Biophysics and Physicobiology

Regular Article

Molecular dynamics simulation study on the structural instability of the most common cystic fibrosis-associated mutant ΔF508-CFTR

Mitsuhiko Odera[1], Tadaomi Furuta[1], Yoshiro Sohma[2] and Minoru Sakurai[1]

[1]Center for Biological Resources and Informatics, Tokyo Institute of Technology, Yokohama, Kanagawa 226-8501, Japan
[2]Department of Pharmaceutical Sciences, Graduate School of Pharmacy and Center for Medical Science, International University of Health and Welfare, Ohtawara, Tochigi 324-8501, Japan

Cystic fibrosis transmembrane conductance regulator (CFTR) is an anion channel that belongs to the ATP binding cassette protein superfamily. Deletion of phenylalanine at position 508 (ΔF508) is the most common CF-associated mutation and is present in nearly 90% of CF patients. Currently, atomistic level studies are insufficient for understanding the mechanism by which the deletion of a single amino acid causes greatly reduced folding as well as trafficking and gating defects. To clarify this mechanism, we first constructed an atomic model of the inward-facing ΔF508-CFTR and performed all-atom molecular dynamics (MD) simulations of the protein in a membrane environment. All of the computational methodologies used are based on those developed in our previous study for wild-type CFTR. Two important findings were obtained. First, consistent with several previous computational results, the deletion of F508 causes a disruption of a hydrophobic cluster located at the interface between the nucleotide binding domain 1 (NBD1) and intracellular loop 4 (ICL4). This exerts unfavorable influences on the correlated motion between ICLs and transmembrane domains (TMDs), likely resulting in gating defects. Second, the F508 deletion affected the NBD1–NBD2 interface via allosteric communication originating from the correlated motion between NBDs and ICLs. As a result, several unusual inter-residue interactions are caused at the NBD1–NBD2 interface. In other words, their correct dimerization is impaired. This study provided insight into the atomic-level details of structural and dynamics changes caused by the ΔF508 mutation and thus provides good insight for drug design.

Key words: cystic fibrosis transmembrane conductance regulator, ion channel, gating defect, hydrophobic cluster

Cystic fibrosis transmembrane conductance regulator (CFTR) is an anion channel that belongs to the ATP binding cassette (ABC) protein superfamily, and its mutations are known to cause CF, which is commonly found in Caucasians [1,2]. The deletion of phenylalanine at position 508 (ΔF508) is the most common CF-associated mutation and is present in nearly 90% of CF patients [2]. The F508 deletion results in greatly reduced folding, as well as both trafficking and

Corresponding author: Minoru Sakurai, Center for Biological Resources and Informatics, Tokyo Institute of Technology, 4259-B-62, Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan.
e-mail: masakurai@bio.titech.ac.jp

◀ *Significance* ▶

CFTR is an anion channel that belongs to the ABC transporter superfamily. Deletion of F508 (ΔF508) is known to cause serious gating defects. Our MD simulations revealed two posible origins of the functional defects of this mutant. First, the deletion of F508 causes a disruption of a hydrophobic cluster located at the interface between the nucleotide binding domain 1 (NBD1) and intracellular loop 4. Second, the effect of the F508 deletion propagates into the NBD1–NBD2 interface, resulting in the impairment of their correct dimerization. These findings are useful for drug design.

gating defects [3–5]. In addition, when ΔF508-CFTR reaches the cell membrane, it exhibits a shorter half-life (~4 h) than wild-type (WT) (>48 h) channels either because of more rapid ubiquitin-dependent degradation or, in some cases, defective recycling [5-8].

CFTR consists of two nucleotide-binding domains (NBD1 and NBD2) and two transmembrane domains (TMD1 and TMD2) (Supplementary Fig. S1). Each TMD has six transmembrane helices (TMs). NBDs and TMDs make contact via four intracellular loops (ICL1–ICL4) [9]. In addition, there is a regulatory domain (RD) unique to CFTR between NBD1 and TMD2, which regulates CFTR channel activity through phosphorylation by protein kinase A. The F508 residue is located in NBD1 close to its interface with ICL4 [9]. The ΔF508 mutation likely destabilizes NBD1 and its interaction with other domains [10,11], resulting in trafficking and gating defects [12]. However, early biochemical studies of isolated NBD1 with the ΔF508 mutation showed that the deletion of F508 did not affect the backbone structure, whereas its refolding was impaired [13]. Indeed, both crystallographic data [14–16] and mass spectral measurements [17] suggested minimal structural differences between WT- and ΔF508-NBD1. On the other hand, MD simulations [18–21] and a NMR study [22] indicated that dynamic differences between WT- and ΔF508-NBD1 spread all over the protein.

Although the atomic level mechanism by which deletion of F508 causes observed functional defects of CFTR currently remains unclear, at least three structural and thermodynamic factors have been implicated. First, the ΔF508 mutation causes destabilization of NBD1 itself. Indeed, recent isothermal calorimetric studies suggested that the mutation results in both kinetic and thermodynamic folding defects in the isolated NBD1 domain [23,24]. Second, deletion of ΔF508 generates a void at the NBD1/ICL4 interface and thus impairs proper assembly of the NBD–TMD interfaces, which is needed to acquire the final native fold of the whole protein [16,25–28]. Simultaneous introduction of NBD1-thermostabilizing and NBD1/ICL4 interface-correcting mutations synergistically improves ΔF508 processing to nearly WT levels [10,29], indicating that both defects contribute to ΔF508 misprocessing. Third, deletion of F508 likely causes a significant structural perturbation to the NBD1–NBD2 dimer interface, although it is distant from the mutation site [30,31]. Jih *et al.* [30] employed a functional assay to assess the stability of the NBD dimer, and the results suggested that the ΔF508 mutation significantly destabilized the NBD dimer. In addition, a recent NMR study on a ΔF508 NBD1 indicated that the dimerization ability of NBD1 is reduced by the mutation through a conformational change of the Q-loop, which leads to unfavorable homodimerization [22]. This result may be extended to the case of heterodimerization with NBD2, which is needed for channel gating.

NBD heterodimerization is thought to be essential for CFTR folding and processing [26], as well as channel gating [32–35]. However, the effect of ΔF508 mutation on NBD heterodimerization is not well studied due to the difficulty in directly measuring dimerization and producing isolated NBD2. Molecular dynamics (MD) simulation is a unique method that enables us to observe the above dimerization process at the atomic level. Our recent MD simulations for the full WT-CFTR model, except for the R domain, successfully reproduced the dimerization of NBDs induced by the binding of ATP molecules as well as the resultant structural change of TMD, although channel opening was not yet observed due to the short simulation duration time (~100 ns) [36]. In addition, we recently applied similar MD simulations to another heterodimeric ABC transporter, TM287/288, and indicated that the binding of two ATP molecules to NBDs induced the formation of the consensus ATP-binding pocket or NBD dimerization [37]. In these MD simulations, no artificial bias potential was used to enhance protein structural changes [36,37]. The successful reproduction of the heterodimerization of the NBDs implies that both the modelling structure and MD simulation protocol used are accurate enough to describe the dynamics of full-size heterodimeric ABC transporters. We expect that the application of such a well-established methodology to ΔF508-CFTR will elucidate the molecular mechanism by which ΔF508-CFTR causes various functional defects, as mentioned above. In this study, we obtain a structural model of ΔF508-CFTR by modifying the WT-CFTR model used in the previous study [36] and conduct multiple MD simulations of it. We reveal that deletion of F508 causes the disruption of a hydrophobic cluster located at the NBD1/ICL4 interface and that this defect influences the structure of the NBD1–NBD2 interface, consequently leading to incomplete heterodimerization.

Materials and Methods

Homology Models of WT- and ΔF508-CFTR

CFTR has a loop-like structure called the "regulatory insertion (RI)" (residues 403–438) in NBD1 located between strands β1 and β2 of the ABC-specific β (or RecA) subdomain (Supplementary Fig. S1). NBD1, NBD2, TMD1 and TMD2 were constructed through 3D homology modeling, the details of which were described in our previous paper [36] and are not repeated here. RI of WT-CFTR was constructed as follows. We used the PDB structure 1XMI as a template for modeling of NBD1. In this structure, the atomic coordinates of part of RI are determined, although those of A412–L428 are missing. We performed modeling only for this missing part using Discovery Studio 4.0 (Accelrys), which generated energetically reasonable structures for the input primary sequence (A412–L428) without any template. As a result, this part was determined to be disordered. The resultant structure was optimized by energy minimization calculation before MD simulation. The homology modeling was not applied for RD (residues 646–843) because the split

channels with no RD were highly active without phosphorylation [38]. The N-terminal (residues 1–56) and C-terminal (residues 1428–1480) regions were also not considered in this study because no appropriate templates have been found. We compared our WT-CFTR model with other homology models [39,40] and a recently reported experimental structure for human CFTR [41] in Supplementary Figures S2–S4.

The structure of ΔF508-CFTR was obtained as follows. First, we conducted a MD simulation for a system in which the homology modelling structure of WT-CFTR was embedded in the lipid bilayer and obtained the equilibrium structure of the WT protein under the heavy-atom restraints (see below). Next, we carried out homology modelling in which this WT structure was used as a template. Then, the modelling calculation was performed using Discovery Studio 4.0.

The protonation states of all the charged residues were determined on the assumption of pH=7. However, His1402 in the H-loop was doubly protonated.

MD simulations

MD simulations were performed according to the protocol developed in the previous study [36]. First, the homology model of WT-CFTR or ΔF508-CFTR was inserted into a 1-palmitoyl-2-oleylphosphatidylcholine (POPC) lipid bilayer. The bilayer structure was prepared by quadrupling a bilayer consisting of 128 POPC molecules downloaded from Tieleman's Web site (available for download from http://moose.bio.ucalgary.ca/). Consequently, 484 POPC molecules were present in the bilayer. Subsequently, this system (protein and membrane) was solvated by water molecules, where Na^+ and Cl^- ions were inserted randomly into the solvent to preserve the electric neutrality of the whole system. The resultant system has 59,991 water molecules as well as 133 Na^+ and 150 Cl^- ions, a solution corresponding to 100 mM. The total number of atoms in the systems for the WT-CFTR and ΔF508 mutant was 217,538 and 217,521, respectively.

Each system, including WT-CFTR or ΔF508-CFTR, was relaxed by the steepest descent minimization until the maximum force on any atom was less than 100 kJ/(mol Å). The MD simulation was carried out as follows: the energy-minimized system was first equilibrated for 40 ns with harmonic restraints on all of the heavy atoms of proteins with a strength of 10 kJ/(mol Å²); then, production runs were carried out for a total of 100 ns without the restraints. We performed one and three 100-ns-long MD simulations for WT-CFTR or ΔF508-CFTR, respectively. All replicates were initiated with different sets of random velocities. All simulations were conducted under the NPT conditions with a Nosé−Hoover thermostat [42,43] and Parrinello−Rahman barostat [44]. The temperature of the system was kept at 310 K, with separate coupling of solutes (e.g., protein, nucleotides, and lipids) and solvents. The pressure was kept at 1 bar. During the MD simulations, the covalent bonds were constrained by the LINCS algorithm [45,46] and SPC water molecules [47] were constrained with the SETTLE algo-

rithm [48]. The van der Waals interactions were used with a cutoff distance of 12 Å. The electrostatic interactions were calculated using the particle-mesh Ewald method with a real-space cutoff of 12 Å. The simulation time step was set to 2 fs.

All of the MD simulations were performed using the GROMACS v4.0.7. package [49]. The GROMOS-87 force field [50–52] was used for the protein, and the SPC model was used for the water molecules. The force field for ATP with a net charge of −4e was obtained by modifying the original GROMOS force field. We removed the hydrogen atom from the protonated γ-phosphate and redistributed the partial charges evenly over the phosphate oxygen atoms. As for the POPC molecules, we used the topology developed by Berger *et al.* [53] and later modified by Tieleman and Berendsen.[54] Analysis of the MD trajectories was carried out using GROMACS tools. 3D graphics were produced using Discovery Studio 4.0 (Accelrys).

Results

Structural Stability of WT- and ΔF508-CFTR

To check the structural stabilities of the WT and ΔF508 systems during their MD simulations, we calculated the root mean square deviations (RMSDs) of the Cα atoms from their initial modeling structures. The time courses of the RMSDs obtained from a production run of the WT and from three production runs of ΔF508 are shown in Figure 1A and 1B–1D, respectively. In both the WT and ΔF508 systems, the RMSD values tend to increase with time. In more detail, their changes became slow beyond 70 ns, although the values for TMD2 (C, ΔF508 (2nd)) and NBD1 (D, ΔF508 (3rd)) still continue to increase significantly. Additionally, it is of note that in both systems, the structural change of NBD1 is remarkably larger than that of NBD2. The larger change of NBD1 is likely ascribed to its template-free modeling part of RI (see below). Since the RMSD values of the whole protein ("All" in Fig. 1A–1D) continue to increase slightly, the protein is not thought to reach the true equilibrium state. In spite of this, in both systems, the two NBDs dimerized in a head-to-tail fashion, where the RecA subdomain of one NBD faced the helical subdomain of the other NBD (Fig. 2A and 2B). Therefore, the 100 ns simulations performed here seem to satisfy the minimal conditions that would be required for our main purposes, i.e., here we investigate how the NBD dimerization is affected by the deletion of F508, and consequently how the NBD-TMD and NBD-NBD interactions responsible for the gating function are disturbed.

Figure 3 shows close-up views of the vicinity of RI, which correspond to the snapshots at 100 ns in the four MD runs. As already described in our previous paper [36], a part of RI (residues 424−436) in NBD1 of WT-CFTR contacts an upstream portion (residues 1329−1332) of the X-loop in NBD2, which is called the "safety catch" [39] and likely
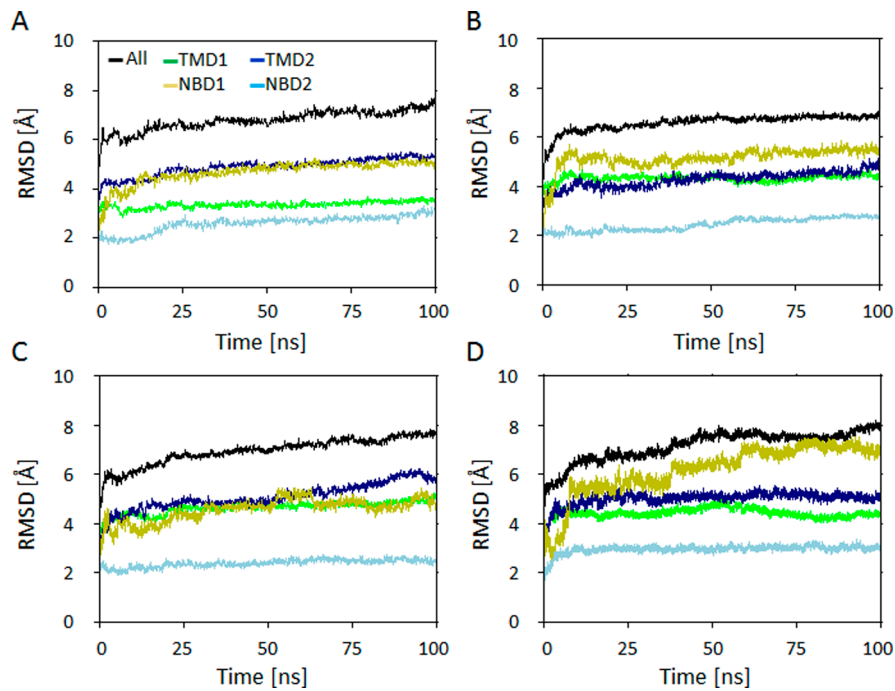
Figure 1 Time course of RMSD. A: WT, B: ΔF508 (1st), C: ΔF508 (2nd), D: ΔF508 (3rd). In each figure, the data for the whole protein structure (black), TMD1 (green), TMD2 (violet), NBD1 (gold) and NBD2 (cyan) are separately illustrated.
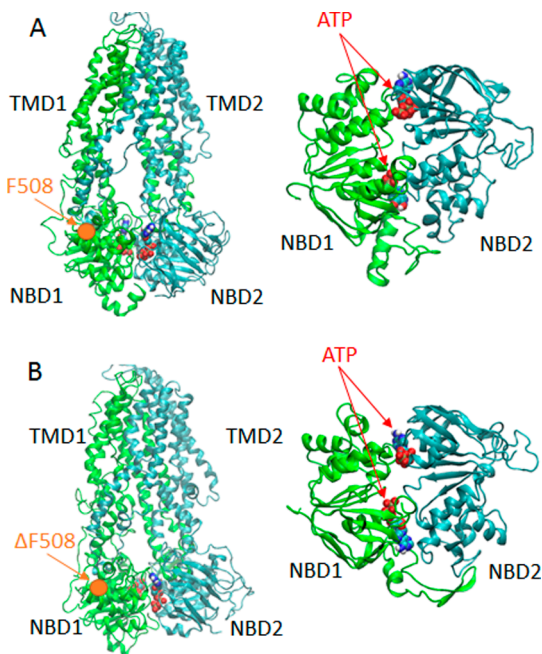

Figure 2 Snapshot structures at 100 ns. A: WT system, B: ΔF508 (1st) system. In both figures, the left is the front view of the whole protein and the right is the view of the two NBDs from the top. The first and the second halves of the protein are depicted in green and cyan ribbons, and ATPs are in CPK models. The position of F508 (WT) and the deleted position of F508 (ΔF508) are indicated by orange circles.

limits the NBD sliding motion. In contrast, RI of ΔF508-CFTR contacts NBD2 in a different fashion for each run and its conformation tends to be more disordered. In the 2nd and 3rd runs, RI did not catch NBD2, and in the 3rd run, it invades into the NBD1-NBD2 interface and reaches ICL2 to hamper their ideal dimerization. Thus, it is likely that in ΔF508-CFTR the role of RI in catching NBD2 does not work well.

NBD dimerization in the WT- and ΔF508-CFTR systems

To examine the structure of the NBD1–NBD2 hetero-dimer in more detail, we measured the center of mass (COM) distance between the Walker A motif of one NBD and the signature motif of the other NBD (Fig. 4B–E). These motifs should be paired with each other to form an ATP-binding pocket (ABP), as illustrated schematically in Figure 4A. Since two ATP molecules bind to the NBDs, two ABPs (red and green lines in Fig. 4A) are formed, which are hereafter denoted as ABP1 and ABP2, respectively. In the initial state of the WT system, the distance of the Walker A-signature was 20 and 18 Å at the ABP1 and ABP2 sites, respectively. Beyond 10 ns, these distances converged to approximately 10 and 11 Å, respectively, which implies that the two NBDs approached each other and dimerized in an almost ideal fashion. On the other hand, in the ΔF508 system, the Walker A-signature distances in all three different MD runs reached values larger than those in the WT system. In both the 1st and 3rd runs, the values at the ABP1 and ABP2 sites were approximately 14 Å and 12 Å, respectively, and the 100 ns conformations of the NBD dimers in these two runs were
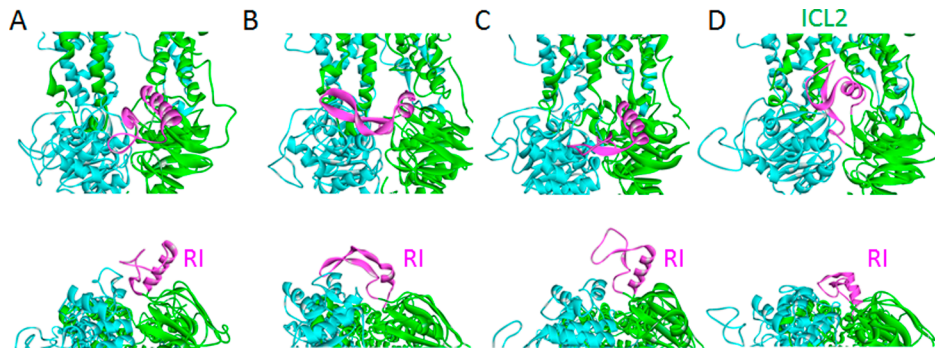
Figure 3   Close-up views of the vicinity of RI. A: WT, B–D: 1st, 2nd, and 3rd runs of ΔF508. Each figure corresponds to the snapshot at 100 ns in each run. RI in each figure are depicted in magenta ribbon.
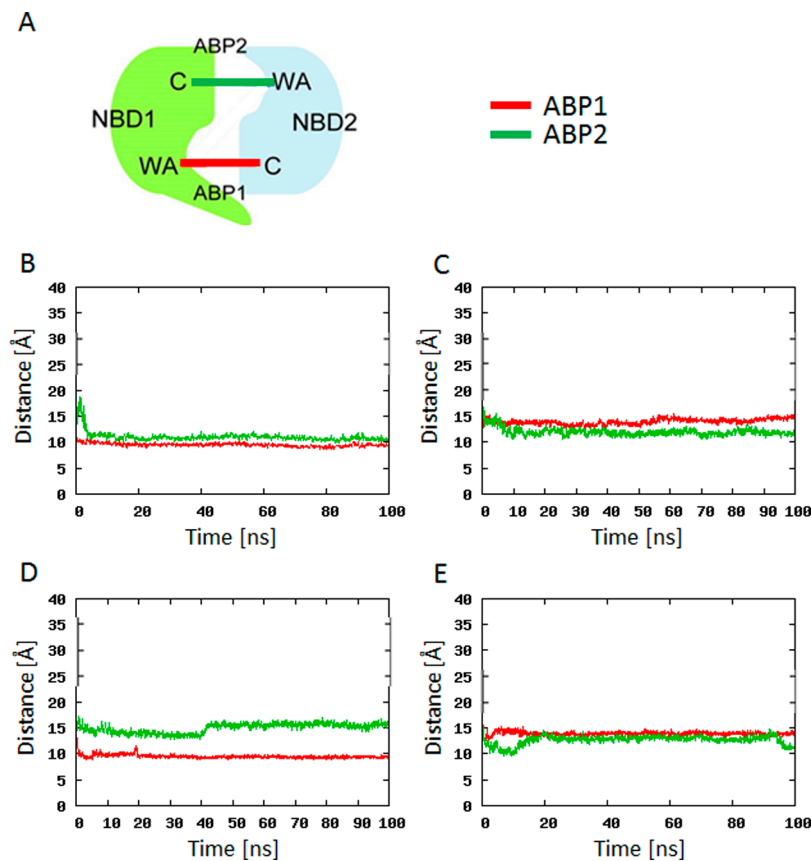

Figure 4   Time course of the COM distance between the Walker A and Signature motifs. A: Schematic representation of NBD dimer. B: WT, C: ΔF508(1st), D: ΔF508(2nd), E: ΔF508(3rd). The lines for ABP1 and ABP2 are plotted in red and green colors, respectively.

very similar to each other with a RMSD of 3.48 Å (Supplementary Fig. S5). However, in the 2nd run, the corresponding distances at the ABP1 and ABP2 sites were 10 Å and 16 Å, respectively, indicating less perfect dimerization in this case. Thus, it is reasonable to choose the 1st or 3rd run as a representative of the ΔF508 system. Hereinafter, we focus mainly on the 1st run, and its results will be compared with those for the WT system.

Motional coupling between the NBDs and TMDs through the ICL

To examine the global protein motion in the WT and ΔF508 systems, we carried out principal component analysis (PCA) for the trajectory data from 90 to 100 ns. The first principal component is illustrated in Figure 5, where the movement of each Cα atom is indicated by a thin arrow drawn from each Cα position. In the WT system, the NBD1
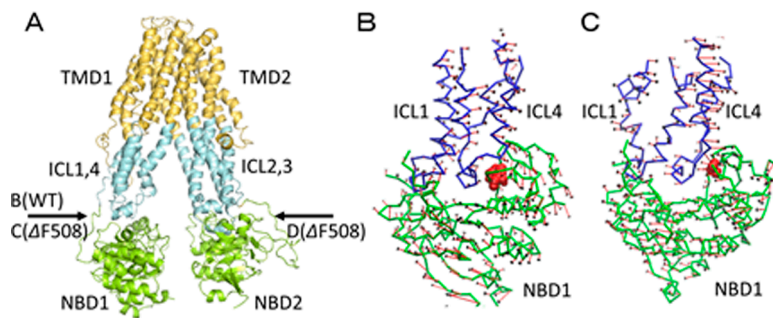
Figure 5 Results of PCA. A: Visual directions of the whole structure are indicated by black arrows, B: Movements of the NBD1 and ICLs in the WT system, C: Movements of the NBD1 and ICLs in the ΔF508 system. NBDs, ICLs, and F508 are colored by green, blue and red, respectively. In the case of ΔF508, instead of F508, G509 is indicated by red.
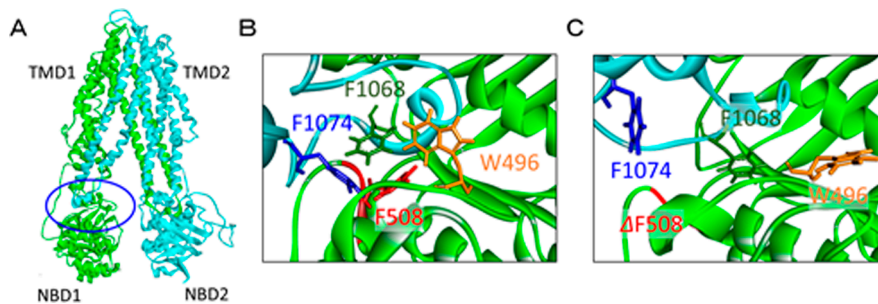
Figure 6 Hydrophobic interaction network (hydrophobic cluster) at the NBD1/ICL4 interface. A: The position of the NBD1/ICL4 interface, B and C: close-up views of the NBD1/ICL4 interfaces in the WT and ΔF508 systems, respectively. Red: F508 and its neighboring region, orange: W496, green: F1068, blue: F1074.

rotates in a clock-wise direction. The ICLs (ICL1 and ICL4) move in a manner synchronized with the motion of the NBDs: the ICLs move to the right. This rotational motion is exactly like a "nodding-like" motion upon conformational transition from an inward-facing to an outward-facing conformation, which we discussed in our previous study [55]. Such a characteristic correlated motion was also found on the NBD2 side (data not shown). By contrast, on the NBD1 side of the ΔF508 system, the motional correlation between the NBD1 and ICLs disappeared, so that these regions moved in an anti-parallel fashion and the motion of ICLs was randomly disturbed. On the NBD2 side, however, the correlated motion was maintained between the NBD2 and ICLs, and was similar to the WT system (Supplementary Fig. S6). Therefore, the decorrelation on the NBD1 side might arise from a structural defect, e.g., deletion of F508, in the NBD1. It has been hypothesized that the ICLs act as transmitters, propagating the motion generated in the NBD engine to the TMDs. These PCA results suggest that such a function of the ICL is seriously affected by the defect of the partner NBD.

## Structural changes at the NBD-ICL interface caused by the deletion of F508

F508 is located at the interface between the helical sub-domain in NBD and ICL4 (Fig. 6A), a region responsible for

the communication between NBDs and TMDs. Here, we analyze how the deletion of F508 influences the local structure around the original F508 position.

In Figure 6B and 6C, a close-up view of the helical sub-domain side of the interface between NBD1 and ICL4 is shown. In the WT, aromatic amino acid residues (F508, W496, F1068 and F1074) form a hydrophobic cluster (Figs. 6B and 7A), where Figure 7 shows the distances between F508 and these residues. Among these four aromatic residues, F508, W496 and F1068 already formed a cluster in the starting homology model, while F1074 participated in a cluster at the initial stage of the MD simulation. The inter-residue distance between any two amino acids forming this cluster was kept almost constant throughout the MD simulation, which indicates the good stability of the cluster. Thus, this cluster likely plays an important role in the mechanical connection between NBD1 and ICL4. On the other hand, such a hydrophobic cluster was disrupted by deletion of F508 (Figs. 6C and 7B), and a similar disruption was also observed in the 3rd run of the ΔF508 system (data not shown), in which the conformation of the NBD dimer was similar to that in the 1st run as described above. The disruption of the hydrophobic cluster might lead to an unfavorable inter-action between NBD1 and ICL4, resulting in the decorrelated motion between them (Fig. 5C).
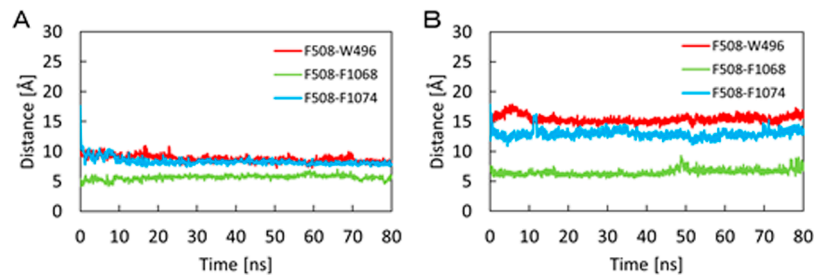
Figure 7   COM distances between two of the aromatic residues forming the hydrophobic cluster. A: WT, B: ΔF508. In ΔF508, the center of mass between I507 and G509 was defined as the hypothetical position of F508.
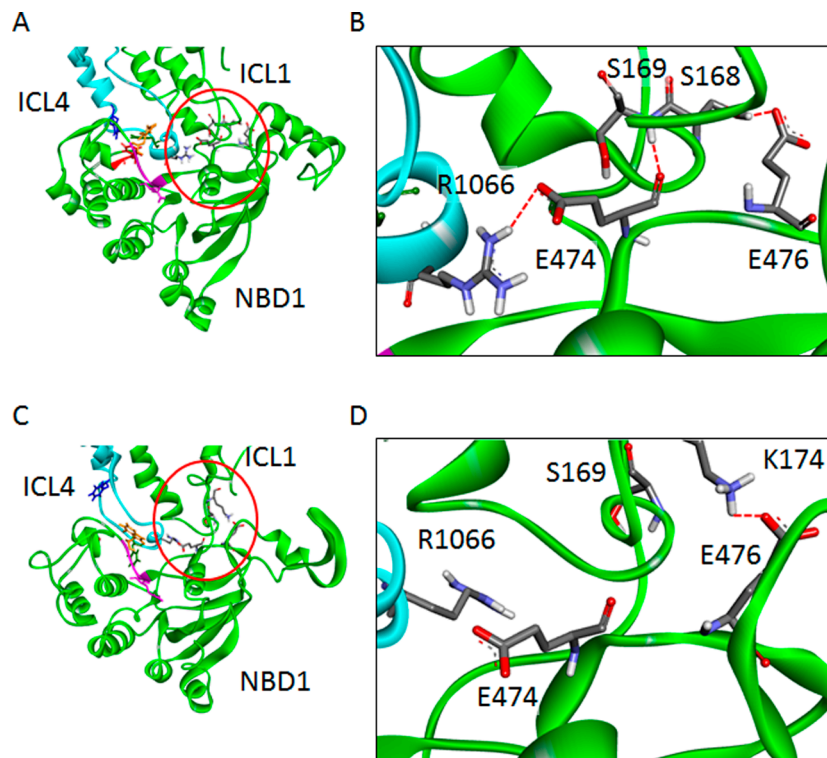

Figure 8   Inter-residue interaction network among ICL1 and NBD1 in the WT and ΔF508 systems. A: The interface region of NBD1/ICL1 in the WT system. Red circle: RecA side of NBD1/ICL1. B: close-up view of the red circle region of A. Five residues (S168, S169, E474, E476, R1066): stick model, dotted red lines: hydrogen bonds. C: The interface region of NBD1/ICL1 in the ΔF508 system. RecA side of NBD1/ICL1. D: close-up view of the red circle region of B. Five residues (S169, K174, E474, E476, R1066): stick model, dotted red line: hydrogen bond.

To understand the interaction between the NBDs and ICLs in more detail, we analyzed the interaction between the RecA subdomains of NBD1 and the ICL1. Hereinafter, we introduce a symbol @xxx to show where a given amino acid residue is located: for example, F508@NBD1 means that F508 is located in NBD1. Figure 8A and 8C show the cross section of the NBD1/ICL1 interface in the WT and mutant systems, respectively. In the WT, on the RecA side of NBD1, E474@NBD1, located in the downstream region of the Walker A motif, formed a hydrogen bond network including R1066@ICL4 and S169@ICL1 (Fig. 8B). Furthermore, it was suggested by the model that S168@ICL1 interacts with

E476@NBD1. In ΔF508, on the RecA subdomain side, the interaction between E476 and K174 of ICL1 was observed (Fig. 8D), whereas the hydrogen bonds present in the WT system disappeared. In summary, in the mutant, in addition to the disruption of the hydrophobic cluster (Fig. 6C), the mode and number of the inter-residue interactions were also significantly changed on the RecA sides of the NBD1/ICL1 interface.

Structural changes at the NBD1-NBD2 interface induced by the ΔF508 mutation

The major parts of the structural changes induced by dele-
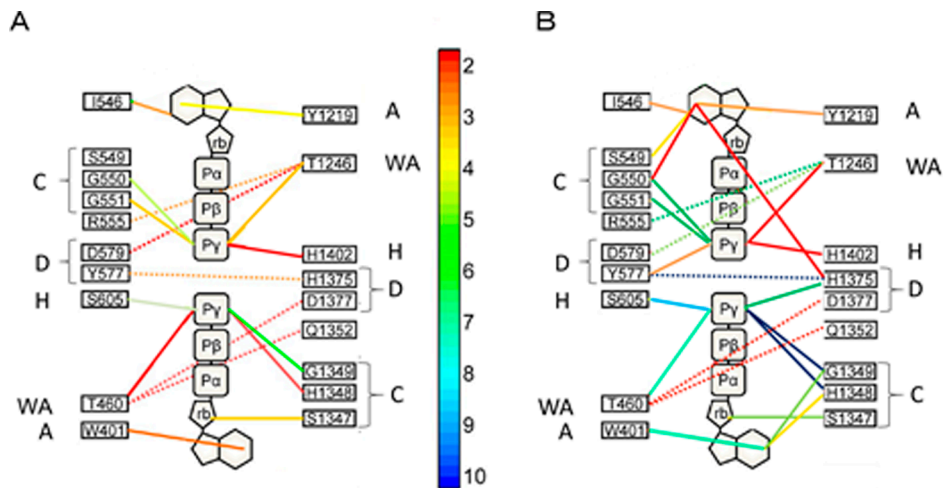
Figure 9   The interactions among ATP and its surrounding amino acid residues. A: WT, B: ΔF508. Each solid line connects ATP with its interacting amino acid residue. Each dotted line represents an amino acid pair that directly contributes to the interaction between the two NBDs. The interaction distances are shown in the color gradation from 2.0 (red) to 10.0 (blue) Å.

tion of F508 occurred at the initial phase (0–5 ns) of the MD simulation. To investigate collective motions in this phase, we performed PCA analysis to reveal which modes of molecular motion caused such a structural change. As a result, the first principal component (PC1) in the WT corresponds to the closure motion of the two NBDs and can be supposed to govern the total motion of the protein because its relative contribution was 39.4% of the total protein motion (Supplementary Fig. S7). This is consistent with the observation that NBD dimerization occurred at 0–5 ns (Fig. 4B). On the other hand, in the ΔF508 system, the main motion was assigned to a tilting movement of NBD1 against the ICLs (ICL1 and ICL4), consequently leading to a distortion of the relative position between NBD1 and ICL4.

It is of great interest to clarify whether the mutation-induced structural change of the NBD1/ICL4 interface influences the NBD1–NBD2 interface. For this purpose, we examined the interaction between the ATP molecule accommodated in the ABP and its surrounding amino acid residues in both the WT and ΔF508 systems. The result is shown in Figure 9, where Walker A and signature motifs are abbreviated as WA and C, respectively. In the WT system, ATP molecules are sandwiched between the Walker A motif of one NBD as well as the signature motif and D-loop of the other NBD in both ABPs. There were many hydrogen bonds that can be found between the γ-phosphate moieties of the ATP molecules and the neighboring residues of the signature motifs (shown in Fig. 9 left). In particular, hydrogen bond formation between an ATP and G550 (and G551) and between the other ATP and H1348 are thought to make important contributions to the stabilization of the NBD dimer. In both ABPs, T460 (ABP1) and T1246 (ABP2) of the Walker A interact with D1377 (ABP1) and D579 (ABP2) of the D loops, respectively, and simultaneously interact

with the amino acid residues (Q1352 and R555, respectively) located in the downstream regions of the signature motifs. Namely, the NBD dimer is stabilized not only by the "glue" effect of ATP (hydrogen bond formation between ATP and its surrounding residues) but also by the inter-residue interactions between the two NBDs [36]. On the other hand, in the ΔF508 system, the NBD1–NBD2 interaction network was considerably modified. In particular, the distances between the γ-phosphate of an ATP and G550 (G551) and between that of the other ATP and H1348 (G1349) are considerably elongated, so that these interaction pairs were no longer regarded to be hydrogen bonds. Instead, these residues in the signature motifs interact with the adenine ring of the nearby ATP. In addition, the inter-residue interactions between the Walker A motif and D loop (dotted line in Fig. 9) were modified. Consequently, it can safely be said that the NBD1–NBD2 dimer in the ΔF508 system is far from the ideal form found in the WT.

Discussion

The regulatory domain (RD) with multiple consensus sequences for PKA-dependent phosphorylation initially regulates the CFTR channel activity [1]. Once the RD is phosphorylated, the CFTR channel is opened by ATP binding to NBDs and subsequent NBD dimerization. Deletion of F508 impairs channel opening, which implies that this mutation perturbs NBD heterodimerization and the NBD-TMD mechanical coupling. However, the former is not immediately obvious because ΔF508 caused limited alterations to the NBD1 structure [13–17], and additionally, the mutation is not located near the dimer interface. In addition, MD simulations have indicated that dynamic differences between WT- and ΔF508-NBD1 spread all over the protein [18–21]. Thus,

deletion of F508 should affect multiple steps of the reaction cycle of CFTR. Elucidation of the atomic level mechanism of dysfunction caused by the deletion of F508 is important not only for basic protein science but also for medical treatment. It has not been well studied due to the difficulty of directly monitoring NBD heterodimerization. Recently, for WT-CFTR, Rahman *et al.* successfully investigated the inward- to outward-facing conformational transition induced by NBD heterodimerization using targeted MD simulation, but they have not yet applied this method to ΔF508-CFTR. In this study, the atomic-level structure and dynamics of ΔF508-CFTR were investigated using the MD simulation protocol that successfully reproduced the NBD heterodimerization of CFTR [36] and TM287/288 [37] in our previous studies.

The clustering of several aromatic amino acid residues around F508 in CFTR has been found first by cysteine cross-linking experimental study [11] as well as by a homology modeling study based on Sav1866 [56], and recently by cryo-EM studies of zebrafish and human CFTRs [9,41] In Supplementary Figure S8, we provided a comparison between the aromatic acid residues cluster found in this study with that in the cryo-EM study on human CFTR [41]. In addition, a recent computer modeling study has already predicted that the hydrophobic cluster is disrupted by the deletion of F508, resulting in the formation of a cavity at the NBD1/ICL4 interface [57]. Therefore, the present results for the hydrophobic cluster are consistent with those of the previous studies.

According to the present study, in addition to the helical subdomain side, the RecA side at the NBD1/ICL1 interface is also influenced by the deletion of F508 (Fig. 8). In general, ICLs act as transmitters connecting NBDs and TMDs, and thus, it is likely that the irregular motion of ICLs causes significant modulations on TMD motion, resulting in opening and gating defects. This model may be supported by some experiments that investigated the effect of chemical modifications at the NBD-ICL interface on the protein activity. For example, cross-linking between NBD1 and ICL3 (or ICL4) or between NBD1 and ICL3 inhibits channel activity [58,59]. These observations suggest that TMD-NBD cross-talk is disrupted by the mutation at the NBD-ICL interface.

Next, we discuss how defects at the NBD1/ICL interface influence dimerization of NBD1 and NBD2. Jih *et al.* performed a functional assay to assess the stability of the NBD dimer [30], and found that the pyrophosphate (PPi)-induced locked-open time of the ΔF508 system was dramatically shortened, suggesting that the ΔF508 mutation destabilizes the NBD dimer states. According to the present MD simulation results, the inter-residue interaction network around the bound ATP is broken in the mutant (Fig. 9): ATP is not surrounded by the canonical sandwich structure formed by the Walker A and signature motifs in the ΔF508 system. This would result in a reduction of the ATP hydrolysis activity

and in turn exert a harmful influence on the channel gating in the TMDs.

It is of great interest to elucidate the mechanism by which the deletion of F508 might impair NBD heterodimerization despite the fact that F508 is relatively distant from the dimer interface. A recent NMR study on two NBD1 dimerization showed that residues in the Q-loop segment of the WT NBD1 exist in equilibrium between two states [22]. From the chemical shift data, it was inferred that the deletion of F508, which interacts with W496 in the Q-loop, pushes the conformation of the Q-loop toward the state that is less compatible with dimerization. According to our MD simulation results, W496 is a member of the hydrophobic cluster around F508 (Fig. 6B) and its side chain orientation significantly changed as a result of the deletion of F508 (Fig. 6C). It was found that the orientation of Q493 and the position of the Q-loop is changed with the mutation (Supplementary Fig. S9). Thus, our MD data seem to partly support the above mechanism based on the NMR data. In addition, as described below, we propose other mechanisms responsible for the impairment of NBD heterodimerization in the ΔF508 system.

The change of the dynamic behavior of RI seems to be a factor hampering the ideal NBD heterodimerization. As shown in Figure 3B–3D, RI in the ΔF508 system does not catch NBD2 firmly, while RI in the WT system catches the upstream portion of the X-loop in NBD2, known as "safety catch" [39]. The observation that the dynamics of RI is dramatically changed with the deletion of F508 can be accounted for by considering the correlated motion within NBD1. A recent study on the correlated motion of NBD1 has indicated that both F508 and RI move in a highly correlated fashion and thus the deletion of F508 affects the dynamics of RI, resulting in an increase of its thermal fluctuation [20]. If RI really hamper NBD heterodimerization in the ΔF508 system, the deletion of RI should restore the function of the protein. Indeed, Aleksandrov *et al.* have indicated that RI deletion restores ATP occlusion by NBD1 of the ΔF508 system and has a strong thermostabilizing influence on the channel [60]. This implies that restriction of RI to a particular conformational state may ameliorate the impact of the ΔF508 mutation.

So far, several MD simulation studies on the effect of F508 deletion have been reported for NBD1 alone [18–21] or an open-state CFTR model [61], and indicated that the dynamics change induced by the mutation tends to spread almost all over the protein. Such a tendency was found in the present MD simulations. One of the characteristic features of the present study is that the MD simulations were applied to the initial molecular process starting from an inward-facing conformation (channel closed state) to the ATP-occluded state (NBD-heterodimerized state). As a result, we could propose a new hypothesis for the role of RI as discussed above. However, it is important to note that the initial protein structures based on the present homology-modeling might include somewhat unfavorable distortions.

Unfortunately, it is difficult to discuss to what extent these potential defects affected the final results. Probably, this problem will be solved not until the present results are compared with those from new simulation studies for recently reported cryo-EM structures of CFTR which is beyond the scope of this study.

Conclusion

In summary, our MD simulations provided insights into the atomistic level mechanisms causing the structural defects, including the allosteric communication between the mutation site and NBD1–NBD2 interface. It should be stressed that these results were obtained by an all-atom MD simulation in which no artificial bias potential to accelerate the conformation sampling was applied. Therefore, we believe that the molecular events observed in this study reliably reproduce the essential aspects of the real system and thus provide good insight for drug design.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Acknowledgement

Author Contributions

M. O. performed all the simulations and analyses. T. F., Y. S. and M. S. directed the entire project and co-wrote the manuscript.

References

[1] Riordan, J. R., Rommens, J. M., Kerem, B.-S., Alon, N., Rozmahel, R., Grzelczak, Z., *et al.* Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* **245**, 1066–1073 (1989).

[2] Welsh, M. J. & Smith, A. E. Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. *Cell* **73**, 1251–1254 (1993).

[3] Dalemans, W., Barbry, P., Champigny, G., Jallat, S., Dott, K., Dreyer, D., *et al.* Altered chloride ion channel kinetics associated with the ΔF508 cystic fibrosis mutation. *Nature* **354**, 526–528 (1991).

[4] Denning, G. M., Anderson, M. P., Amara, J. F., Marshall, J., Smith, A. E. & Welsh, M. J. Processing of mutant cystic fibrosis transmembrane conductance regulator is temperature-sensitive. *Nature* **358**, 761–764 (1992).

[5] Lukacs, G. L., Chang, X. B., Bear, C., Kartner, N., Mohamed, A., Riordan, J. R., *et al.* The ΔF508 mutation decreases the stability of cystic fibrosis transmembrane conductance regulator in the plasma membrane. Determination of functional half-lives on transfected cells. *J. Biol. Chem.* **268**, 21592–21598 (1993).

[6] Chang, X. B., Mengos, A., Hou, Y. X., Cui, L., Jensen, T. J., Aleksandrov, A., *et al.* Role of N-linked oligosaccharides in the biosynthetic processing of the cystic fibrosis membrane conductance regulator. *J. Cell Sci.* **121**, 2814–2823 (2008).

[7] Okiyoneda, T., Barriere, H., Bagdany, M., Rabeh, W. M., Du, K., Hohfeld, J., *et al.* Peripheral protein quality control removes unfolded CFTR from the plasma membrane. *Science* **329**, 805–810 (2010).

[8] Swiatecka-Urban, A., Brown, A., Moreau-Marquis, S., Renuka, J., Coutermarsh, B., Barnaby, R., *et al.* The short apical membrane half-life of rescued ΔF508-cystic fibrosis transmembrane conductance regulator (CFTR) results from accelerated endocytosis of ΔF508-CFTR in polarized human airway epithelial cells. *J. Biol. Chem.* **280**, 36762–36772 (2005).

[9] Zhang, Z. & Chen, J. Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator. *Cell* **167**, 1586–1597.e9 (2016).

[10] Rabeh, W. M., Bossard, F., Xu, H., Okiyoneda, T., Bagdany, M., Mulvihill, C. M., *et al.* Correction of both NBD1 energetics and domain interface is required to restore ΔF508 CFTR folding and function. *Cell* **148**, 150–163 (2012).

[11] Serohijos, A. W., Hegedus, T., Aleksandrov, A. A., He, L., Cui, L., Dokholyan, N. V., *et al.* Phenylalanine-508 mediates a cytoplasmic-membrane domain contact in the CFTR 3D structure crucial to assembly and channel function. *Proc. Natl. Acad. Sci. USA* **105**, 3256–3261 (2008).

[12] Cui, L., Aleksandrov, L., Hou, Y. X., Gentzsch, M., Chen, J. H., Riordan, J. R., *et al.* The role of cystic fibrosis transmembrane conductance regulator phenylalanine 508 side chain in ion channel gating. *J. Physiol.* **572**, 347–358 (2006).

[13] Qu, B. H. & Thomas, P. J. Alteration of the cystic fibrosis transmembrane conductance regulator folding pathway. *J. Biol. Chem.* **271**, 7261–7264 (1996).

[14] Atwell, S., Brouillette, C. G., Conners, K., Emtage, S., Gheyi, T., Guggino, W. B., *et al.* Structures of a minimal human CFTR first nucleotide-binding domain as a monomer, head-to-tail homodimer, and pathogenic mutant. *Protein Eng. Des. Sci.* **23**, 375–384 (2010).

[15] Lewis, H. A., Zhao, X., Wang, C., Sauder, J. M., Rooney, I., Noland, B. W., *et al.* Impact of the ΔF508 mutation in first nucleotide-binding domain of human cystic fibrosis transmembrane conductance regulator on domain folding and structure. *J. Biol. Chem.* **280**, 1346–1353 (2005).

[16] Thibodeau, P. H., Brautigam, C. A., Machius, M. & Thomas, P. J. Side chain and backbone contributions of Phe508 to CFTR folding. *Nat. Struct. Mol. Biol.* **12**, 10–16 (2005).

[17] Lewis, H. A., Wang, C., Zhao, X., Hamuro, Y., Conners, K., Kearins, M. C., *et al.* Structure and dynamics of NBD1 from CFTR characterized using crystallography and hydrogen/deuterium exchange mass spectrometry. *J. Mol. Biol.* **396**, 406–430 (2010).

[18] Bisignano, P. & Moran, O. Molecular dynamics analysis of the wild type and dF508 mutant structures of the human CFTR-nucleotide binding domain 1. *Biochimie* **92**, 51–57 (2010).

[19] Wieczorek, G. & Zielenkiewicz, P. ΔF508 mutation increases conformational flexibility of CFTR protein. *J. Cyst. Fibros.* **7**, 295–300 (2008).

[20] Proctor, E. A., Kota, P., Aleksandrov, A. A., He, L., Riordan, J. R. & Dokholyan, N. V. Rational Coupled Dynamics Network Manipulation Rescues Disease-Relevant Mutant Cystic Fibrosis Transmembrane Conductance Regulator. *Chem. Sci.* **6**, 1237–1246 (2015).

[21] Zhenin, M., Noy, E. & Senderowitz, H. REMD Simulations Reveal the Dynamic Profile and Mechanism of Action of

Deleterious, Rescuing, and Stabilizing Perturbations to NBD1 from CFTR. *J. Chem. Inf. Model.* **55**, 2349–2364 (2015).

[22] Chong, P. A., Farber, P. J., Vernon, R. M., Hudson, R. P., Mittermaier, A. K. & Forman-Kay, J. D. Deletion of Phenyl-alanine 508 in the First Nucleotide-binding Domain of the Cystic Fibrosis Transmembrane Conductance Regulator Increases Conformational Exchange and Inhibits Dimeriza-tion. *J. Biol. Chem.* **290**, 22862–22878 (2015).

[23] Protasevich, I., Yang, Z., Wang, C., Atwell, S., Zhao, X., Emtage, S., *et al.* Thermal unfolding studies show the disease causing F508del mutation in CFTR thermodynamically desta-bilizes nucleotide-binding domain 1. *Protein Sci.* **19**, 1917–1931 (2010).

[24] Wang, W., Wu, J., Bernard, K., Li, G., Wang, G., Bevensee, M. O., *et al.* ATP-independent CFTR channel gating and allosteric modulation by phosphorylation. *Proc. Natl. Acad. Sci. USA* **107**, 3888–3893 (2010).

[25] Cui, L., Aleksandrov, L., Chang, X. B., Hou, Y. X., He, L., Hegedus, T., *et al.* Domain interdependence in the biosyn-thetic assembly of CFTR. *J. Mol. Biol.* **365**, 981–994 (2007).

[26] Du, K. & Lukacs, G. L. Cooperative assembly and misfolding of CFTR domains in vivo. *Mol. Biol. Cell* **20**, 1903–1915 (2009).

[27] Du, K., Sharma, M. & Lukacs, G. L. The ΔF508 cystic fibrosis mutation impairs domain-domain interactions and arrests post-translational folding of CFTR. *Nat. Struct. Mol. Biol.* **12**, 17–25 (2005).

[28] Rosser, M. F., Grove, D. E., Chen, L. & Cyr, D. M. Assembly and misassembly of cystic fibrosis transmembrane conduc-tance regulator: folding defects caused by deletion of F508 occur before and after the calnexin-dependent association of membrane spanning domain (MSD) 1 and MSD2. *Mol. Biol. Cell* **19**, 4570–4579 (2008).

[29] Mendoza, J. L., Schmidt, A., Li, Q., Nuvaga, E., Barrett, T., Bridges, R. J., *et al.* Requirements for efficient correction of ΔF508 CFTR revealed by analyses of evolved sequences. *Cell* **148**, 164–174 (2012).

[30] Jih, K. Y., Li, M., Hwang, T. C. & Bompadre, S. G. The most common cystic fibrosis-associated mutation destabilizes the dimeric state of the nucleotide-binding domains of CFTR. *J. Physiol.* **589**, 2719–2731 (2011).

[31] Miki, H., Zhou, Z., Li, M., Hwang, T. C. & Bompadre, S. G. Potentiation of disease-associated cystic fibrosis transmem-brane conductance regulator mutants by hydrolyzable ATP analogs. *J. Biol. Chem.* **285**, 19967–19975 (2010).

[32] Aleksandrov, A. A., Cui, L. & Riordan, J. R. Relationship between nucleotide binding and ion channel gating in cystic fibrosis transmembrane conductance regulator. *J. Physiol.* **587**, 2875–2886 (2009).

[33] Aleksandrov, L., Aleksandrov, A. A., Chang, X. B. & Riordan, J. R. The First Nucleotide Binding Domain of Cystic Fibrosis Transmembrane Conductance Regulator Is a Site of Stable Nucleotide Interaction, whereas the Second Is a Site of Rapid Turnover. *J. Biol. Chem.* **277**, 15419–15425 (2002).

[34] Berger, A. L., Ikuma, M. & Welsh, M. J. Normal gating of CFTR requires ATP binding to both nucleotide-binding domains and hydrolysis at the second nucleotide-binding domain. *Proc. Natl. Acad. Sci. USA* **102**, 455–460 (2005).

[35] Vergani, P., Lockless, S. W., Nairn, A. C. & Gadsby, D. C. CFTR channel opening by ATP-driven tight dimerization of its nucleotide-binding domains. *Nature* **433**, 876–880 (2005).

[36] Furukawa-Hagiya, T., Furuta, T., Chiba, S., Sohma, Y. & Sakurai, M. The power stroke driven by ATP binding in CFTR as studied by molecular dynamics simulations. *J. Phys. Chem. B* **117**, 83–93 (2013).

[37] Furuta, T., Sato, Y. & Sakurai, M. Structural Dynamics of the Heterodimeric ABC Transporter TM287/288 Induced by ATP and Substrate Binding. *Biochemistry* **55**, 6730–6738 (2016).

[38] Csanady, L., Chan, K. W., Seto-Young, D., Kopsco, D. C., Nairn, A. C. & Gadsby, D. C. Severed channels probe regula-tion of gating of cystic fibrosis transmembrane conductance regulator by its cytoplasmic domains. *J. Gen. Physiol.* **116**, 477–500 (2000).

[39] Mornon, J. P., Lehn, P. & Callebaut, I. Molecular models of the open and closed states of the whole human CFTR protein. *Cell. Mol. Life Sci.* **66**, 3469–3486 (2009).

[40] Rahman, K. S., Cui, G., Harvey, S. C. & McCarty, N. A. Modeling the conformational changes underlying channel opening in CFTR. *PLoS ONE* **8**, e74574 (2013).

[41] Liu, F., Zhang, Z., Csanády, L., Gadsby, D. C. & Chen, J. Molecular Structure of the Human CFTR Ion Channel. *Cell* **169**, 85–95 (2017).

[42] Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).

[43] Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A Gen. Phys.* **31**, 1695–1697 (1985).

[44] Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).

[45] Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simula-tions. *J. Comput. Chem.* **18**, 1463–1472 (1997).

[46] Hess-Coelho, T. A. A Redundant Parallel Spherical Mecha-nism for Robotic Wrist Applications. *J. Mech. Des.* **129**, 891–895 (2006).

[47] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F. & Hermans, J. Interaction models for water in relation to pro-tein hydration. in *Intermolecular Forces* (Pullman, B., ed.) pp. 331–342 (D. Reidel Publishing Co., Dordrecht, The Nether-lands, 1981).

[48] Miyamoto, S. & Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**, 952–962 (1992).

[49] van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. & Berendsen, H. J. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).

[50] Mark, A. E., Vanhelden, S. P., Smith, P. E., Janssen, L. H. M. & van Gunsteren, W. F. Convergence Properties of Free Energy Calculations: α-Cyclodextrin Complexes as a Case Study. *J. Am. Chem. Soc.* **116**, 6293–6302 (1994).

[51] van Buuren, A. R., Marrink, S. J. & Berendsen, H. J. C. A molecular dynamics study of the decane/water interface. *J. Phys. Chem.* **97**, 9206–9212 (1993).

[52] van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krüger, P. K., Mark, A. E., *et al. Biomolecular Simu-lation: The GROMOS96 Manual and User Guide*, (Vdf Hochschulverlag AG an der ETH Zurich, Zurich, Switzerland, 1996).

[53] Berger, O., Edholm, O. & Jahnig, F. Molecular dynamics sim-ulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* **72**, 2002–2013 (1997).

[54] Tieleman, D. P. & Berendsen, H. J. A molecular dynamics study of the pores formed by Escherichia coli OmpF porin in a fully hydrated palmitoyloleoylphosphatidylcholine bilayer. *Biophys. J.* **74**, 2786–2801 (1998).

[55] Furuta, T., Yamaguchi, T., Kato, H. & Sakurai, M. Analysis of the Structural and Functional Roles of Coupling Helices in the ATP-Binding Cassette Transporter MsbA through Enzyme Assays and Molecular Dynamics Simulations. *Biochemistry* **53**, 4261–4272 (2014).

[56] Mornon, J. P., Lehn, P. & Callebaut, I. Atomic model of

human cystic fibrosis transmembrane conductance regulator: membrane-spanning domains and coupling interfaces. *Cell. Mol. Life Sci.* **65**, 2594–2612 (2008).

[57] Kalid, O., Mense, M., Fischman, S., Shitrit, A., Bihler, H., Ben-Zeev, E., *et al.* Small molecule correctors of F508del-CFTR discovered by structure-based virtual screening. *J. Comput. Aided Mol. Des.* **24**, 971–991 (2010).

[58] He, L., Aleksandrov, A. A., Serohijos, A. W., Hegedus, T., Aleksandrov, L. A., Cui, L., *et al.* Multiple membrane-cytoplasmic domain contacts in the cystic fibrosis transmembrane conductance regulator (CFTR) mediate regulation of channel gating. *J. Biol. Chem.* **283**, 26383–26390 (2008).

[59] He, L., Aleksandrov, L. A., Cui, L., Jensen, T. J., Nesbitt, K. L. & Riordan, J. R. Restoration of domain folding and interdomain assembly by second-site suppressors of the ΔF508

mutation in CFTR. *FASEB J.* **24**, 3103–3112 (2010).

[60] Aleksandrov, A. A., Kota, P., Aleksandrov, L. A., He, L., Jensen, T., Cui, L., *et al.* Regulatory insertion removal restores maturation, stability and function of ΔF508 CFTR. *J. Mol. Biol.* **401**, 194–210 (2010).

[61] Belmonte, L. & Moran, O. On the interactions between nucleotide binding domains and membrane spanning domains in cystic fibrosis transmembrane regulator: A molecular dynamic study. *Biochimie* **111**, 19–29 (2015).