pH 7.4 with NaOH. The pipette solution contains 140 mM NMDG, 5 mM $CaCl_2$, 2 mM $MgCl_2$, and 10 mM HEPES (pH 7.4 with HCl). The perfusion solution contains 150 mM NMDG, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 10 mM EGTA, and 8 mM Tris (pH 7.4 with HCl). Borosilicate micropipettes (OD 1.5 mm, ID 0.86 mm, Sutter) were pulled to 2-5 MΩ resistance. After a gigaseal was formed, inside-out patches were excised and exposed to 25 units/ml PKA (Sigma-Aldrich) and 2 mM ATP to phosphorylate CFTR. In subsequent perfusions, 6 units/ml PKA and 2 mM ATP was used to maintain phosphorylation. Currents were recorded at -30 mV and 25°C using an Axopatch 200B amplifier, a Digidata 1550 digitizer and pCLAMP software (Molecular Devices). The recordings were low pass filtered at 1 kHz and sampled at 20 kHz. Data were analyzed with Clampfit and GraphPad Prism. To derive the EC50 values of GLPG1837, currents were normalized with the amount of current at 3 μM GLPG1837 for each patch as described [167], and the dose-response curves were fitted using the Hill equation.

5.3.7 Planar lipid bilayer recording

A lipid mixture containing 1-palmitoyl-2-oleyl-sn-glycero-3-phosphoethanolamine (POPE), 1-palmitoyl-2-oleyl-sn-glycero-3-phospho-(1'-rac-glycerol) (POPG) 3:1 (w/w) was resuspended at 20 mg/ml by sonication in buffer containing 200 mM NaCl, 2 mM $MgCl_2$, and 20 mM HEPES (pH 7.2 with NaOH). Lipids were mixed with 2 % glyco- diosgenin (GDN) for 1 hour at 25°C. Purified CFTR was added to the lipid mixture at a protein-to-lipid ratio of 1:100 (w/w) and incubated for 2 hours at 4°C. β-cyclodextrin was added to the lipid/protein mixture at a 1.2x molar ratio to GDN. β-cyclodextrin addition was repeated twice with each addition separated by 6 hours. Proteoliposomes were pelleted and resuspended in buffer containing 200 mM NaCl, 2 mM $MgCl_2$, and 20 mM HEPES (pH 7.2 with NaOH).

Planar lipid bilayers were made by painting a 1,2-dioleoyl-sn-glycero-3-phosphoetanolamine (DOPE), 1-palmitoyl-2-oleyl-sn-glycero-3-phosphocholine (POPC), and 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS) 2:1:1 (w/w/w) lipid mixture solubilized in decane, across a ~100 μm diameter hole on a plastic transparency. Proteoliposomes containing CFTR were phosphorylated with protein kinase A (NEB) for one hour at 25°C and fused with planar lipid bilayers. Currents were recorded at 25°C in symmetric buffer containing 150 mM NaCl, 2 mM $MgCl_2$, 2 mM ATP, and 20 mM HEPES (pH 7.2 with NaOH). Voltage was clamped at 150 mV with an Axopatch 200B amplifier (Molecular Devices). Currents were low pass filtered at 1 kHz, digitized at 20 kHz with a Digidata 1440A digitizer (Molecular Devices), and recorded using pCLAMP software (Molecular Devices). Displayed recordings were low-pass filtered at 100 Hz. Data were analyzed with Clampfit and GraphPad Prism.

5.3.8 Scintillation proximity assay

CFTR constructs used in this assay contain a C-terminal Strep-tag, followed by a PreScission Protease cleavage site, and a GFP tag. The GFP tag is removed during purification whereas the Strep tag is retained to attach CFTR to the SPA beads.

To measure ivacaftor binding, 5 nM CFTR was incubated with 0.5 mg/ml YSi- streptavidin SPA beads (PerkinElmer) in the presence of 0.06-200 nM ivacaftor (the molar ratio of [^{3}H] ivacaftor (21 Ci/mmol, synthesized by Moravek): cold ivacaftor is 1:3) in Buffer B (20 mM Tris-HCl pH 7.5, 200 mM NaCl, 0.06% digitonin, 2 mM DTT and 0.1% Tween 20) at 4°C for 1 hr. The reactions were carried out in 96-well non-binding surface microplates (Corning) and data were recorded using a Microbeta Trilux plate reader (PerkinElmer). Specific binding was obtained by subtracting background radioligand binding in the absence of protein. The K_d values of ivacaftor were calculated by fitting the data with a single-site saturation binding model accounting for ligand depletion by 5 nM of CFTR in GraphPad Prism version 8.0.0 for Mac, GraphPad Software, San Diego, California USA, www.graphpad.com. The K_i of GLPG1837 was calculated by fitting the data with a single-site competitive binding model accounting for ligand depletion [179]. The readings were normalized by dividing the specific binding with the total binding (Bmax) and represented in percentages.

CHAPTER 6. Structural study of ABCA4

6.1 Summary

ABCA4 is the only ABC importer in mammals. Mutations in ABCA4 are linked to several eye diseases such as Stargardt disease, the most common inherited single-gene retinal disease. In this chapter, we report two cryo–electron microscopy (cryo-EM) structures of ABCA4 in the absence or presence of ATP to 3.3 Å. These structures reveal the conformational change of ABCA4 along its transport cycle. Careful structural analyses provide a structural interpretation of known disease-causing mutations. Moreover, both structures reveal a hydrophilic pocket in the transmembrane domains and a hydrophobic extracellular tunnel that may play a critical role for ABCA4's transport activity. With these structures, we can begin to design experiments to understand the unique functional properties of this transporter.

6.2 Results

6.2.1 Biochemical Characterization

Full-length human ABCA4 was expressed in mammalian cells and purified to homogeneity as evidenced by its gel filtration profile (Fig. 6.1A). The protein ran at a size of around 250 kDa by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), about the expected molecular weight of full-length ABCA4 (Fig. 6.1A).

To examine the function of the purified ABCA4, I measured the ATPase activity of wild-type ABCA4 (WT ABCA4) and the double-mutant E1087Q/E2096Q (ABCA4-MM) (Fig. 6.1B). The conserved glutamic acid at position 1087 and 2096 are general bases catalyzing ATP hydrolysis. Mutation of the corresponding residues in other ABC transporters such as in P-glycoprotein, CFTR, and maltose transporters abolished ATP hydrolysis but not ATP binding [112, 180, 181]. WT ABCA4 has a K_m of 0.08 ± 0.01 mM for ATP, comparable to the reported values of 0.09 mM or 0.075 mM [182, 183]. The maximal ATPase activity of ABCA4 reconstituted in brain polar lipid liposomes was reported to be around 100 nmol/mg/min [184]. Our detergent purified WT ABCA4 has a similar maximal ATPase activity (112.5 ± 2.8 nmol/mg/min, corresponding to a specific turnover rate of 0.5 ATP per second). The ABCA4-MM construct has a similar size exclusion profile as the WT protein, with no measurable ATPase activity (Fig. 6.1). These data indicate the samples purified in detergent function similarly to the membrane embedded ABCA4 and are suitable for structural studies.

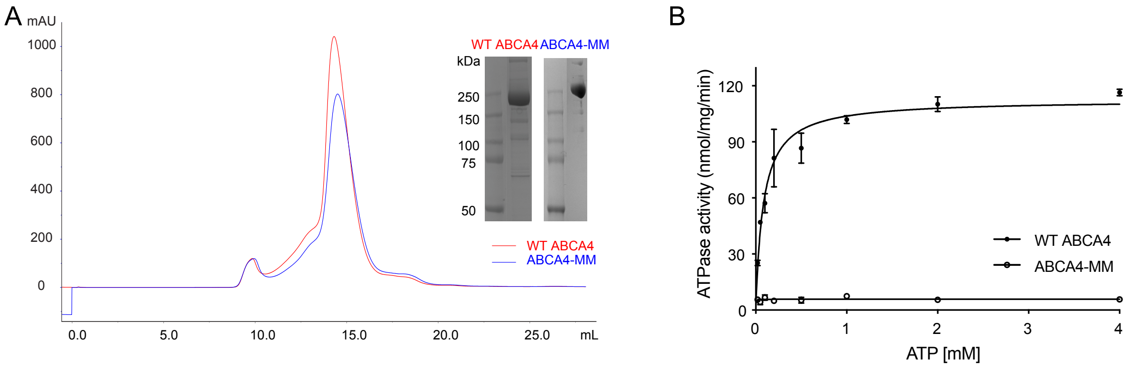


Fig. 6. 1 Biochemical characterization of ABCA4

(A) Size exclusion profile and SDS-PAGE of the purified WT ABCA4 and the E1087Q/E2096Q mutant (ABCA-MM). The protein is purified in Buffer A (20 mM HEPES pH 8, 200 mM NaCl, 0.06% Digitonin).

(B) ATPase activities as a function of ATP concentration at 28°C. Data points represent the means and standard deviations (SDs) of three measurements from a single preparation. WT ABCA4 shows a K_m of 0.08 ± 0.01 mM for ATP. The maximal ATPase activity is determined to be 112 ± 3 nmol/mg/min (assuming a molecular weight of 256 kDa, the specific turnover rate is 0.5 per second).

6.2.2 Overall overview of ATP-free ABCA4 structure

The structure of WT ABCA4 was determined by single molecule cryo-EM in the absence of ATP and substrate. The overall resolution of the EM density is estimated to be 3.3 Å (Fig. 6.2A). The overall EM density is of high quality except for the part of the exocytoplasmic region, suggesting local high mobility. The final model contains aa 3-2253, excluding aa 138-272, and loop regions between residues 468-473, 487-497, 873-916, 937-947, 1162-1201, 1277-1340, 1900-1911 and 2172-2178. Polyalanines are built for regions aa273-279, 473-480, 484-487.

ABCA4, a single polypeptide of 2273 residues, is one of the largest molecules in the ABC transporter superfamily. It consists of two homologous halves, each containing a TMD, a large extracellular domain (ECD), an NBD, and a cytoplasmic regulatory domain (RD) (Fig. 6.3B).

The two extracellular domains (ECDs), located between the first two helices in TMD1 and TMD2, respectively, extend 130 Å away from the membrane surface (Fig. 6.3C). The large size of the ECDs is consistent with their localization in the rim regions of the disc lumens, as the central region of the discs is not large enough to accommodate ABCA4 (ranges from 0 to 120 Å in height across different species) [185]. Each TMD contains six transmembrane helices (TM1-6 and 7-12) and two short helices (EHs) that penetrate into the outer leaflet of the membrane bilayer. In the absence of ATP, the two NBDs are apart from each other; similarly, the two TMDs are separated. In contrast, extensive interdomain interactions were observed between the two ECDs and RDs (Fig. 6.3C).

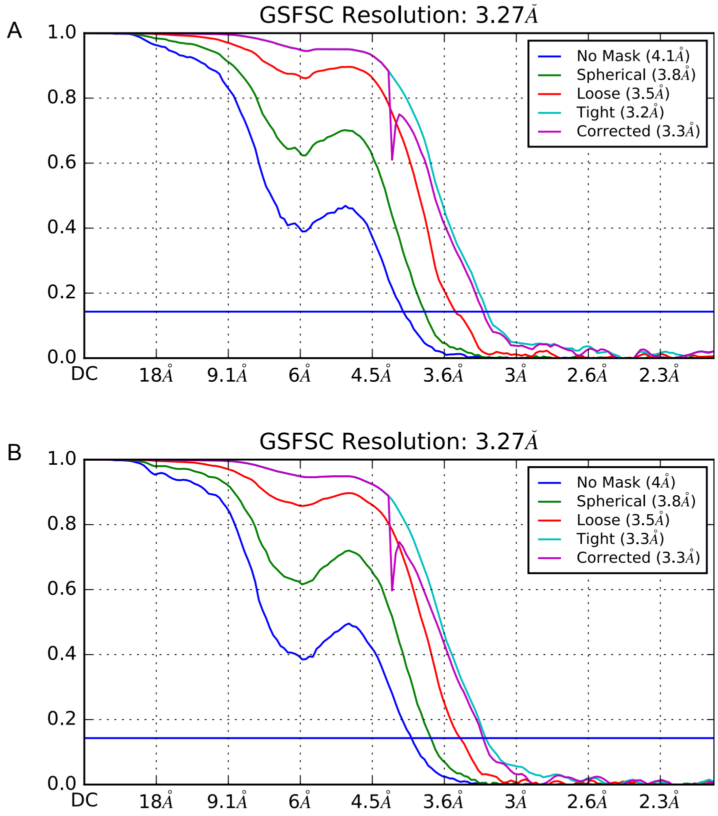


Fig. 6. 2 Data and model quality assessment
Plots of the Fourier Shell Correlation (FSC) between reconstruction of two half-datasets for WT-ABCA4 (A) or ABCA4-MM (B).

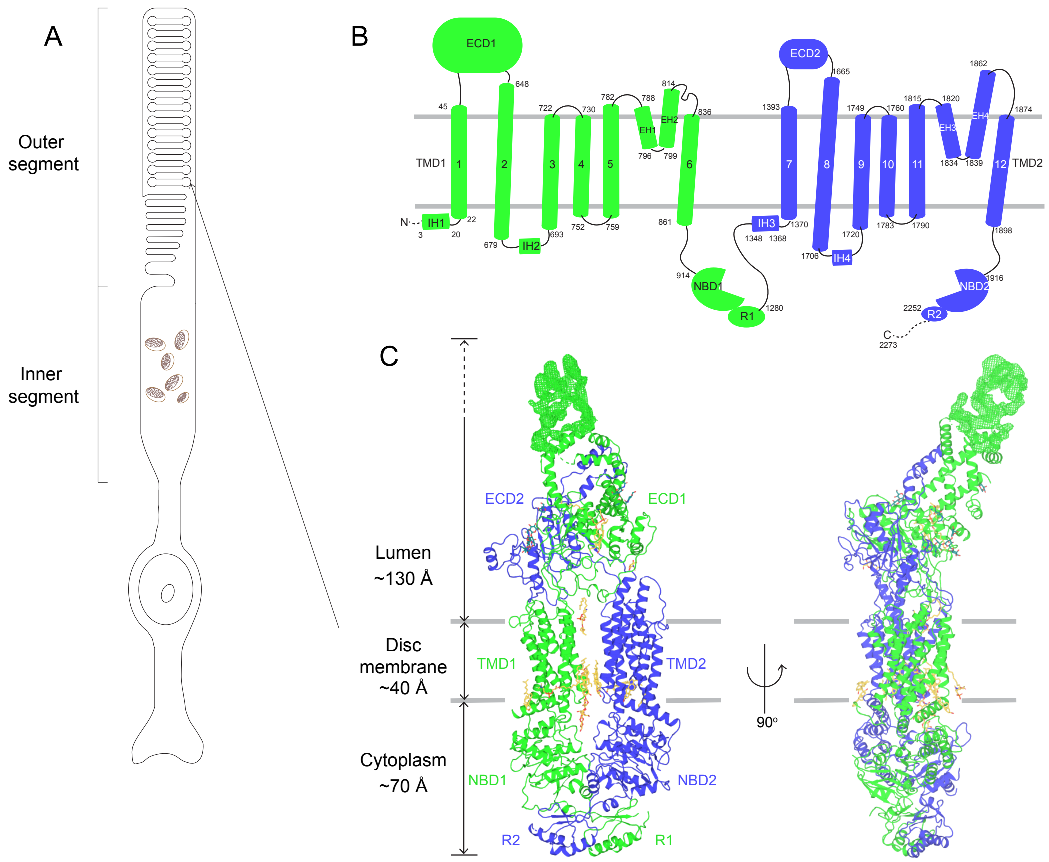


Fig. 6. 3 The overall structure of ATP-free ABCA4

(A) An illustration showing a rod photoreceptor cell. A black arrow indicates the cellular localization of ABCA4 to the rod photoreceptor outer segment discs.

(B) The domain structure of ABCA4

(C) Two orthogonal views of ABCA4. Also shown in green the EM density corresponding to part of ECD1 (residues: 138-272). That are highly flexible. Ordered detergents, lipids and N-linked glycans are shown in stick model.

6.2.3. The extracellular domains (ECDs) are stabilized by disulfide bonds and form a hydrophobic tunnel

ECD consists of 603 residues between TM1/2. ECD2, located between TM7/8, has 285 residues. They do not contain any known enzymatic or functional motives. However, many disease mutations are mapped to the ECDs [186-237], indicating these domains are important in function.

The two ECDs pack closely together near the extracellular surface of the membrane, capping the outward-facing TM cavity (Fig. 6.4A). The structure of the ECDs is stabilized by extensive disulfide bonds. One pair of interdomain (C641-C1490) and five pairs of intradomain disulfide bonds were observed in the structure (C54-C81, C75-C324, C370-C519 in ECD1 and C1444-C1455, C1488-C1502 in ECD2) (Fig. 6.4A, right panel). Out of those cysteines, seven missense mutations, C54Y, C75G, C519R, C641S, C1455R, C1488R, C1490Y, have been identified in patients with retinal diseases [188, 238-240]. Consistent with functional studies [241], N-linked glycosylation were observed on residues N98, N415, N444, N504, N1469, N1529, N1588 and N1662. Most of the glycosylation sites are highly conserved and missense mutations at positions 415, 504, 1588 are linked to retinal diseases [224, 242, 243].

A prominent structural feature of the ECDs is the hollow interior. A tunnel, approximately 50 Å in length and 6-15 Å in diameter, is enclosed by helices from both ECDs (Fig 6.4B). The interior surface of the tunnel is largely hydrophobic.

The role of extracellular domain of ABCA4 has remained a conundrum. Previous functional study of the isolated ECD2 shows that it binds to all trans-retinal with a K_d of 2-4 μM [244]. In the structure, we do observe strong densities inside the hydrophobic tunnel that may correspond to lipid or detergent molecules (Fig 6.4C); therefore, it is possible that the ECDs function to scavenge trans-retinal that mis-localize to the luminal space of the outer segment discs.

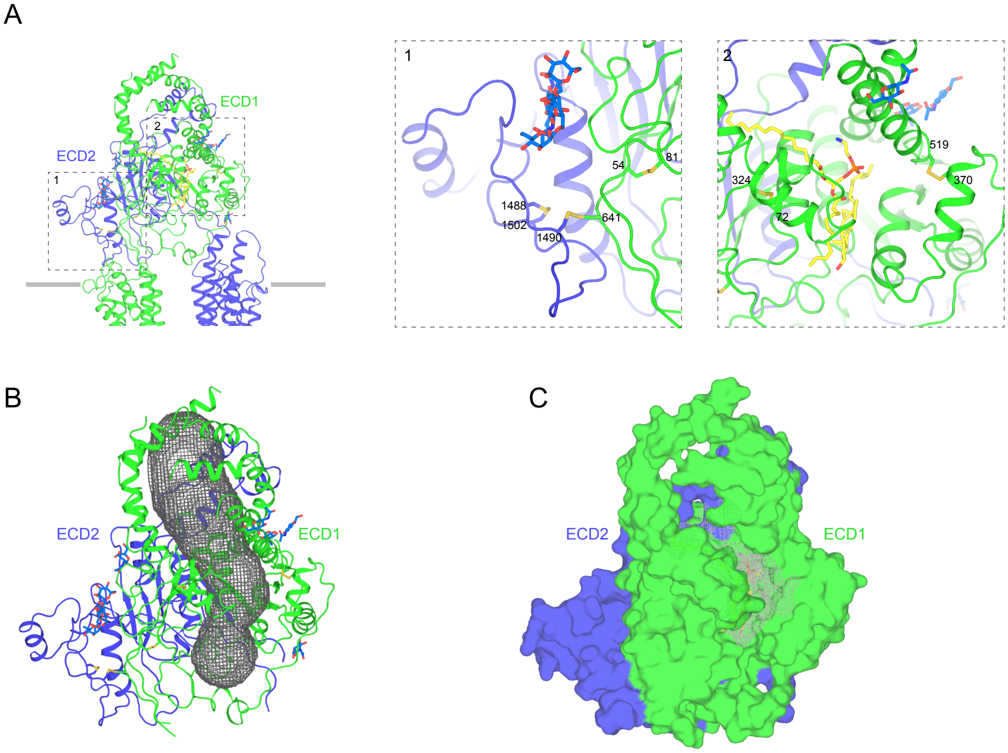


Fig. 6. 4 hydrophobic tunnel formed by extracellular domains (ECDs)
(A) The overall architecture of ECDs is stabilized by inter- and intra-domain disulfide bonds. The density for the disulfide bonds is shown in gray mesh. Five out of six disulfide bonds pairs are shown (C641-C1490, C54-C81, C1488-1502, C370-C519, C72-C324).
(B) ECD1 and ECD2 fold together to form a tunnel-like structure in the luminal space of outer membrane segment discs. The diameter of the tunnel is mapped by the program HOLE and shown in gray mesh.
(C) Extra EM densities inside the tunnel that may correspond to bound lipids. The ECDs are represented in surface. The extra densities are shown in mesh.

6.2.4. The transmembrane domains (TMDs) contains a hydrophilic pocket

Different from most eukaryotic ABC transporters in which the TM helices domain-swap to pack against the opposite TMD, the TMDs of ABCA4 does not exhibit a domain swapping configuration. Each TMD folds into a compact structure without contacting the other TMD (Fig. 6.3C, 6.5A). A large cavity is formed at the interface of two TMDs, opening to the luminal media and the outer leaflet of the membrane. The gap between the TMDs at the level of the inner leaflet is filled with ordered lipid and detergent molecules. The overall configuration of the TMDs is outward facing, but unlike any other known transporters, the cytoplasmic gate is formed by lipid molecules instead of protein. In addition, since several ordered lipid molecules are also found at protein/bilayer, it is likely that they play a role in regulating the function of ABCA4. In line with this hypothesis, digitonin was reported to be allosteric stimulator of ABCA4 [245].

Another unusual structural feature of ABCA4 is a large hydrophilic pocket formed by EH3 and 4 and TM12 of TMD2. Many charged residues on the surface of the indentation are exposed to lipids, as far as 15 Å from the membrane surface (Fig. 6.5B, C). We imagine that the structure of the lipid bilayer would be locally deformed to accommodate these charged residues. The density in this pocket is relatively poor comparing to the rest the TMDs, suggesting high mobility. It is tempting to hypothesize that this region binds the head group of the lipid substrate to facilitate translocation. Several clinical mutations, R1820P, R1843W and E1885, have been reported to cause Stargardt disease [188, 194]. Future experiments will be carried out to test the functional role of this region.

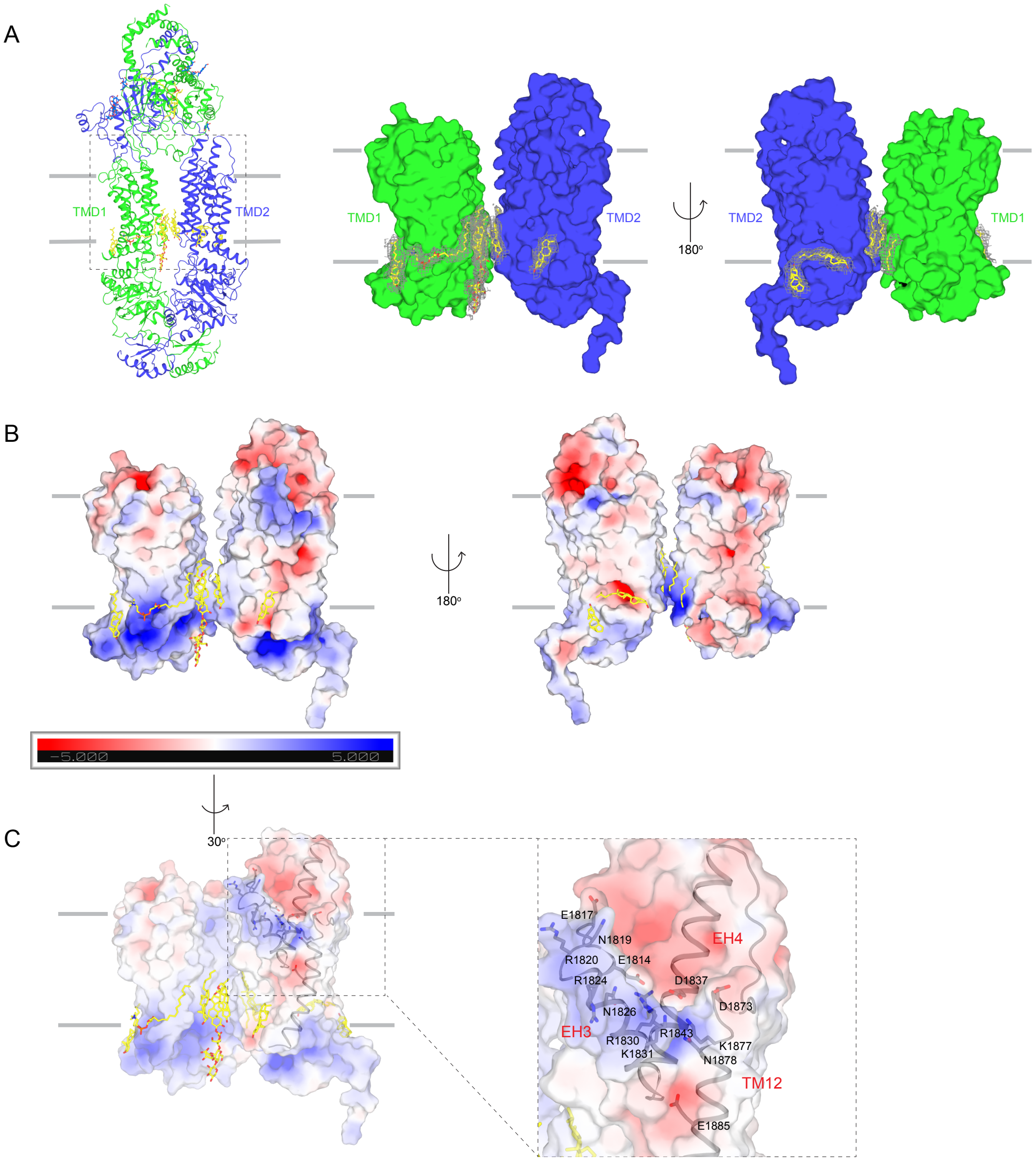


Fig. 6. 5 Transmembrane domains (TMDs) adopt an outward facing conformation

(A) Zoomed in view of the transmembrane domains shown in surface. The TMD1 is in green and TMD2 is in blue. The transmembrane domains maintain their contact mainly through the lipids that pack in the inner leaflet (lipids and a digitonin are shown in yellow with their densities shown in gray).

(B) Electrostatic surface representation of the TMDs, calculated assuming a pH of 7 and a concentration of 0.15 M of both (+1) and (-1) ions. Scale: red, negative (-5 kT/e); blue, positive (+5 kT/e).

(C) The hydrophilic indentation below the boundary of the outer membrane. The electrostatic surface is in 50% transparency. The side chains of the hydrophilic residues in the pocket are shown and labeled (E1814, E1817, N1819, R1820, R1824, N1826, R1830, K1831, R1843, D1837, D1873, K1877, N1878, E1885). EH3, EH4 and TM12 are shown in thin, grey ribbons.

6.2.5. Regulatory domains (RDs) stabilize NBD dimers and contain ACT-fold domains

The cytoplasmic region of ABCA4 consists of two NBDs and R domains, which forms a dimer in a domain-swapped fashion (Fig. 6.6A and B). The regulatory domain C-terminal to NBD1 (i.e. R1) crosses the dimer interface to interact with NBD2, and R2 crosses over the interface to pack against NBD1. In the absence of ATP, the NBD dimer is stabilized by interactions between charged residues. For example, R1097 and R1098 from NBD1 form salt bridges with D2242 in R2. R2107 in NBD2 salt bridges with E1270 in R1 (Fig. 6.6A). This stabilization is likely important for ABCA4's function as disease causing mutations such as R2107H [246], R1097S [209], R1098C [194] were observed in Stargardt patients.

The two R domains have very similar structure, both exhibiting an ACT-like fold which adopts a βαββαβ topology [247]. Typical ACT motifs bind regulatory small molecules [248]. The only other ABC transporter known to have ACT motifs is the bacterial methionine importer MetNI [249]. In MetNI, the two ACT motifs also form a dimer but with an interface very different from that of ABCA4 (Fig. 6.6C). In ABCA4, the interface is formed by α2 and β4 but in MetIN, α1 and β2 mediate the domain-domain interaction (Fig. 6.6B and C). It is shown that when the intracellular concentration of methionine is high, it binds to the ACT motifs of NetNI to inhibit further uptake of methionine [249]. Sequence alignment indicates that the ACT motif is conserved among the 12 members of the ABCA subfamily. Whether they with any regulatory ligands is currently unknown.

It was shown that an NBD1-R1 construct bind cis-retinal with 80 nM affinity. It also interacts with trans-retinal ($K_d > 2\ \mu M$) in a dose-dependent manner [250]; however, further investigation is required to understand whether these interacts are mediated through R1 or is physiologically relevant.

In addition to the ACT motif, R2 in the ABCA family also contains a highly conserved VFVNFA motif [251, 252]. Mutations that result in removal of the VFVNFA motif of A1, A3 and A4 have been reported in patients with Tangier disease, surfactant deficiency syndrome and Stargardt disease, respectively [251, 253-256]. Functional studies have shown that the VFVNFA motif is necessary for proper folding of ABCA1 and ABCA4 [257]. In ABCA4, this motif forms a helix pinning at the interface of two NBDs, interacting with two highly conserved sequences: The H loop of one NBD and the D loop of the other NBD (Fig. 6.6D). In the transport cycle, the two NBDs open and close like a pair of tweezers akin to the maltose transporter [124], and the helix containing the VFVNFA motif is located at the base of the tweezer. Consistently, FRET analysis showed that the VFVNFA motif is important in promoting NBD association [258].

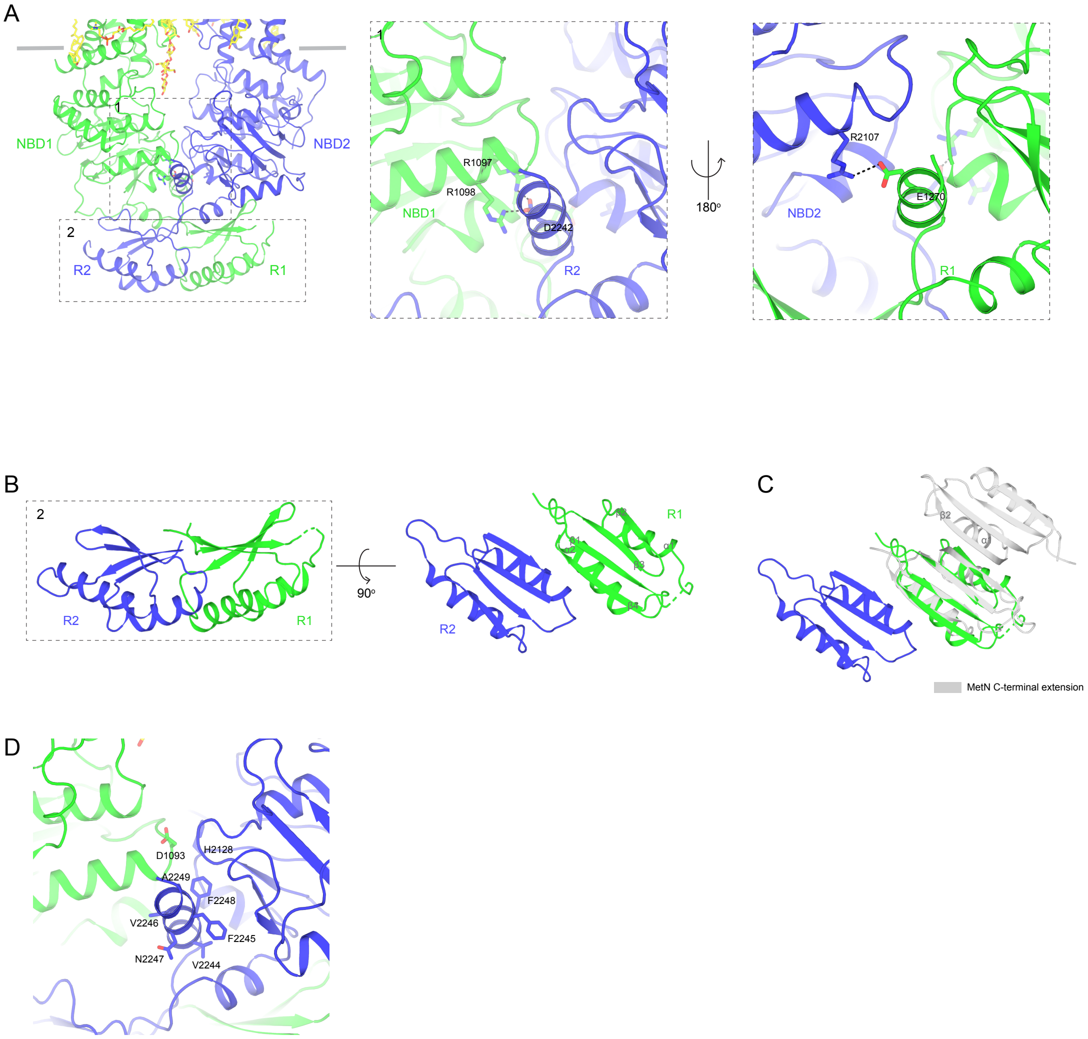


Fig. 6. 6 Regulatory domains (RDs) stabilize NBD dimers and each contains an ACT-fold domain

(A) R1 and R2 are domain swapped relative to the NBDs. NBD1 interact with R2 and NBD2 interact with R1. Disease-causing mutants, R1097, R1098 and R2107, are engaged in stabilizing the interactions between NBD and R domains [194, 209, 259].

(B) Zoomed-in view of R1 and R2 dimers for both front and top view. Both R1 and R2 adopts an ACT-fold with a ferredoxin-like $\beta\alpha\beta\beta\alpha\beta$ topology. The interface between RDs are formed by $\alpha 2$ and $\beta 4$.

(C) Superposition of RD-dimer with a MetN C-terminal extension dimer which also adopts ACT fold (PDB: 3DHW). MetN is shown in gray. The alignment between R1 and one MetN C-terminal extension yields a RMSD of 2.74 Å. Unlike RD-dimer, MetN C-terminal extension dimerizes through $\alpha 1$ and $\beta 2$.

(D) The VFVNFA motif. The VFVNFA motif (residues 2244-2249) in the C-terminus of ABCA4 is interacting with conserved H loop (H2128) of NBD2 and D loop (D1093) of NBD1.

6.2.6. Conformational change of ABCA4 upon ATP binding

To capture ABCA4 in its ATP-bound state, we used the ABCA4-MM mutant that binds but not hydrolyzing ATP. The protein was purified in the presence of 1 mM ATP and frozen onto cryo-EM grids with an additional 9 mM of ATP-Mg^{2+}. Analyzing the cryo-EM data through 3D classification showed that nearly 100% of the 333,731 are in the NBD-dimerized configuration. The overall resolution was about 3.3 Å (Fig. 6.2B). Two ATP molecules, 1,933 residues, and 5 lipid molecules are well resolved in the density map.

In comparison to the ATP-free structure, the presence of ATP brought upon large conformational changes that can be described as rigid body rotations of individual domains (Fig. 6.7A). The two ECDs rotate as a unit for about 30^{o} along an axis normal to the membrane surface. The two TMDs move towards the molecular center to form contact with each other. No cavity nor lipids are observed between the two TMDs (Fig. 6.7B). the two NBDs form a closed dimer with two ATP bound at its interface. The R domains rotate towards the membrane, forming a continuous eight-stranded antiparallel β sheet (Fig. 6.7C). As a result, new contacts are formed between the R domains and NBDs (Fig. 6.7D), which appear to be functionally important. For example, H2202 forms hydrogen bond with D1102, and R2107 forms salt bridges with E1270 and E1223 simultaneously. Both D1102Y and R2107H are observed in patients, with R2107H being the most frequent mutation in African American (19.3%).

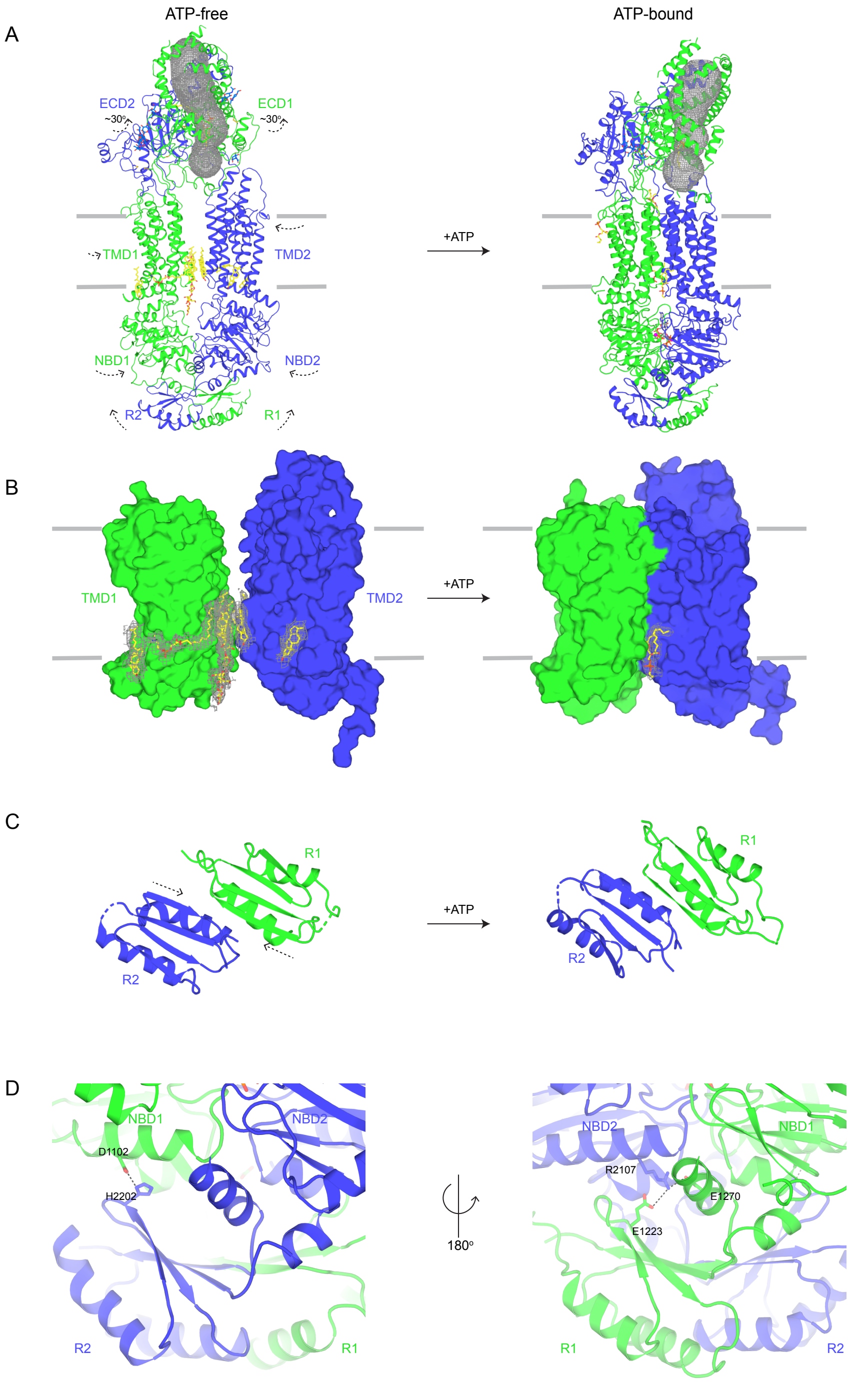


Fig. 6. 7 Conformational change upon ATP binding

(A) Overall conformational change. The relative motion of each domain is indicated in dashed arrow. The tunnel in the ECDs is shown in gray mesh.
(B) Conformational change of the TMDs. The TMDs are shown in surface representation.
(C) Conformational change of the R domains. The dashed arrows indicate the relative motion of each R domain upon ATP binding.
(D) Salt bridge interactions between R domain and NBDs in the ATP-bound conformation.

6.2.7 Molecular details of ATP binding

The signature motifs of ABCA4 deviate from the consensus sequence. Instead of LSGGQ, it has a sequence of YSGGN and LSGGM in NBDs, respectively; however, the two ATP molecules are bound in a manner typical of ABC transporters (Fig. 6.8A and B). Each ATP interacts residues in the walker A, B motifs of one NBD and signature motif of the other NBD. The Mg^{2+} ions are coordinated between beta, gamma phosphates and residues from walker A motif and the Q loop (Fig. 6.8A and B, right panel). In addition, residues outside of conserved motifs also contribute to ATP binding. T2070 makes a hydrogen bond with the adenine ring of the ATP in ATPase site 1, and K1054 hydrogen bonds with the sugar group in ATPase site 2 (Fig. 6.8A and B).

The integrity of the ATP binding composite sites is extremely important for the correct function of ABCA4 as both sites are hot spots for Stargardt disease mutations. For example, N965Y [196], N965S [8], G966E [243], T970P [260], T970I [226] , T971N [202] on walker A motif, H1118D [261], H1118Y [262] of the H loop, Y2071F [202], S2072N [263], G2074V [264], G2074S [226], G2074D [243] in the signature motif and T2070R [220] of the first ATPase site, and S1063P [194], M1066R [263] in the signature motif, G1972R [204], N1974S [226], G1975S [194], K1978E [234], T1979I [204] in the Walker A motif, D2095V [217] in the Walker B, H2128R [198] of the H loop at the second ATPase site were all observed in patients. Substitutions of the catalytic glutamic acids, E1087K [8], E1087D [204], E1087G [195] and E2096K [188], E2096G [191], E2096V [191], were also identified in patients.

It was shown that ABCA4 can bind both ATP and GTP with the same affinity [182, 183]. Consistently, our structures indicate that GTP can bind at both ATP-binding sites without any steric clashes. As rod photoreceptors contain similar levels of ATP and GTP [265], this finding suggests that ABCA4 can utilize both nucleotides to power retinal translocation.

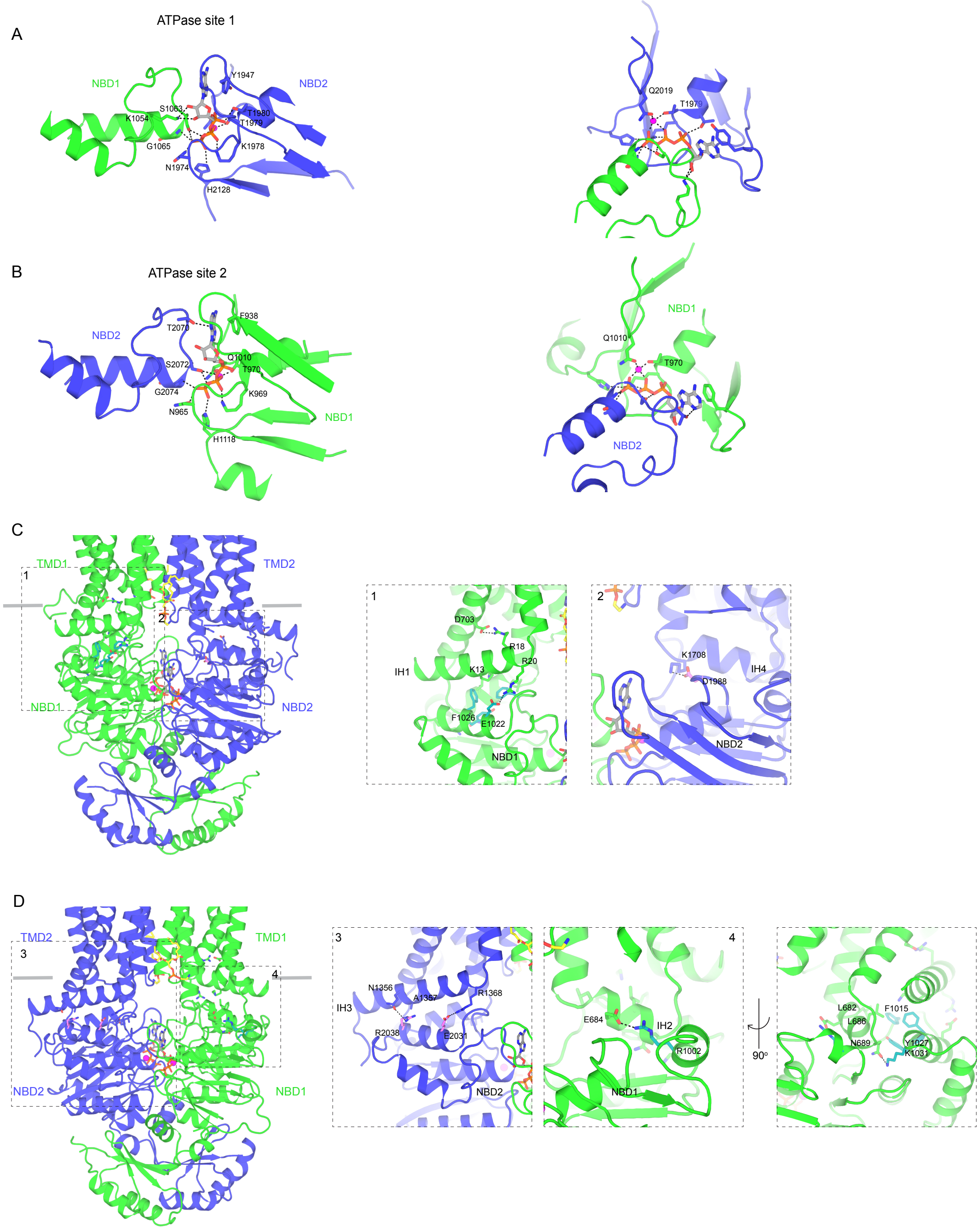


Fig. 6. 8 Molecular details of ATP binding and the TMD/NBD interface

(A-B) Symmetric ATP binding at the ATP-binding site 1 (A) and ATP-binding site 2 (B). Side chains of the interacting residues are shown in stick model. Hydrogen bonds are indicated by dashed lines. The coordination of the Mg^{2+} is shown in the right panel. Mg^{2+} ions are in magenta. (C-D) Energy coupling sites from NBDs to TMDs are hotspots for disease-causing mutations. Inner helices (IH1 to IH4) interact with NBDs through ionic and hydrophobic interactions.

6.2.8 Molecular details of the TMD/NBD interface

Conformational changes brought upon ATP binding are transmitted to the TMDs through TMD/NBD interface. IH1 and IH3 maintain favorable interactions with NBDs before or after ATP binding, for example, through the salt bridges between R20 on IH1 and E1022 in NBD1, R1368 on IH3 and E2031 in NBD2 (Fig. 6.8C and D). On the other hand, IH1 and IH3 interact with NBDs more closely upon ATP binding. When ATP binds to ATPase site 1, R1002 in NBD1 moves ~7 Å towards IH1 to form a salt bridge with E684, and in ATPase site 2, D1988 in NBD2 shifts towards IH4 to form a salt bridge with K1708 on IH3 (Fig. 6.8C and D). These newly formed favorable salt bridges and the surrounding hydrophobic interactions help relay the signal from ATP binding to the transmembrane movement.

Residues that involve in energetic coupling are hotspots for disease mutations. So far, mutations are identified in three out of four interfaces—E1022G (interact with IH1), F1026L (interact with K13 on IH1 through cation-pi interaction), R18P (on IH1 and salt bridges with D703 in TM helix 3); R1002W (interact with IH2 in ATP-bound state); F1015I/L (form tight hydrophobic packing with L682/L686 on IH2 and Y1027 in NBD1); R1031E (forms ionic interaction with Q689); A1356T/E/V (pack tightly with surrounding residues; mutation may lead to disruption of the ionic interaction between Q1356 on IH3 and R2028); R1368C (salt bridges with E2031) (Fig. 6.8D).

6.2.9 Structural comparison of ABCA4 with ABCA1

The human ABCA has 12 members, many are involved in lipid transport and are associated with hereditary diseases [8, 266]. Despite high degrees of sequence similarity within the subfamily, ABCA4 is the only member that transport substrate inward. All other members are exporters. To identify any structural basis underlying the unique functional property of ABCA4, we compared its structure with ABCA1, which shares 51% identical sequence to ABCA4 but functions as a lipid exporter. The structure of ABCA1 in its ATP-free, substrate-free conformation was determined to 4 Å resolution [54], which is very similar to that of the *apo* form of ABCA4, with overall RMSD of 2.64 Å. The largest differences occur in the ECDs. In residues 136-257, ABCA1 are better structured than ABCA4, where only amorphous densities were observed for residues 138-272 in ABCA4 (Fig. 6.3C, green density). Furthermore, residues 100-323, 435-504 and 1541-1613 show large displacement upon superposition (Fig. 6.9A, dashed arrow). Consistently, the sequence identity of these regions (25%, 34% and 31% respectively) are lower than the overall value of 51%. Biochemical studies suggest ECDs in ABCA1 are involved in binding to its ligand apolipoprotein A-I for efficient cholesterol efflux [267]. As we have discussed earlier, ECD1 in ABCA4 was shown to bind retinal [244]. It remains to be tested if the structural differences correlate with their role in binding to substrates or ligands.

One of the prominent features of ABCA4 is the hydrophilic indentation at the luminal side of the membrane surface (Fig. 6.5C). The corresponding region in ABCA1 was poorly resolved; most of the residues were built as poly alanines [54]. However, a conservation analysis between A1 and A4 (https://consurf.tau.ac.il/) (Fig. 6.8B) indicates that many hydrophilic residues forming the indentation are conserved, suggesting that ABCA1 also has a conserved hydrophilic indentation like ABCA4.

The structure of the ABCA1's R domain is very different from that of ABCA4 (Fig. 6.9C). Instead of the ACT-fold (βαββαβ), the two R domains of ABCA1 has a topology of αββαββ (Fig. 6.9C). Moreover, the R domains are in non-domain swapped fashion relative to the NBDs (Fig. 6.9D). Given the high sequence identities (43% for R1 and 58% for R2), such large structural difference seems improbable. The reconstruction of the ABCA1 shows poor density in the R domain, thus it is likely that the published structure is incorrect.

Thus far, our structural analyses have not revealed how these two very similar proteins transport substrates in opposite directions. We look forward to future studies to address this question.

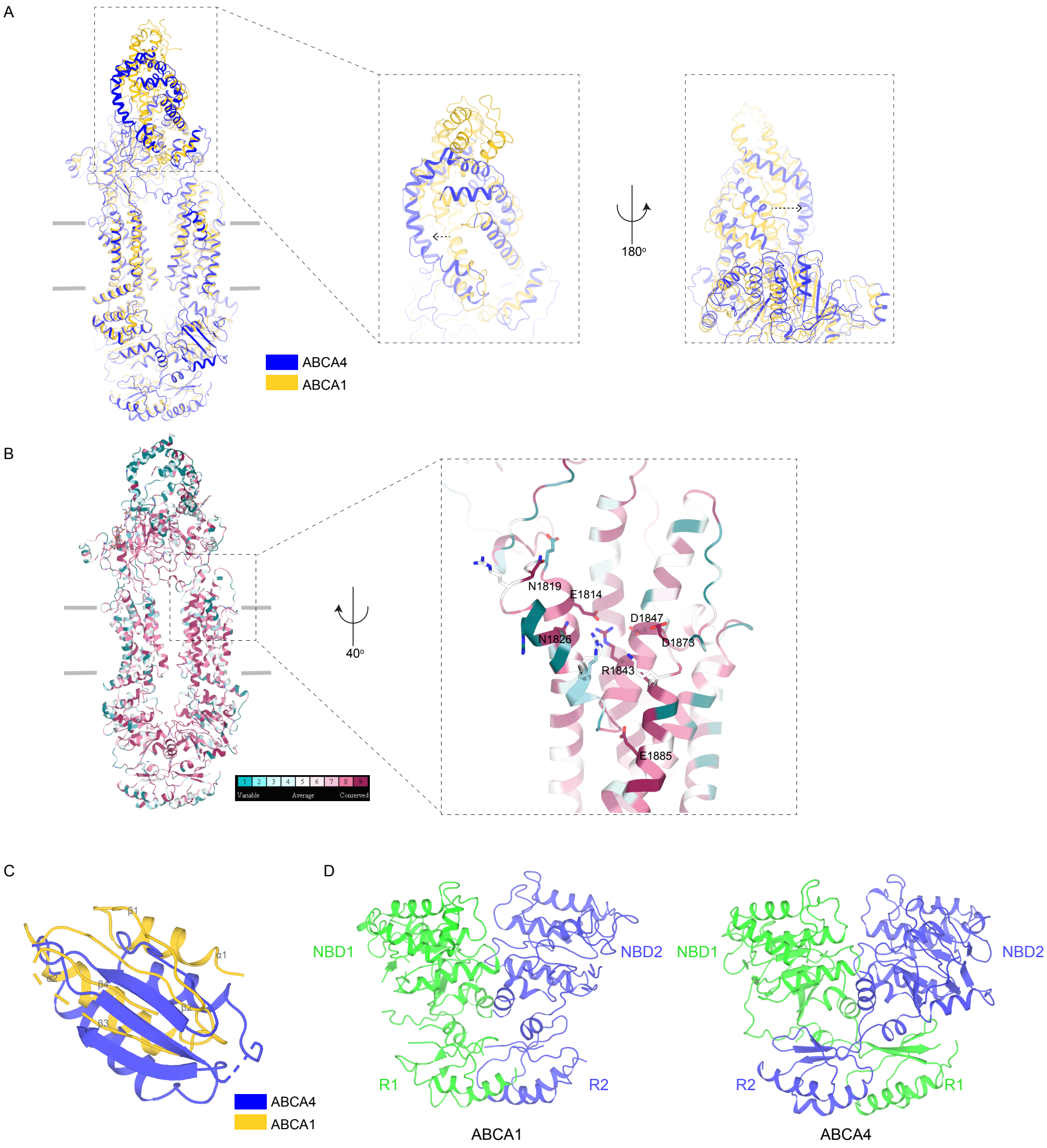


Fig. 6. 9 Structural comparison of ABCA4 with ABCA1

(A) Overall superposition of ABCA1 and ABCA4. The structural rearrangement of ABCA4 relative to ABCA1 is indicated in dashed arrow. ABCA4 is shown in blue and ABCA1 is in gold.

(B) Sequence conservation of ABCA1 and ABCA4 from mouse, bovine and human mapped on the structure of hABCA4. The residues in the hydrophilic indentation are shown in sticks, and the conserved residues are labeled.

(C) Structural superposition of one R domain of ABCA1 and ABCA4. The secondary structure of ABCA1 is labeled.

(D) Position of R domains relative to NBDs in ABCA1 and ABCA4.

6.3 Material and methods

6.3.1 Cell culture

Sf9 cells were cultured in Sf-900 II SFM medium (GIBCO) supplemented with 5% FBS and 1% Antibiotic-Antimycotic. HEK293S GnTI$^-$ cells were cultured in Freestyle 293 (GIBCO) supplemented with 2% FBS and 1% Antibiotic-Antimycotic.

6.3.2 Mutagenesis

All mutations were introduced using QuikChange Site-Directed Mutagenesis System (Stratagene).

6.3.3 Protein expression and purification

DNA encoding ABCA4 gene was synthesized and codon-optimized for expression in mammalian cells (BioBasic) and subcloned into a BacMam expression vector [92] with a C-terminal green fluorescent protein (GFP) tag. The resulting plasmid was transformed to DH10Bac *E. coli* cells to produce bacmid DNA. Recombinant baculoviruses were first generated in Sf9 cells maintained in sf-900 II SFM medium (GIBCO) plus 5% FBS using Cellfectin II reagents (Invitrogen). The resulting P1 viruses were amplified two more generations and 10% P3 virus was used to infect HEK293S GnTI$^-$ suspension cells at $3x10^6$ cells/mL in Freestyle 293 medium (GIBCO) supplemented with 2% FBS. Infected cells were incubated at 37^{o}C for 12 hours. Next, the temperature was decreased to 30^{o}C and 10 mM sodium butyrate was added to induce protein expression for 48 hours [92]. To purify the protein, cells were resuspended in lysis buffer (20 mM HEPES pH 8, 2 mM $MgCl_2$, 200 mM NaCl, 20% Glycerol) supplemented with protease inhibitors (1 μg/ml Leupepetin, 1 μg/ml Pepstatin, 1 μg/ml Aprotonin, 100 μg/ml Trypsin inhibitor, 1 mM Benzamidine, and 1 mM PMSF) and DNase (3 μg/ml) and solubilized with 1.25% (w/v) 2,2-didecylpropane-1,3-bis-b-D-maltopyranoside (LMNG) and 0.25% (w/v) Cholesteryl hemisuccinate (CHS) for 2 hours at 4^{o}C. The cell lysates were centrifuged at 70,000 g for 1 hour and the supernatant was added to CNBR-activated Sepharose beads (GE Healthcare) coupled to a GFP nanobody [268]. The beads were washed with Buffer A (20 mM HEPES pH 8, 200 mM NaCl, 0.06% Digitonin) and then incubated with PreScission protease (5:1 w/w ratio) at 4^{o}C for 3 hours to remove the C-terminal GFP tag. The protein was eluted with Buffer A, and further purified with gel filtration chromatography using a Superose 6 10/300 column (GE Healthcare) equilibrated with Buffer A. For purification of ABCA4-MM construct, 1 mM of ATP was supplemented to both lysis buffer and buffer A.

6.3.4 EM data acquisition and processing

The ABCA4-MM sample (5.5 mg/mL in 0.06% digitonin) was incubated with 9 mM ATP and $MgCl_2$ plus on ice for 15 min before sample freezing. About 3 mM fluorinated Fos-choline-8 was added to the sample right before freezing on Quantifoil R1.2/1.3 400 mesh Au grids using Vitrobot Mark IV (FEI). EM images were collected on a 300 kV Titian Krios (FEI) with a K2

Summit detector (Gatan) in super- resolution mode using Serial EM. The defocus ranged from -0.8 to -2.5 µm and the dose rate was 8 e^-/pixel/sec.

The images of both datasets were first corrected for gain reference and binned by 2 to obtain a physical pixel size of 1.03 Å. Beam-induced sample motion was corrected using MotionCor2 [170]. CTF estimation was performed using Gctf [171]. The particle auto-picking was done by Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang).

For ABCA4-WT dataset, 5226 images were collected. 1,375,284 particles were picked by Gautomatch and imported in cryoSPARC v2 [269] for 2D classification, and 425,348 particles were selected for ab-initio reconstruction. The best class (320,102 particles) was chosen for Non-uniform refinement to yield a map of 3.64 Å. The half maps, masks, and particles from cryoSPARC were imported to Relion [58] through pyem suite (https://github.com/asarnow/pyem) for post-processing and polishing. The polished particles were then imported to cryoSPARC for another round of Non-uniform refinement to yield a final map of 3.27 Å.

The ABCA4-EQ datasets were processed similarly with the following differences and the best map was yielded from two separate datasets. For the first dataset, 6,766 images were collected, and 857,074 particles were picked by Gautomatch. 2D classification yielded 455,456 particles for ab-initio reconstruction, and the best class (287,299 particles) was chosen for Non-uniform refinement to obtain a 3.57 Å map. Further postprocessing and polishing in Relion [58] yielded a map of 3.37 Å. For the second dataset, 2,304 images were collected, and 278,932 particles were picked. 2D classification yielded 121,894 particles for ab-initio reconstruction, and the best class (46,432 particles) was chosen for Non-uniform refinement to obtain a 3.83 Å map. Further postprocessing and polishing in Relion [58] yielded a map of 3.52 Å. The best classes from both datasets (287,299 particles and 46,432 particles) were combined after polishing and imported into cryoSPARC to obtain a final map of 3.27 Å with Non-uniform refinement.

6.3.5 Model building and refinement

The model building and refinement of ABCA4 was carried out as described [39]. In brief, the dataset was randomly split into two halves, one half for model building and refinement (working set) and the other half for validation (free set). The model was built in Chimera [109] and Coot [99] and refined in real space with Phenix [270]. Phenix Elbow was used for adding constraints for the lipid and detergent molecules. Final structure of ABCA4-WT contains residues 3-45, 645-873 of TMD1; 46-137, 273-468, 473-487, 497-644 of ECD1; 916-937, 947-1151 of NBD1; 1152-1162, 1201-1277 of R1; 1340-1395, 1664-1900 of TMD2; 1396-1663 of ECD2; 1911-2161 of NBD2; 2162-2172, 2178-2253 of R2; eight cholesterol; two phosphatidylcholines. Out of the built residues, aa273-279, 473-480 and 484-487 are built as alanines. Final structure of ABCA4-MM contains residues 3-45, 645-881 of TMD1; 46-114, 117-136, 266-303, 308-331, 339-350, 362-456, 459-470, 493-644 of ECD1; 914-1151 of NBD1; 1152-1164, 1199-1280 of R1; 1342-1395, 1664-1902 of TMD2; 1396-1663 of ECD2; 1916-2161 of NBD2; 2162-2252 of R2; two ATPs; two Mg^{2+}; five phosphatidylethanolamine. Out of the built residues, aa117-136, 310-313, 447-456 and 493-503 are built as alanines. Figures were generated with PyMOL and Chimera.

6.3.6 ATPase assay

The ATPase activity was measured using an ATP/NADH consuming coupled method [93] using Buffer B (50 mM HEPES pH 8.0, 150 mM KCl, 0.06% Digitonin, 2 mM $MgCl_2$). The ATPase reaction buffer consists of 0.2 μM ABCA4, 60 μg/ml pyruvate kinase, 32 μg/ml lactate dehydrogenase, 4 mM phosphoenolpyruvate, 0.15 mM NADH, and different amounts of ATP/Mg^{2+}. The fluorescence of NADH was excited at 340 nm and monitored at 445 nm for 45 min using an Infinite M1000 Microplate Reader (TECAN). Reaction mixtures in the absence of CFTR were also recorded to measure the background signal. All experiments were repeated three times and data were fitted to the Michaelis–Menten equation to calculate the Km and Vmax values using GraphPad Prism 8.

CHAPTER 7. Outlook and future directions

7.1 Outlook and future directions of CFTR project

Over the years, we have learned a tremendous amount about the structure-function relationship of CFTR. However, there are still many questions unanswered.

First, we have identified the binding site for two CFTR potentiators; however, since the binding of potentiators does not induce additional conformational change on the ATP-bound CFTR, it is hard to propose a detailed mechanism for their potentiating effect. Based on electrophysiological measurements, ivacaftor increases ion conductance in the presence and absence of ATP [161]. It seems possible that addition of ivacaftor to CFTR would induce an open state without the necessity of NBD dimerization. To test this hypothesis, movements of NBDs and TM helices induced by ivacaftor can be directly measured through single molecule fluorescence resonance energy transfer (FRET) or double electron-electron resonance (DEER). If additional conformational states are observed upon addition of potentiators, such distributions could be analyzed to derive a mechanism for the action of potentiators. Furthermore, such states may be captured through cryo-EM.

Second, our results indicate that the binding pocket for potentiators is likely a hotspot for regulation, therefore it is possible to utilize this binding pocket for *de novo* drug discovery. We are currently collaborating with Prof. Brian Shoichet's lab for virtual screening and have identified two novel compounds that can potentiate CFTR's activity; however, in order for a more accurate and rational design of novel CFTR potentiators, more factors need to be considered. For example, the potentiators bind at the protein lipid interface with half of the molecule exposed to the membrane, but the current docking campaign did not take the lipid component into the consideration. Moreover, there is an anonymous density found in this binding pocket near the residue R933. How this density changes the microenvironment and docking accuracy needs to be further investigated.

Beyond extension of my current project, many interesting questions may still be explored. For example, my thesis centered around wild type CFTR, but in a disease context, determining how different mutations affect CFTR may be more relevant in a pathological context. However, many CF causing mutations such as ΔF508 cause CFTR prone to degradation, it is therefore difficult to produce large amounts of recombinant protein for structural studies. In the future, it would be ideal to analyze the structures of these mutants in their native membrane environment. Advancements in cryo-electron tomography could help achieve this goal.

Outside of the structural field, technological advancements in recent years are very encouraging and provide opportunities to treat cystic fibrosis through different venues. For instance, gene therapy is not available to patients due to difficulties such as off-target effects of vectors or misidentification of correct cell types to target [271]. Nowadays, with the development of Crisper-Cas technology and single-cell sequencing, people are likely to push gene therapy further for cystic fibrosis treatment [272-274].

7.2 Outlook and future direction of ABCA4 project

For the ABCA4 project, even with the two structures, the transport mechanism of ABCA4 is not completely understood.

First, it is unclear how ABCA4 binds and releases its substrate. ABCA4 does not have defined substrate binding pockets like canonical ABC transporters. In the ATP-free state, ABCA4 adopts an outward facing conformation, while the TMDs are separated to contact with the lipid bilayers. Upon ATP binding, TMDs move towards each other such that no cavity is found in between. Obtaining the structure of ABCA4 in the substrate-bound state would be an important step towards understanding ABCA4 function. I have made several attempts to stabilize a substrate bound ABCA4, but so far without success. Through our structures, we identified a hydrophilic pocket that could be involved in binding the lipid headgroups of the substrate. Future mutagenesis experiments will be performed to test this hypothesis. If mutation in this pocket interferes with substrate translocation, it is possible to identify mutants that would lock ABCA4 in its pre-translocation state for structural determination.

Another important question to address is: Why do ABCA4 and ABCA1, while having very similar structures, transport substrates in opposite directions (ABCA1 is an exporter, while ABCA4 is an importer)? PE, with different length or degree of unsaturation, can all stimulate ATPase activity of ABCA4, which suggests the headgroup of PE is a determining factor for substrate transport [183]. The structures of the TMDs in ABCA1 and ABCA4 both permit lateral entry of lipids from both sides of the membrane. It is plausible that different amino acid sequences in the TMDs interact with PE versus PC head groups for substrate recognition and transport. One way to test this hypothesis is to use an ABCA1 and ABCA4 chimera protein. For example, we can make a construct with ABCA4's ECDs and NBDs and ABCA1's TMDs. If swapping of TMDs transforms ABCA4 to ABCA1, then we know that the difference in TMDs determines the specificity in directionality of transport. Then more detailed mutagenesis experiments can be done to nail down the specific residues responsible for substrate recognition and directionality of transport.

Finally, the function of the ECDs in ABCA4 requires investigation. The ECDs are very important, as there are hundreds of mutations found in ECDs that cause vision diseases. Based on our structures, we propose that the ECDs can act as a scavenging apparatus for trans-retinal mislocalized to the disc space. To test those hypotheses, photo crosslink experiments can be carried out to test if ECDs binds trans-retinal specifically.

In summary, the structures answered many questions in the field. From the structures of CFTR, we now understand how phosphorylation regulates CFTR and what transformations CFTR takes to function as an ion channel. ABCA4 structures offer interpretations of many missense mutations that lead to various vision diseases. But these structures are just a starting point to go deeper into mechanism.

References:

1. Theodoulou, F.L. and I.D. Kerr, *ABC transporter research: going strong 40 years on.* Biochem Soc Trans, 2015. **43**(5): p. 1033-40.
2. Quazi, F., S. Lenevich, and R.S. Molday, *ABCA4 is an N-retinylidene-phosphatidylethanolamine and phosphatidylethanolamine importer.* Nat Commun, 2012. **3**: p. 925.
3. Nichols, C.G., *KATP channels as molecular sensors of cellular metabolism.* Nature, 2006. **440**(7083): p. 470-6.
4. Boat, T.F., M.J. Welsh, and A.L. Beaudet, *The Metabolic Basis of Inherited Disease, 6th Edition - Scriver,Cr, Beaudet,Al, Sly,Ws, Valle,D.* New York: McGraw-Hill, 1989: p. 2649-2680.
5. Rommens, J.M., et al., *Identification of the cystic fibrosis gene: chromosome walking and jumping.* Science, 1989. **245**(4922): p. 1059-65.
6. Molday, R.S., *Insights into the Molecular Properties of ABCA4 and Its Role in the Visual Cycle and Stargardt Disease.* Prog Mol Biol Transl Sci, 2015. **134**: p. 415-31.
7. Burke, T.R. and S.H. Tsang, *Allelic and phenotypic heterogeneity in ABCA4 mutations.* Ophthalmic Genet, 2011. **32**(3): p. 165-74.
8. Allikmets, R., et al., *A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy.* Nat Genet, 1997. **15**(3): p. 236-46.
9. Gunderson, K.L. and R.R. Kopito, *Conformational states of CFTR associated with channel gating: the role ATP binding and hydrolysis.* Cell, 1995. **82**(2): p. 231-9.
10. Rees, D.C., E. Johnson, and O. Lewinson, *ABC transporters: the power to change.* Nat Rev Mol Cell Biol, 2009. **10**(3): p. 218-27.
11. Higgins, C.F., *ABC transporters: from microorganisms to man.* Annu Rev Cell Biol, 1992. **8**: p. 67-113.
12. ANDERSEN, D.H., *CYSTIC FIBROSIS OF THE PANCREAS AND ITS RELATION TO CELIAC DISEASE: A CLINICAL AND PATHOLOGIC STUDY.* American Journal of Diseases of Children, 1938. **56**(2): p. 344-399.
13. Kerem, B., et al., *Identification of the cystic fibrosis gene: genetic analysis.* Science, 1989. **245**(4922): p. 1073-80.
14. Riordan, J.R., et al., *Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA.* Science, 1989. **245**(4922): p. 1066-73.
15. O'Sullivan, B.P. and S.D. Freedman, *Cystic fibrosis.* Lancet, 2009. **373**(9678): p. 1891-904.
16. Cheng, S.H., et al., *Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis.* Cell, 1990. **63**(4): p. 827-34.
17. Welsh, M.J. and A.E. Smith, *Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis.* Cell, 1993. **73**(7): p. 1251-4.
18. Anderson, M.P., et al., *Demonstration that CFTR is a chloride channel by alteration of its anion selectivity.* Science, 1991. **253**(5016): p. 202-5.
19. Cheng, S.H., et al., *Phosphorylation of the R domain by cAMP-dependent protein kinase regulates the CFTR chloride channel.* Cell, 1991. **66**(5): p. 1027-36.
20. Bear, C.E., et al., *Purification and functional reconstitution of the cystic fibrosis transmembrane conductance regulator (CFTR).* Cell, 1992. **68**(4): p. 809-18.
21. Vergani, P., et al., *CFTR channel opening by ATP-driven tight dimerization of its nucleotide-binding domains.* Nature, 2005. **433**(7028): p. 876-80.

22. Cui, G., et al., *Mutations at arginine 352 alter the pore architecture of CFTR.* J Membr Biol, 2008. **222**(2): p. 91-106.
23. Linsdell, P., A. Evagelidis, and J.W. Hanrahan, *Molecular determinants of anion selectivity in the cystic fibrosis transmembrane conductance regulator chloride channel pore.* Biophys J, 2000. **78**(6): p. 2973-82.
24. Linsdell, P., *Relationship between anion binding and anion permeability revealed by mutagenesis within the cystic fibrosis transmembrane conductance regulator chloride channel pore.* J Physiol, 2001. **531**(Pt 1): p. 51-66.
25. Gong, X. and P. Linsdell, *Mutation-induced blocker permeability and multiion block of the CFTR chloride channel pore.* J Gen Physiol, 2003. **122**(6): p. 673-87.
26. Ge, N., et al., *Direct comparison of the functional roles played by different transmembrane regions in the cystic fibrosis transmembrane conductance regulator chloride channel pore.* J Biol Chem, 2004. **279**(53): p. 55283-9.
27. Aubin, C.N. and P. Linsdell, *Positive charges at the intracellular mouth of the pore regulate anion conduction in the CFTR chloride channel.* J Gen Physiol, 2006. **128**(5): p. 535-45.
28. Zhou, J.J., M. Fatehi, and P. Linsdell, *Direct and indirect effects of mutations at the outer mouth of the cystic fibrosis transmembrane conductance regulator chloride channel pore.* J Membr Biol, 2007. **216**(2-3): p. 129-42.
29. Zhou, J.J., M. Fatehi, and P. Linsdell, *Identification of positive charges situated at the outer mouth of the CFTR chloride channel pore.* Pflugers Arch, 2008. **457**(2): p. 351-60.
30. El Hiani, Y. and P. Linsdell, *Changes in accessibility of cytoplasmic substances to the pore associated with activation of the cystic fibrosis transmembrane conductance regulator chloride channel.* J Biol Chem, 2010. **285**(42): p. 32126-40.
31. Zhou, J.J., et al., *Regulation of conductance by the number of fixed positive charges in the intracellular vestibule of the CFTR chloride channel pore.* J Gen Physiol, 2010. **135**(3): p. 229-45.
32. Qian, F., Y. El Hiani, and P. Linsdell, *Functional arrangement of the 12th transmembrane region in the CFTR chloride channel pore based on functional investigation of a cysteine-less CFTR variant.* Pflugers Arch, 2011. **462**(4): p. 559-71.
33. Wang, W., Y. El Hiani, and P. Linsdell, *Alignment of transmembrane regions in the cystic fibrosis transmembrane conductance regulator chloride channel pore.* J Gen Physiol, 2011. **138**(2): p. 165-78.
34. Wang, W. and P. Linsdell, *Relative movements of transmembrane regions at the outer mouth of the cystic fibrosis transmembrane conductance regulator channel pore during channel gating.* J Biol Chem, 2012. **287**(38): p. 32136-46.
35. El Hiani, Y. and P. Linsdell, *Conformational changes opening and closing the CFTR chloride channel: insights from cysteine scanning mutagenesis.* Biochem Cell Biol, 2014. **92**(6): p. 481-8.
36. El Hiani, Y. and P. Linsdell, *Functional Architecture of the Cytoplasmic Entrance to the Cystic Fibrosis Transmembrane Conductance Regulator Chloride Channel Pore.* J Biol Chem, 2015. **290**(25): p. 15855-65.
37. Linsdell, P., *Anion conductance selectivity mechanism of the CFTR chloride channel.* Biochim Biophys Acta, 2016. **1858**(4): p. 740-7.
38. Linsdell, P., *Architecture and functional properties of the CFTR channel pore.* Cell Mol Life Sci, 2016.

39. Zhang, Z. and J. Chen, *Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator.* Cell, 2016. **167**(6): p. 1586-1597 e9.

40. Van Goor, F., et al., *Effect of ivacaftor on CFTR forms with missense mutations associated with defects in protein processing or function.* J Cyst Fibros, 2014. **13**(1): p. 29-36.

41. Hadida, S., et al., *Discovery of N-(2,4-di-tert-butyl-5-hydroxyphenyl)-4-oxo-1,4-dihydroquinoline-3-carboxamide (VX-770, ivacaftor), a potent and orally bioavailable CFTR potentiator.* J Med Chem, 2014. **57**(23): p. 9776-95.

42. Van der Plas, S.E., et al., *Discovery of N-(3-Carbamoyl-5,5,7,7-tetramethyl-5,7-dihydro-4H-thieno[2,3-c]pyran-2-yl)-lH-pyr azole-5-carboxamide (GLPG1837), a Novel Potentiator Which Can Open Class III Mutant Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Channels to a High Extent.* J Med Chem, 2018. **61**(4): p. 1425-1435.

43. Van Goor, F., et al., *Correction of the F508del-CFTR protein processing defect in vitro by the investigational drug VX-809.* Proc Natl Acad Sci U S A, 2011. **108**(46): p. 18843-8.

44. Taylor-Cousar, J.L., et al., *Tezacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del.* N Engl J Med, 2017. **377**(21): p. 2013-2023.

45. Wainwright, C.E., J.S. Elborn, and B.W. Ramsey, *Lumacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR.* N Engl J Med, 2015. **373**(18): p. 1783-4.

46. Ramsey, B.W., et al., *A CFTR potentiator in patients with cystic fibrosis and the G551D mutation.* N Engl J Med, 2011. **365**(18): p. 1663-72.

47. Keating, D., et al., *VX-445-Tezacaftor-Ivacaftor in Patients with Cystic Fibrosis and One or Two Phe508del Alleles.* N Engl J Med, 2018. **379**(17): p. 1612-1620.

48. Davies, J.C., et al., *VX-659-Tezacaftor-Ivacaftor in Patients with Cystic Fibrosis and One or Two Phe508del Alleles.* N Engl J Med, 2018. **379**(17): p. 1599-1611.

49. Maeda, T., M. Golczak, and A. Maeda, *Retinal photodamage mediated by all-trans-retinal.* Photochem Photobiol, 2012. **88**(6): p. 1309-19.

50. Allikmets, R., et al., *Mutation of the Stargardt disease gene (ABCR) in age-related macular degeneration.* Science, 1997. **277**(5333): p. 1805-7.

51. Cremers, F.P., et al., *Autosomal recessive retinitis pigmentosa and cone-rod dystrophy caused by splice site mutations in the Stargardt's disease gene ABCR.* Hum Mol Genet, 1998. **7**(3): p. 355-62.

52. Bodzioch, M., et al., *The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease.* Nat Genet, 1999. **22**(4): p. 347-51.

53. Shulenin, S., et al., *ABCA3 gene mutations in newborns with fatal surfactant deficiency.* N Engl J Med, 2004. **350**(13): p. 1296-303.

54. Qian, H., et al., *Structure of the Human Lipid Exporter ABCA1.* Cell, 2017. **169**(7): p. 1228-1239 e10.

55. Tsybovsky, Y., et al., *Molecular organization and ATP-induced conformational changes of ABCA4, the photoreceptor-specific ABC transporter.* Structure, 2013. **21**(5): p. 854-60.

56. Liu, F., et al., *Molecular structure of the human CFTR ion channel.* Cell, 2017. **169**(1): p. 85-95. e8.

57. Heymann, J.B. and D.M. Belnap, *Bsoft: image processing and molecular modeling for electron microscopy.* J Struct Biol, 2007. **157**(1): p. 3-18.

58. Scheres, S.H., *RELION: implementation of a Bayesian approach to cryo-EM structure determination.* J Struct Biol, 2012. **180**(3): p. 519-30.

59. Grigorieff, N., *Frealign: An Exploratory Tool for Single-Particle Cryo-EM.* Methods Enzymol, 2016. **579**: p. 191-226.
60. Ward, A.B., et al., *Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide-binding domain.* Proc Natl Acad Sci U S A, 2013. **110**(33): p. 13386-91.
61. Frank, G.A., et al., *Cryo-EM Analysis of the Conformational Landscape of Human P-glycoprotein (ABCB1) During its Catalytic Cycle.* Mol Pharmacol, 2016. **90**(1): p. 35-41.
62. Moeller, A., et al., *Distinct conformational spectrum of homologous multidrug ABC transporters.* Structure, 2015. **23**(3): p. 450-60.
63. Smart, O.S., et al., *HOLE: a program for the analysis of the pore dimensions of ion channel structural models.* J Mol Graph, 1996. **14**(6): p. 354-60, 376.
64. Heijne, G., *The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology.* EMBO J, 1986. **5**(11): p. 3021-7.
65. Nilsson, J., B. Persson, and G. von Heijne, *Comparative analysis of amino acid distributions in integral membrane proteins from 107 genomes.* Proteins, 2005. **60**(4): p. 606-16.
66. Linsdell, P., *Architecture and functional properties of the CFTR channel pore.* Cell Mol Life Sci, 2017. **74**(1): p. 67-83.
67. Smith, S.S., et al., *CFTR: covalent and noncovalent modification suggests a role for fixed charges in anion conduction.* J Gen Physiol, 2001. **118**(4): p. 407-31.
68. Zhang, J. and T.C. Hwang, *The Fifth Transmembrane Segment of Cystic Fibrosis Transmembrane Conductance Regulator Contributes to Its Anion Permeation Pathway.* Biochemistry, 2015. **54**(24): p. 3839-50.
69. Doyle, D.A., et al., *The structure of the potassium channel: molecular basis of K+ conduction and selectivity.* Science, 1998. **280**(5360): p. 69-77.
70. Mihalyi, C., B. Torocsik, and L. Csanady, *Obligate coupling of CFTR pore opening to tight nucleotide-binding domain dimerization.* Elife, 2016. **5**.
71. Csanady, L., et al., *Severed channels probe regulation of gating of cystic fibrosis transmembrane conductance regulator by its cytoplasmic domains.* J Gen Physiol, 2000. **116**(3): p. 477-500.
72. Wang, W., et al., *ATP-independent CFTR channel gating and allosteric modulation by phosphorylation.* Proc Natl Acad Sci U S A, 2010. **107**(8): p. 3888-93.
73. Dawson, R.J. and K.P. Locher, *Structure of the multidrug ABC transporter Sav1866 from Staphylococcus aureus in complex with AMP-PNP.* FEBS Lett, 2007. **581**(5): p. 935-8.
74. Csanady, L., et al., *Preferential phosphorylation of R-domain Serine 768 dampens activation of CFTR channels by PKA.* J Gen Physiol, 2005. **125**(2): p. 171-86.
75. Csanady, L., et al., *Functional roles of nonconserved structural segments in CFTR's NH2-terminal nucleotide binding domain.* J Gen Physiol, 2005. **125**(1): p. 43-55.
76. Csanady, L., P. Vergani, and D.C. Gadsby, *Strict coupling between CFTR's catalytic cycle and gating of its Cl- ion pore revealed by distributions of open channel burst durations.* Proc Natl Acad Sci U S A, 2010. **107**(3): p. 1241-6.
77. Vergani, P., A.C. Nairn, and D.C. Gadsby, *On the mechanism of MgATP-dependent gating of CFTR Cl- channels.* J Gen Physiol, 2003. **121**(1): p. 17-36.
78. Aleksandrov, A.A., et al., *Allosteric modulation balances thermodynamic stability and restores function of DeltaF508 CFTR.* J Mol Biol, 2012. **419**(1-2): p. 41-60.

79. Cai, Z., et al., *Impact of the F508del mutation on ovine CFTR, a Cl- channel with enhanced conductance and ATP-dependent gating.* J Physiol, 2015. **593**(11): p. 2427-46.
80. Clarke, L.L., et al., *Defective epithelial chloride transport in a gene-targeted mouse model of cystic fibrosis.* Science, 1992. **257**(5073): p. 1125-8.
81. Kopelman, H., et al., *Rabbit pancreatic acini express CFTR as a cAMP-activated chloride efflux pathway.* Am J Physiol, 1995. **269**(3 Pt 1): p. C626-31.
82. Marshall, J., et al., *Identification and localization of a dogfish homolog of human cystic fibrosis transmembrane conductance regulator.* J Biol Chem, 1991. **266**(33): p. 22749-54.
83. Price, M.P., et al., *Function of Xenopus cystic fibrosis transmembrane conductance regulator (CFTR) Cl channels and use of human-Xenopus chimeras to investigate the pore properties of CFTR.* J Biol Chem, 1996. **271**(41): p. 25184-91.
84. Rogers, C.S., et al., *Disruption of the CFTR gene produces a model of cystic fibrosis in newborn pigs.* Science, 2008. **321**(5897): p. 1837-41.
85. Singer, T.D., et al., *A divergent CFTR homologue: highly regulated salt transport in the euryhaline teleost F. heteroclitus.* Am J Physiol, 1998. **274**(3 Pt 1): p. C715-23.
86. Sun, X., et al., *Disease phenotype of a ferret CFTR-knockout model of cystic fibrosis.* J Clin Invest, 2010. **120**(9): p. 3149-60.
87. Vasiliou, V., K. Vasiliou, and D.W. Nebert, *Human ATP-binding cassette (ABC) transporter family.* Hum Genomics, 2009. **3**(3): p. 281-90.
88. Inagaki, N., et al., *A family of sulfonylurea receptors determines the pharmacological properties of ATP-sensitive K+ channels.* Neuron, 1996. **16**(5): p. 1011-7.
89. Martin, G.M., et al., *Cryo-EM structure of the ATP-sensitive potassium channel illuminates mechanisms of assembly and gating.* Elife, 2017. **6**.
90. Johnson, Z.L. and J. Chen, *Structural Basis of Substrate Recognition by the Multidrug Resistance Protein MRP1.* Cell, 2017. **168**(6): p. 1075-1085 e9.
91. Dutzler, R., et al., *X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity.* Nature, 2002. **415**(6869): p. 287-94.
92. Goehring, A., et al., *Screening and large-scale expression of membrane proteins in mammalian cells for structural studies.* Nat Protoc, 2014. **9**(11): p. 2574-85.
93. Scharschmidt, B.F., et al., *Validation of a recording spectrophotometric method for measurement of membrane-associated Mg- and NaK-ATPase activity.* J Lab Clin Med, 1979. **93**(5): p. 790-9.
94. Magleby, K.L. and B.S. Pallotta, *Burst kinetics of single calcium-activated potassium channels in cultured rat muscle.* J Physiol, 1983. **344**: p. 605-23.
95. Grant, T. and N. Grigorieff, *Measuring the optimal exposure for single particle cryo-EM using a 2.6 A reconstruction of rotavirus VP6.* Elife, 2015. **4**: p. e06980.
96. Rohou, A. and N. Grigorieff, *CTFFIND4: Fast and accurate defocus estimation from electron micrographs.* J Struct Biol, 2015. **192**(2): p. 216-21.
97. Rubinstein, J.L. and M.A. Brubaker, *Alignment of cryo-EM movies of individual particles by optimization of image translations.* J Struct Biol, 2015. **192**(2): p. 188-95.
98. Webb, B. and A. Sali, *Comparative Protein Structure Modeling Using MODELLER.* Curr Protoc Bioinformatics, 2014. **47**: p. 5 6 1-32.
99. Emsley, P., et al., *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
100. Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis.* Nat Protoc, 2015. **10**(6): p. 845-58.

101. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

102. Brown, A., et al., *Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions.* Acta Crystallogr D Biol Crystallogr, 2015. **71**(Pt 1): p. 136-53.

103. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method.* Acta Crystallogr D Biol Crystallogr, 1997. **53**(Pt 3): p. 240-55.

104. Eyck, L.F.T., *Efficient structure-factor calculation for large molecules by the fast fourier transform.* Acta Cryst. , 1977. **A33**: p. 486-492.

105. Scheres, S.H. and S. Chen, *Prevention of overfitting in cryo-EM structure determination.* Nat Methods, 2012. **9**(9): p. 853-4.

106. Wang, Z., et al., *An atomic model of brome mosaic virus using direct electron detection and real-space optimization.* Nat Commun, 2014. **5**: p. 4808.

107. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.

108. Davis, I.W., et al., *MolProbity: all-atom contacts and structure validation for proteins and nucleic acids.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W375-83.

109. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.

110. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.

111. Ashkenazy, H., et al., *ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules.* Nucleic Acids Res, 2016. **44**(W1): p. W344-50.

112. Zhang, Z., F. Liu, and J. Chen, *Conformational Changes of CFTR upon Phosphorylation and ATP Binding.* Cell, 2017. **170**(3): p. 483-491 e8.

113. Aleksandrov, L., et al., *Differential interactions of nucleotides at the two nucleotide binding domains of the cystic fibrosis transmembrane conductance regulator.* J Biol Chem, 2001. **276**(16): p. 12918-23.

114. Berger, A.L., M. Ikuma, and M.J. Welsh, *Normal gating of CFTR requires ATP binding to both nucleotide-binding domains and hydrolysis at the second nucleotide-binding domain.* Proc Natl Acad Sci U S A, 2005. **102**(2): p. 455-60.

115. Dawson, R.J. and K.P. Locher, *Structure of a bacterial multidrug ABC transporter.* Nature, 2006. **443**(7108): p. 180-5.

116. Mense, M., et al., *In vivo phosphorylation of CFTR promotes formation of a nucleotide-binding domain heterodimer.* EMBO J, 2006. **25**(20): p. 4728-39.

117. Vergani, P., et al., *Control of the CFTR channel's gates.* Biochem Soc Trans, 2005. **33**(Pt 5): p. 1003-7.

118. De la Rosa, M.B. and S.W. Nelson, *An interaction between the Walker A and D-loop motifs is critical to ATP hydrolysis and cooperativity in bacteriophage T4 Rad50.* J Biol Chem, 2011. **286**(29): p. 26258-66.

119. Grossmann, N., et al., *Mechanistic determinants of the directionality and energetics of active export by a heterodimeric ABC transporter.* Nat Commun, 2014. **5**: p. 5419.

120. Bompadre, S.G., et al., *G551D and G1349D, two CF-associated mutations in the signature sequences of CFTR, exhibit distinct gating defects.* J Gen Physiol, 2007. **129**(4): p. 285-98.

121. Cai, Z., A. Taddei, and D.N. Sheppard, *Differential sensitivity of the cystic fibrosis (CF)-associated mutants G551D and G1349D to potentiators of the cystic fibrosis transmembrane conductance regulator (CFTR) Cl- channel.* J Biol Chem, 2006. **281**(4): p. 1970-7.
122. Lin, W.Y., K.Y. Jih, and T.C. Hwang, *A single amino acid substitution in CFTR converts ATP to an inhibitory ligand.* J Gen Physiol, 2014. **144**(4): p. 311-20.
123. Chaves, L.A. and D.C. Gadsby, *Cysteine accessibility probes timing and extent of NBD separation along the dimer interface in gating CFTR channels.* J Gen Physiol, 2015. **145**(4): p. 261-83.
124. Oldham, M.L., et al., *Crystal structure of a catalytic intermediate of the maltose transporter.* Nature, 2007. **450**(7169): p. 515-21.
125. Aleksandrov, L., et al., *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, 2002. **277**(18): p. 15419-25.
126. Basso, C., et al., *Prolonged nonhydrolytic interaction of nucleotide with CFTR's NH2-terminal nucleotide binding domain and its role in channel gating.* J Gen Physiol, 2003. **122**(3): p. 333-48.
127. Szollosi, A., et al., *Mutant cycles at CFTR's non-canonical ATP-binding site support little interface separation during gating.* J Gen Physiol, 2011. **137**(6): p. 549-62.
128. Tsai, M.F., M. Li, and T.C. Hwang, *Stable ATP binding mediated by a partial NBD dimer of the CFTR chloride channel.* J Gen Physiol, 2010. **135**(5): p. 399-414.
129. Al-Shawi, M.K., *Catalytic and transport cycles of ABC exporters.* Essays Biochem, 2011. **50**(1): p. 63-83.
130. Gadsby, D.C., *Ion channels versus ion pumps: the principal difference, in principle.* Nat Rev Mol Cell Biol, 2009. **10**(5): p. 344-52.
131. Jordan, I.K., et al., *Evolutionary and functional divergence between the cystic fibrosis transmembrane conductance regulator and related ATP-binding cassette transporters.* Proc Natl Acad Sci U S A, 2008. **105**(48): p. 18865-70.
132. Muallem, D. and P. Vergani, *Review. ATP hydrolysis-driven gating in cystic fibrosis transmembrane conductance regulator.* Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1514): p. 247-55.
133. Alexander, C., et al., *Cystic fibrosis transmembrane conductance regulator: using differential reactivity toward channel-permeant and channel-impermeant thiol-reactive probes to test a molecular model for the pore.* Biochemistry, 2009. **48**(42): p. 10078-88.
134. Hwang, T.C. and K.L. Kirk, *The CFTR ion channel: gating, regulation, and anion permeation.* Cold Spring Harb Perspect Med, 2013. **3**(1): p. a009498.
135. Locher, K.P., *Mechanistic diversity in ATP-binding cassette (ABC) transporters.* Nat Struct Mol Biol, 2016. **23**(6): p. 487-93.
136. Baker, N.A., et al., *Electrostatics of nanosystems: application to microtubules and the ribosome.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10037-41.
137. Dolinsky, T.J., et al., *PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W522-5.
138. Sorum, B., D. Czege, and L. Csanady, *Timing of CFTR pore opening and structure of its transition state.* Cell, 2015. **163**(3): p. 724-33.

139. Gadsby, D.C., P. Vergani, and L. Csanady, *The ABC protein turned chloride channel whose failure causes cystic fibrosis.* Nature, 2006. **440**(7083): p. 477-83.
140. Haws, C., et al., *CFTR channels in immortalized human airway cells.* Am J Physiol, 1992. **263**(6 Pt 1): p. L692-707.
141. Winter, M.C., et al., *Effect of ATP concentration on CFTR Cl- channels: a kinetic analysis of channel regulation.* Biophys J, 1994. **66**(5): p. 1398-403.
142. Cotten, J.F. and M.J. Welsh, *Cystic fibrosis-associated mutations at arginine 347 alter the pore architecture of CFTR. Evidence for disruption of a salt bridge.* J Biol Chem, 1999. **274**(9): p. 5429-35.
143. Bear, C.E., et al., *Cl- channel activity in Xenopus oocytes expressing the cystic fibrosis gene.* J Biol Chem, 1991. **266**(29): p. 19142-5.
144. Cliff, W.H., R.A. Schoumacher, and R.A. Frizzell, *cAMP-activated Cl channels in CFTR-transfected cystic fibrosis pancreatic epithelial cells.* Am J Physiol, 1992. **262**(5 Pt 1): p. C1154-60.
145. Tilly, B.C., et al., *Cyclic AMP-dependent protein kinase activation of cystic fibrosis transmembrane conductance regulator chloride channels in planar lipid bilayers.* J Biol Chem, 1992. **267**(14): p. 9470-3.
146. Xie, J., et al., *Intracellular loop between transmembrane segments IV and V of cystic fibrosis transmembrane conductance regulator is involved in regulation of chloride channel conductance state.* J Biol Chem, 1995. **270**(47): p. 28084-91.
147. Zhang, Z., F. Liu, and J. Chen, *Molecular structure of the ATP-bound, phosphorylated human CFTR.* Proc Natl Acad Sci U S A, 2018.
148. Baker, J.M., et al., *CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices.* Nat Struct Mol Biol, 2007. **14**(8): p. 738-45.
149. Dulhanty, A.M. and J.R. Riordan, *Phosphorylation by cAMP-dependent protein kinase causes a conformational change in the R domain of the cystic fibrosis transmembrane conductance regulator.* Biochemistry, 1994. **33**(13): p. 4072-9.
150. Bozoky, Z., et al., *Regulatory R region of the CFTR chloride channel is a dynamic integrator of phospho-dependent intra- and intermolecular interactions.* Proc Natl Acad Sci U S A, 2013. **110**(47): p. E4427-36.
151. Naren, A.P., et al., *CFTR chloride channel regulation by an interdomain interaction.* Science, 1999. **286**(5439): p. 544-8.
152. Gao, X. and T.C. Hwang, *Localizing a gate in CFTR.* Proc Natl Acad Sci U S A, 2015. **112**(8): p. 2461-6.
153. Johnson, Z.L. and J. Chen, *ATP Binding Enables Substrate Release from Multidrug Resistance Protein 1.* Cell, 2018. **172**(1-2): p. 81-89 e10.
154. Cui, G., et al., *Three charged amino acids in extracellular loop 1 are involved in maintaining the outer pore architecture of CFTR.* J Gen Physiol, 2014. **144**(2): p. 159-79.
155. Cheung, M. and M.H. Akabas, *Locating the anion-selectivity filter of the cystic fibrosis transmembrane conductance regulator (CFTR) chloride channel.* J Gen Physiol, 1997. **109**(3): p. 289-99.
156. Guinamard, R. and M.H. Akabas, *Arg352 is a major determinant of charge selectivity in the cystic fibrosis transmembrane conductance regulator chloride channel.* Biochemistry, 1999. **38**(17): p. 5528-37.

157. Seibert, F.S., et al., *Disease-associated mutations in cytoplasmic loops 1 and 2 of cystic fibrosis transmembrane conductance regulator impede processing or opening of the channel.* Biochemistry, 1997. **36**(39): p. 11966-74.
158. Seibert, F.S., et al., *Cytoplasmic loop three of cystic fibrosis transmembrane conductance regulator contributes to regulation of chloride channel activity.* J Biol Chem, 1996. **271**(44): p. 27493-9.
159. Wagner, J.A., et al., *Two novel mutations in a cystic fibrosis patient of Chinese origin.* Hum Genet, 1999. **104**(6): p. 511-5.
160. Liu, F., et al., *Structural identification of a hotspot on CFTR for potentiation.* Science, 2019. **364**(6446): p. 1184-1188.
161. Jih, K.Y. and T.C. Hwang, *Vx-770 potentiates CFTR function by promoting decoupling between the gating cycle and ATP hydrolysis cycle.* Proc Natl Acad Sci U S A, 2013. **110**(11): p. 4404-9.
162. Eckford, P.D., et al., *Cystic fibrosis transmembrane conductance regulator (CFTR) potentiator VX-770 (ivacaftor) opens the defective channel gate of mutant CFTR in a phosphorylation-dependent but ATP-independent manner.* J Biol Chem, 2012. **287**(44): p. 36639-49.
163. Van Goor, F., et al., *Rescue of CF airway epithelial cell function in vitro by a CFTR potentiator, VX-770.* Proc Natl Acad Sci U S A, 2009. **106**(44): p. 18825-30.
164. Coleman, R.G., et al., *Ligand pose and orientational sampling in molecular docking.* PLoS One, 2013. **8**(10): p. e75992.
165. Balius, T.E., et al., *Testing inhomogeneous solvation theory in structure-based ligand discovery.* Proc Natl Acad Sci U S A, 2017. **114**(33): p. E6839-E6846.
166. Yu, Y.C., Y. Sohma, and T.C. Hwang, *On the mechanism of gating defects caused by the R117H mutation in cystic fibrosis transmembrane conductance regulator.* J Physiol, 2016. **594**(12): p. 3227-44.
167. Yeh, H.I., et al., *A common mechanism for CFTR potentiators.* J Gen Physiol, 2017. **149**(12): p. 1105-1118.
168. Liu, H. and J.H. Naismith, *An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol.* BMC Biotechnol, 2008. **8**: p. 91.
169. Kimanius, D., et al., *Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2.* Elife, 2016. **5**.
170. Zheng, S.Q., et al., *MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy.* Nat Methods, 2017. **14**(4): p. 331-332.
171. Zhang, K., *Gctf: Real-time CTF determination and correction.* J Struct Biol, 2016. **193**(1): p. 1-12.
172. Gallagher, K. and K. Sharp, *Electrostatic contributions to heat capacity changes of DNA-ligand binding.* Biophys J, 1998. **75**(2): p. 769-76.
173. Sharp, K.A., *Polyelectrolyte electrostatics: Salt dependence, entropic, and enthalpic contributions to free energy in the nonlinear Poisson–Boltzmann model.* Biopolymers, 1995. **36**(2): p. 227-243.
174. Mysinger, M.M. and B.K. Shoichet, *Rapid context-dependent ligand desolvation in molecular docking.* J Chem Inf Model, 2010. **50**(9): p. 1561-73.
175. Word, J.M., et al., *Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation.* J Mol Biol, 1999. **285**(4): p. 1735-47.

176. Hawkins, P.C., et al., *Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database.* J Chem Inf Model, 2010. **50**(4): p. 572-84.
177. Chambers, C.C., et al., *Model for Aqueous Solvation Based on Class IV Atomic Charges and First Solvation Shell Effects.* The Journal of Physical Chemistry, 1996. **100**(40): p. 16385-16398.
178. Li, J., et al., *New Class IV Charge Model for Extracting Accurate Partial Charges from Wave Functions.* The Journal of Physical Chemistry A, 1998. **102**(10): p. 1820-1831.
179. Wang, Z.X., *An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule.* FEBS Lett, 1995. **360**(2): p. 111-4.
180. Kim, Y. and J. Chen, *Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation.* Science, 2018. **359**(6378): p. 915-919.
181. Zhang, Z., F. Liu, and J. Chen, *Molecular structure of the ATP-bound, phosphorylated human CFTR.* Proc Natl Acad Sci U S A, 2018. **115**(50): p. 12757-12762.
182. Illing, M., L.L. Molday, and R.S. Molday, *The 220-kDa rim protein of retinal rod outer segments is a member of the ABC transporter superfamily.* J Biol Chem, 1997. **272**(15): p. 10303-10.
183. Ahn, J., J.T. Wong, and R.S. Molday, *The effect of lipid environment and retinoids on the ATPase activity of ABCR, the photoreceptor ABC transporter responsible for Stargardt macular dystrophy.* J Biol Chem, 2000. **275**(27): p. 20399-405.
184. Quazi, F. and R.S. Molday, *Differential phospholipid substrates and directional transport by ATP-binding cassette proteins ABCA1, ABCA7, and ABCA4 and disease-causing mutants.* J Biol Chem, 2013. **288**(48): p. 34414-26.
185. Nickell, S., et al., *Three-dimensional architecture of murine rod outer segments determined by cryoelectron tomography.* J Cell Biol, 2007. **177**(5): p. 917-25.
186. Alapati, A., et al., *Molecular diagnostic testing by eyeGENE: analysis of patients with hereditary retinal dystrophy phenotypes involving central vision loss.* Invest Ophthalmol Vis Sci, 2014. **55**(9): p. 5510-21.
187. Ozgul, R.K., et al., *Molecular analysis of the ABCA4 gene in Turkish patients with Stargardt disease and retinitis pigmentosa.* Hum Mutat, 2004. **23**(5): p. 523.
188. Lewis, R.A., et al., *Genotype/Phenotype analysis of a photoreceptor-specific ATP-binding cassette transporter gene, ABCR, in Stargardt disease.* Am J Hum Genet, 1999. **64**(2): p. 422-34.
189. Thiadens, A.A., et al., *Clinical course, genetic etiology, and visual outcome in cone and cone-rod dystrophy.* Ophthalmology, 2012. **119**(4): p. 819-26.
190. Downs, K., et al., *Molecular testing for hereditary retinal disease as part of clinical care.* Arch Ophthalmol, 2007. **125**(2): p. 252-8.
191. Zaneveld, J., et al., *Comprehensive analysis of patients with Stargardt macular dystrophy reveals new genotype-phenotype correlations and unexpected diagnostic revisions.* Genet Med, 2015. **17**(4): p. 262-70.
192. Briggs, C.E., et al., *Mutations in ABCR (ABCA4) in patients with Stargardt macular degeneration or cone-rod degeneration.* Invest Ophthalmol Vis Sci, 2001. **42**(10): p. 2229-36.
193. Sciezynska, A., et al., *Next-generation sequencing of ABCA4: High frequency of complex alleles and novel mutations in patients with retinal dystrophies from Central Europe.* Exp Eye Res, 2016. **145**: p. 93-99.

194. Rivera, A., et al., *A comprehensive survey of sequence variation in the ABCA4 (ABCR) gene in Stargardt disease and age-related macular degeneration.* Am J Hum Genet, 2000. **67**(4): p. 800-13.
195. Fujinami, K., et al., *Detailed genetic characteristics of an international large cohort of patients with Stargardt disease: ProgStar study report 8.* Br J Ophthalmol, 2019. **103**(3): p. 390-397.
196. Zernant, J., et al., *Analysis of the ABCA4 gene by next-generation sequencing.* Invest Ophthalmol Vis Sci, 2011. **52**(11): p. 8479-87.
197. Jiang, F., et al., *Screening of ABCA4 Gene in a Chinese Cohort With Stargardt Disease or Cone-Rod Dystrophy With a Report on 85 Novel Mutations.* Invest Ophthalmol Vis Sci, 2016. **57**(1): p. 145-52.
198. Fishman, G.A., et al., *Variation of clinical expression in patients with Stargardt dystrophy and sequence variations in the ABCR gene.* Arch Ophthalmol, 1999. **117**(4): p. 504-10.
199. Zernant, J., et al., *Frequent hypomorphic alleles account for a significant fraction of ABCA4 disease and distinguish it from age-related macular degeneration.* J Med Genet, 2017. **54**(6): p. 404-412.
200. Papaioannou, M., et al., *An analysis of ABCR mutations in British patients with recessive retinal dystrophies.* Invest Ophthalmol Vis Sci, 2000. **41**(1): p. 16-9.
201. Stenirri, S., et al., *Denaturing HPLC profiling of the ABCA4 gene for reliable detection of allelic variations.* Clin Chem, 2004. **50**(8): p. 1336-43.
202. Webster, A.R., et al., *An analysis of allelic variation in the ABCA4 gene.* Invest Ophthalmol Vis Sci, 2001. **42**(6): p. 1179-89.
203. Schulz, H.L., et al., *Mutation Spectrum of the ABCA4 Gene in 335 Stargardt Disease Patients From a Multicenter German Cohort-Impact of Selected Deep Intronic Variants and Common SNPs.* Invest Ophthalmol Vis Sci, 2017. **58**(1): p. 394-403.
204. Jaakson, K., et al., *Genotyping microarray (gene chip) for the ABCR (ABCA4) gene.* Hum Mutat, 2003. **22**(5): p. 395-403.
205. Paloma, E., et al., *Spectrum of ABCA4 (ABCR) gene mutations in Spanish patients with autosomal recessive macular dystrophies.* Hum Mutat, 2001. **17**(6): p. 504-10.
206. Riveiro-Alvarez, R., et al., *Outcome of ABCA4 disease-associated alleles in autosomal recessive retinal dystrophies: retrospective analysis in 420 Spanish families.* Ophthalmology, 2013. **120**(11): p. 2332-7.
207. Stone, E.M., et al., *Clinically Focused Molecular Investigation of 1000 Consecutive Families with Inherited Retinal Disease.* Ophthalmology, 2017. **124**(9): p. 1314-1331.
208. Klevering, B.J., et al., *Microarray-based mutation analysis of the ABCA4 (ABCR) gene in autosomal recessive cone-rod dystrophy and retinitis pigmentosa.* Eur J Hum Genet, 2004. **12**(12): p. 1024-32.
209. Downes, S.M., et al., *Detection rate of pathogenic mutations in ABCA4 using direct sequencing: clinical and research implications.* Arch Ophthalmol, 2012. **130**(11): p. 1486-90.
210. Birch, D.G., et al., *Visual function in patients with cone-rod dystrophy (CRD) associated with mutations in the ABCA4(ABCR) gene.* Exp Eye Res, 2001. **73**(6): p. 877-86.
211. Joo, K., et al., *Genotypic profile and phenotype correlations of ABCA4-associated retinopathy in Koreans.* Mol Vis, 2019. **25**: p. 679-690.
212. Tracewska, A.M., et al., *Genetic Spectrum of ABCA4-Associated Retinal Degeneration in Poland.* Genes (Basel), 2019. **10**(12).

213. Testa, F., et al., *Macular function and morphologic features in juvenile stargardt disease: longitudinal study.* Ophthalmology, 2014. **121**(12): p. 2399-405.
214. Gerber, S., et al., *Complete exon-intron structure of the retina-specific ATP binding transporter gene (ABCR) allows the identification of novel mutations underlying Stargardt disease.* Genomics, 1998. **48**(1): p. 139-42.
215. Pang, C.P. and D.S. Lam, *Differential occurrence of mutations causative of eye diseases in the Chinese population.* Hum Mutat, 2002. **19**(3): p. 189-208.
216. Riveiro-Alvarez, R., et al., *Frequency of ABCA4 mutations in 278 Spanish controls: an insight into the prevalence of autosomal recessive Stargardt disease.* Br J Ophthalmol, 2009. **93**(10): p. 1359-64.
217. Nassisi, M., et al., *Expanding the Mutation Spectrum in ABCA4: Sixty Novel Disease Causing Variants and Their Associated Phenotype in a Large French Stargardt Cohort.* Int J Mol Sci, 2018. **19**(8).
218. Nassisi, M., et al., *Prevalence of ABCA4 Deep-Intronic Variants and Related Phenotype in An Unsolved "One-Hit" Cohort with Stargardt Disease.* Int J Mol Sci, 2019. **20**(20).
219. Birtel, J., et al., *Clinical and genetic characteristics of 251 consecutive patients with macular and cone/cone-rod dystrophy.* Sci Rep, 2018. **8**(1): p. 4824.
220. Fujinami, K., et al., *ABCA4 gene screening by next-generation sequencing in a British cohort.* Invest Ophthalmol Vis Sci, 2013. **54**(10): p. 6662-74.
221. Ducroq, D., et al., *The ABCA4 gene in autosomal recessive cone-rod dystrophies.* Am J Hum Genet, 2002. **71**(6): p. 1480-2.
222. Wang, J., et al., *Dependable and efficient clinical utility of target capture-based deep sequencing in molecular diagnosis of retinitis pigmentosa.* Invest Ophthalmol Vis Sci, 2014. **55**(10): p. 6213-23.
223. Fujinami, K., et al., *Stargardt disease with preserved central vision: identification of a putative novel mutation in ATP-binding cassette transporter gene.* Acta Ophthalmol, 2011. **89**(3): p. e297-8.
224. Kim, M.S., et al., *Genetic Mutation Profiles in Korean Patients with Inherited Retinal Diseases.* J Korean Med Sci, 2019. **34**(21): p. e161.
225. Ernest, P.J., et al., *Outcome of ABCA4 microarray screening in routine clinical practice.* Mol Vis, 2009. **15**: p. 2841-7.
226. Stenirri, S., et al., *Are microarrays useful in the screening of ABCA4 mutations in Italian patients affected by macular degenerations?* Clin Chem Lab Med, 2008. **46**(9): p. 1250-5.
227. Muller, P.L., et al., *Monoallelic ABCA4 Mutations Appear Insufficient to Cause Retinopathy: A Quantitative Autofluorescence Study.* Invest Ophthalmol Vis Sci, 2015. **56**(13): p. 8179-86.
228. Yatsenko, A.N., et al., *Late-onset Stargardt disease is associated with missense mutations that map outside known functional regions of ABCR (ABCA4).* Hum Genet, 2001. **108**(4): p. 346-55.
229. Carss, K.J., et al., *Comprehensive Rare Variant Analysis via Whole-Genome Sequencing to Determine the Molecular Pathology of Inherited Retinal Disease.* Am J Hum Genet, 2017. **100**(1): p. 75-90.
230. Huang, W.C., et al., *Inner and outer retinal changes in retinal degenerations associated with ABCA4 mutations.* Invest Ophthalmol Vis Sci, 2014. **55**(3): p. 1810-22.
231. Fujinami, K., et al., *Clinical and molecular analysis of Stargardt disease with preserved foveal structure and function.* Am J Ophthalmol, 2013. **156**(3): p. 487-501 e1.

232. Roberts, L., et al., *Molecular Diagnosis of Inherited Retinal Diseases in Indigenous African Populations by Whole-Exome Sequencing.* Invest Ophthalmol Vis Sci, 2016. **57**(14): p. 6374-6381.

233. Testa, F., et al., *Correlation between photoreceptor layer integrity and visual function in patients with Stargardt disease: implications for gene therapy.* Invest Ophthalmol Vis Sci, 2012. **53**(8): p. 4409-15.

234. Zernant, J., et al., *Genetic and clinical analysis of ABCA4-associated disease in African American patients.* Hum Mutat, 2014. **35**(10): p. 1187-94.

235. Aguirre-Lamban, J., et al., *Further associations between mutations and polymorphisms in the ABCA4 gene: clinical implication of allelic variants and their role as protector/risk factors.* Invest Ophthalmol Vis Sci, 2011. **52**(9): p. 6206-12.

236. Ogino, K., et al., *[Genotype screening of retinal dystrophies in the Japanese population using a microarray].* Nippon Ganka Gakkai Zasshi, 2013. **117**(1): p. 12-8.

237. Bauwens, M., et al., *An augmented ABCA4 screen targeting noncoding regions reveals a deep intronic founder variant in Belgian Stargardt patients.* Hum Mutat, 2015. **36**(1): p. 39-42.

238. Hu, F.Y., et al., *ABCA4 Gene Screening in a Chinese Cohort With Stargardt Disease: Identification of 37 Novel Variants.* Front Genet, 2019. **10**: p. 773.

239. Fumagalli, A., et al., *Mutational scanning of the ABCR gene with double-gradient denaturing-gradient gel electrophoresis (DG-DGGE) in Italian Stargardt disease patients.* Hum Genet, 2001. **109**(3): p. 326-38.

240. Rosenberg, T., et al., *N965S is a common ABCA4 variant in Stargardt-related retinopathies in the Danish population.* Mol Vis, 2007. **13**: p. 1962-9.

241. Bungert, S., L.L. Molday, and R.S. Molday, *Membrane topology of the ATP binding cassette transporter ABCR and its relationship to ABC1 and related ABCA transporters: identification of N-linked glycosylation sites.* J Biol Chem, 2001. **276**(26): p. 23539-46.

242. Khan, M., et al., *Cost-effective molecular inversion probe-based ABCA4 sequencing reveals deep-intronic variants in Stargardt disease.* Hum Mutat, 2019. **40**(10): p. 1749-1759.

243. Consugar, M.B., et al., *Panel-based genetic diagnostic testing for inherited eye diseases is highly accurate and reproducible, and more sensitive for variant detection, than exome sequencing.* Genet Med, 2015. **17**(4): p. 253-261.

244. Biswas-Fiss, E.E., et al., *Interaction of extracellular domain 2 of the human retina-specific ATP-binding cassette transporter (ABCA4) with all-trans-retinal.* J Biol Chem, 2010. **285**(25): p. 19372-83.

245. Sun, H., R.S. Molday, and J. Nathans, *Retinal stimulates ATP hydrolysis by purified and reconstituted ABCR, the photoreceptor-specific ATP-binding cassette transporter responsible for Stargardt disease.* J Biol Chem, 1999. **274**(12): p. 8269-81.

246. Rozet, J.M., et al., *Spectrum of ABCR gene mutations in autosomal recessive macular dystrophies.* Eur J Hum Genet, 1998. **6**(3): p. 291-5.

247. Chipman, D.M. and B. Shaanan, *The ACT domain family.* Curr Opin Struct Biol, 2001. **11**(6): p. 694-700.

248. Grant, G.A., *The ACT domain: a small molecule binding domain and its role as a common regulatory element.* J Biol Chem, 2006. **281**(45): p. 33825-9.

249. Kadaba, N.S., et al., *The high-affinity E. coli methionine ABC transporter: structure and allosteric regulation.* Science, 2008. **321**(5886): p. 250-3.

250. Biswas-Fiss, E.E., et al., *Retinoid binding properties of nucleotide binding domain 1 of the Stargardt disease-associated ATP binding cassette (ABC) transporter, ABCA4.* J Biol Chem, 2012. **287**(53): p. 44097-107.
251. Wangtiraumnuay, N., et al., *Novel ABCA4 mutation leads to loss of a conserved C-terminal motif: implications for predicting pathogenicity based on genetic testing.* Eur J Ophthalmol, 2018. **28**(1): p. 123-126.
252. Tsybovsky, Y., R.S. Molday, and K. Palczewski, *The ATP-binding cassette transporter ABCA4: structural and functional properties and role in retinal disease.* Adv Exp Med Biol, 2010. **703**: p. 105-25.
253. Stenirri, S., et al., *De novo deletion removes a conserved motif in the C-terminus of ABCA4 and results in cone-rod dystrophy.* Clin Chem Lab Med, 2006. **44**(5): p. 533-7.
254. Guan, J.Z., et al., *A case of Tangier disease with a novel mutation in the C-terminal region of ATP-binding cassette transporter A1.* Am J Med Genet A, 2004. **130A**(4): p. 398-401.
255. Akil, N. and A.J. Fischer, *Surfactant deficiency syndrome in an infant with a C-terminal frame shift in ABCA3: A case report.* Pediatr Pulmonol, 2018. **53**(5): p. E12-E14.
256. Brousseau, M.E., et al., *Novel mutations in the gene encoding ATP-binding cassette 1 in four tangier disease kindreds.* J Lipid Res, 2000. **41**(3): p. 433-41.
257. Zhong, M., L.L. Molday, and R.S. Molday, *Role of the C terminus of the photoreceptor ABCA4 transporter in protein folding, function, and retinal degenerative diseases.* J Biol Chem, 2009. **284**(6): p. 3640-9.
258. Patel, M.J., S.B. Biswas, and E.E. Biswas-Fiss, *Functional significance of the conserved C-Terminal VFVNFA motif in the retina-specific ABC transporter, ABCA4, and its role in inherited visual disease.* Biochem Biophys Res Commun, 2019. **519**(1): p. 46-52.
259. Garces, F., et al., *Correlating the Expression and Functional Activity of ABCA4 Disease Variants With the Phenotype of Patients With Stargardt Disease.* Invest Ophthalmol Vis Sci, 2018. **59**(6): p. 2305-2315.
260. Sodi, A., et al., *Different patterns of fundus autofluorescence related to ABCA4 gene mutations in Stargardt disease.* Ophthalmic Surg Lasers Imaging, 2010. **41**(1): p. 48-53.
261. Zhao, L., et al., *Next-generation sequencing-based molecular diagnosis of 82 retinitis pigmentosa probands from Northern Ireland.* Hum Genet, 2015. **134**(2): p. 217-30.
262. Wang, L., et al., *Application of Whole Exome and Targeted Panel Sequencing in the Clinical Molecular Diagnosis of 319 Chinese Families with Inherited Retinal Dystrophy and Comparison Study.* Genes (Basel), 2018. **9**(7).
263. Fujinami, K., et al., *Clinical and molecular characteristics of childhood-onset Stargardt disease.* Ophthalmology, 2015. **122**(2): p. 326-34.
264. Chacon-Camacho, O.F., et al., *ABCA4 mutational spectrum in Mexican patients with Stargardt disease: Identification of 12 novel mutations and evidence of a founder effect for the common p.A1773V mutation.* Exp Eye Res, 2013. **109**: p. 77-82.
265. Biernbaum, M.S. and M.D. Bownds, *Light-induced changes in GTP and ATP in frog rod photoreceptors. Comparison with recovery of dark current and light sensitivity during dark adaptation.* J Gen Physiol, 1985. **85**(1): p. 107-21.
266. Rust, S., et al., *Tangier disease is caused by mutations in the gene encoding ATP-binding cassette transporter 1.* Nat Genet, 1999. **22**(4): p. 352-5.
267. Fitzgerald, M.L., et al., *ABCA1 and amphipathic apolipoproteins form high-affinity molecular complexes required for cholesterol efflux.* J Lipid Res, 2004. **45**(2): p. 287-94.

268. Kirchhofer, A., et al., *Modulation of protein properties in living cells using nanobodies.* Nat Struct Mol Biol, 2010. **17**(1): p. 133-8.
269. Punjani, A., et al., *cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination.* Nat Methods, 2017. **14**(3): p. 290-296.
270. Liebschner, D., et al., *Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix.* Acta Crystallogr D Struct Biol, 2019. **75**(Pt 10): p. 861-877.
271. Ostrowski, L.E., et al., *Expression of CFTR from a ciliated cell-specific promoter is ineffective at correcting nasal potential difference in CF mice.* Gene Ther, 2007. **14**(20): p. 1492-501.
272. Montoro, D.T., et al., *A revised airway epithelial hierarchy includes CFTR-expressing ionocytes.* Nature, 2018. **560**(7718): p. 319-324.
273. Plasschaert, L.W., et al., *A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte.* Nature, 2018. **560**(7718): p. 377-381.
274. Liu, J.J., et al., *CasX enzymes comprise a distinct family of RNA-guided genome editors.* Nature, 2019. **566**(7743): p. 218-223.