Investigation of the CFTR chloride channel using molecular dynamics simulations

Theses of the PhD Dissertation

Bianka Vivien Farkas

Supervisors:
Tamás Hegedűs, PhD
Zoltán Gáspári, PhD

Pázmány Péter Catholic University
Faculty of Information Technology and Bionics
Roska Tamás Doctoral School of Sciences and Technology

Budapest
2022

Abstract

Cystic fibrosis (CF) is caused by mutations in the gene of the chloride channel CFTR/ABCC7. Elucidating the mechanism of action of the mutations underlying cystic fibrosis could contribute significantly to the design of targeted therapeutic compounds. For the development of effective therapeutic agents, atomic-level structural and dynamic studies of the CFTR chloride channel are essential.

Mutations causing cystic fibrosis disrupts one or more steps of CFTR biogenesis or impair the function of the ion channel. Consequently, it is important to characterize the chloride pathway at the atomic level and to understand the mechanism of action of chloride conduction. Advances in structure determination methods have enabled atomic-resolution structural characterization of large membrane proteins, such as CFTR. The structures solved by cryo-electron microscopy (cryo-EM) provided an excellent opportunity to study the function of this ion channel, to investigate the structural background of CFTR chloride conduction using *in silico* methods. We described the chloride conducting pathways using molecular dynamics (MD) simulations and analyzed the effect of the atypical broken transmembrane helix 8 (TM8) characteristic to most of the cryo-EM CFTR structures.

The most frequent cystic fibrosis mutation affecting the first nucleotide-binding domain (NBD1) of the CFTR chloride channel is the deletion of phenylalanine in the position 508 (ΔF508/F508del). This mutation prevents the correct folding of NBD1, leading to a premature degradation of the CFTR protein. In addition, ΔF508-CFTR proteins that reach the plasma membrane do not function properly and are characterized by reduced structural stability as well. For this type of mutation, therapeutic efficacy could be enhanced marginally by interfering with the protein folding to promote the formation of correct structures. To explore the mechanism of correction and stabilization of mutant NBD1 structure, our objective was to investigate the unfolding steps and intermediate states of NBD1. Therefore, we investigated the unfolding of NBD1 using single molecule force spectroscopy (AFM) experiments. In addition to the experiments, we characterized the unfolding at the atomic level using MD simulations.

Our simulation results allowed us to describe these processes at the atomic level that are not feasible with acceptable resolution using experiments alone. Furthermore, the methods and results described in my thesis show that the combination and comparison of computational and experimental methods are important for testing the validity of the methods and contribute greatly to understanding of biological processes at the molecular level.

1. Identification of the chloride pathway and investigation of the structural background of CFTR ion channel function

1.1. Motivation and aims of the study

Although several cryo-EM CFTR structures in different conformations have been published recently originating from several organisms, the exact mechanism of function of the ion channel has not yet been described. To overcome the limitations of static structures, the elucidation of protein dynamics using molecular dynamics is a rational possibility. Previously, our research group investigated the structure of zebrafish CFTR determined in the absence of ATP using molecular dynamics (MD) simulations, which revealed separated nucleotide-binding domains (NBDs). It was found that this conformation is not common under physiological conditions, as closure of the NBDs was observed even in short simulations [1]. Since the cryo-EM CFTR structures determined in the active state (ATP-bound and phosphorylated) disclose no pathway suitable for chloride passage, the exact transport pathway of chloride ions was not known. Das *et al.* investigated the CFTR ion channel by generating structural models based on cysteines cross-linking experiments [2]. However, these structural models differ significantly from known ABC protein structures [3]. Others have performed molecular dynamics simulations using homology models based on the ATP-bound zebrafish CFTR structure (PDB ID: 5w81). One study used metadynamics simulations to investigate conformational changes associated with gating [4], and another publication reports on the interactions of CFTR with lipids [5]. However, the detailed description of the chloride pathway was not addressed in these studies.

The zebrafish CFTR structure determined in the active state (ATP-bound and phosphorylated) is referred to as a transient closed state [6]. Thus, we considered MD simulations to cross the energy barrier separating the closed conformation from the open state allowing the identification of the chloride ion pathway. To investigate the structural background of chloride conductance, we used the structure available at the beginning of this study (PDB ID: 5w81 [6]) and we performed molecular dynamics (MD) simulations. In the resulting conformational ensemble, the structures with an open ion channel sufficient for chloride conductance were collected. This method allowed us to study the structural properties of the open chloride channel, to characterize the protein-ion interactions and to explore the energetics of chloride passage through the ion channel.

In most cryo-EM CFTR structures, the TM7 and TM8 transmembrane helices exhibit structural properties different from ABC proteins. In most cryo-EM CFTR structures, the transmembrane helix 8 (TM8) contains a kink in the membrane region (exceptions PDB ID: 6d3r, 6d3s). In addition, the

pseudosymmetry of TM helices is also disrupted by the kink of TM8, and the unusual position of TM8 also displaces transmembrane helix 7 (TM7) from its position typical to homologous ABC proteins. Nevertheless, structural observations suggest that TM8 and its kinked region play an important role in the opening of the ion channel [6]–[9]. However, the unusual properties observed in cryo-EM CFTR structures may cause the closed state of the chloride channel in these structures. To investigate the effect of these unusual features of the TM7 and TM8 observed in the CFTR protein, we generated a structural model with homology modelling, in which both TM8 and TM7 are located similarly to the homologous ABC proteins and the kinked segment of TM8 was straightened. Using the resulting structural model, we performed MD simulations similar to the zebrafish cryo-EM structure.

1.2. New scientific results

In the first part of my thesis, I described our results on the structural background of CFTR chloride conductance. We performed equilibrium molecular dynamics simulations with the ATP-bound and phosphorylated CFTR structure (PDB ID: 5w81) solved in the active state. Using this method, we determined the chloride permeation pathways in the generated conformational ensemble. We characterized the structural properties of the ion channel at the atomic level and investigated the interactions between the protein and the chloride ions. We collected published experimental data on the ion channel function of CFTR and validated our simulation results. We also investigated the effect of TM8 on the ion channel opening. The TM8 is an atypical structural element, kinked in the membrane region that is uncharacteristic of ABC proteins. In addition, the unusual position of TM8 in the structure also displaces TM7 from its position typical of homologous ABC proteins.

Thesis I.1. There are two alternative entry sites in the cytosolic part of the CFTR chloride channel.

We observed two intracellular entry pores (TM4/6 and TM10/12) in the open conformations. Both were surrounded by positively charged amino acids, which may play an important role in the electrostatic attraction of negatively charged chloride ions. In the simulations, chloride ions were frequently present in these regions due to the favorable electrostatic interactions. The determined chloride pathways are also in agreement with published experimental results.

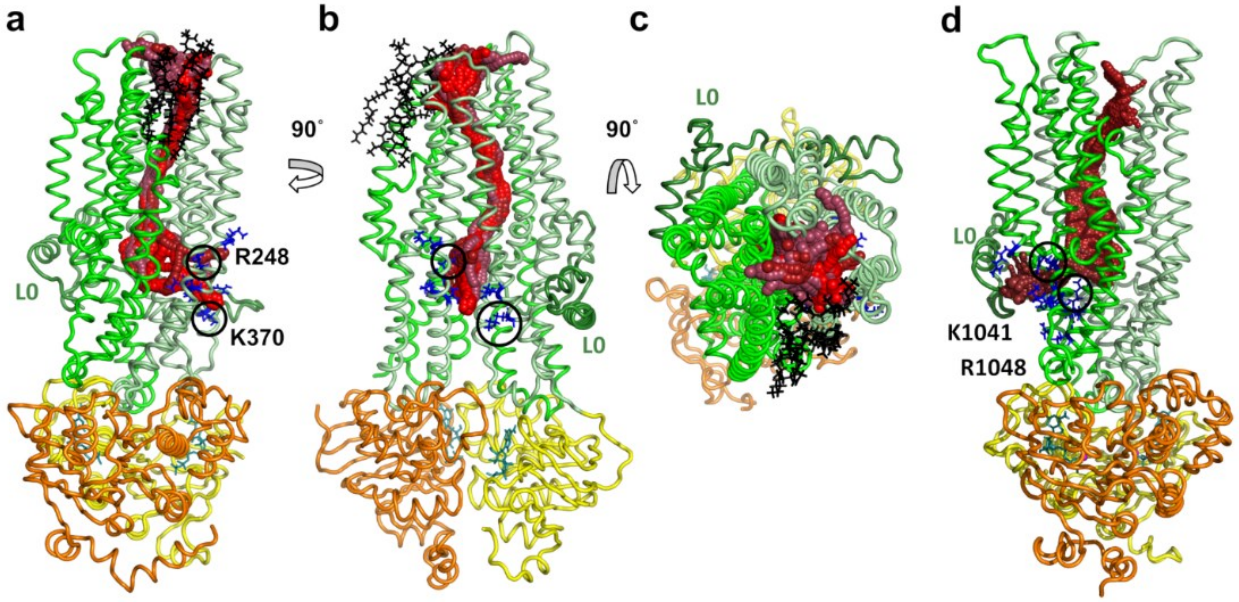

Figure 1. – The chloride pathway detected in the conformational ensemble has two entry sites on the intracellular side. *The chloride conduction pathways were identified in one trajectory from 22 equilibrium simulations with the ATP-bound CFTR structure (PDB ID: 5w81). The top scoring pathways were grouped according to the position of the amino acids lining the ion channel. Red beads represent the pathways identified by Caver. Positively charged amino acids around the intracellular entry sites are shown as blue sticks. The Lasso/L0 region, TMD1, TMD2, NBD1 and NBD2 domains are indicated by dark green, pale green, green, yellow, and orange colors, respectively. In the TM8 region, lipid molecules permeating the pore are indicated by black sticks. (a-c) Pathways opening between TM4/6 helices. (d) Pathways with lower points opening between TM10/12 helices. (a-b, d) Side view from the lipid bilayer plane and (c) top view from the direction of the extracellular space.*

Thesis I.2. The size and the charge of the substrate determine the selectivity at the cytosolic entry sites of the ion channel.

Low ion channel radius values were observed at the intracellular entry of the chloride pathway suggesting that the size of the substrate is limited by this entry region. We determined the chloride ion interaction sites around the cytosolic entry pores in the simulations and showed that chloride ions successfully entered the ion channel. In contrast, positively charged potassium ions did not form significant interactions with the CFTR protein, nor did these ions get into the interior of the protein.

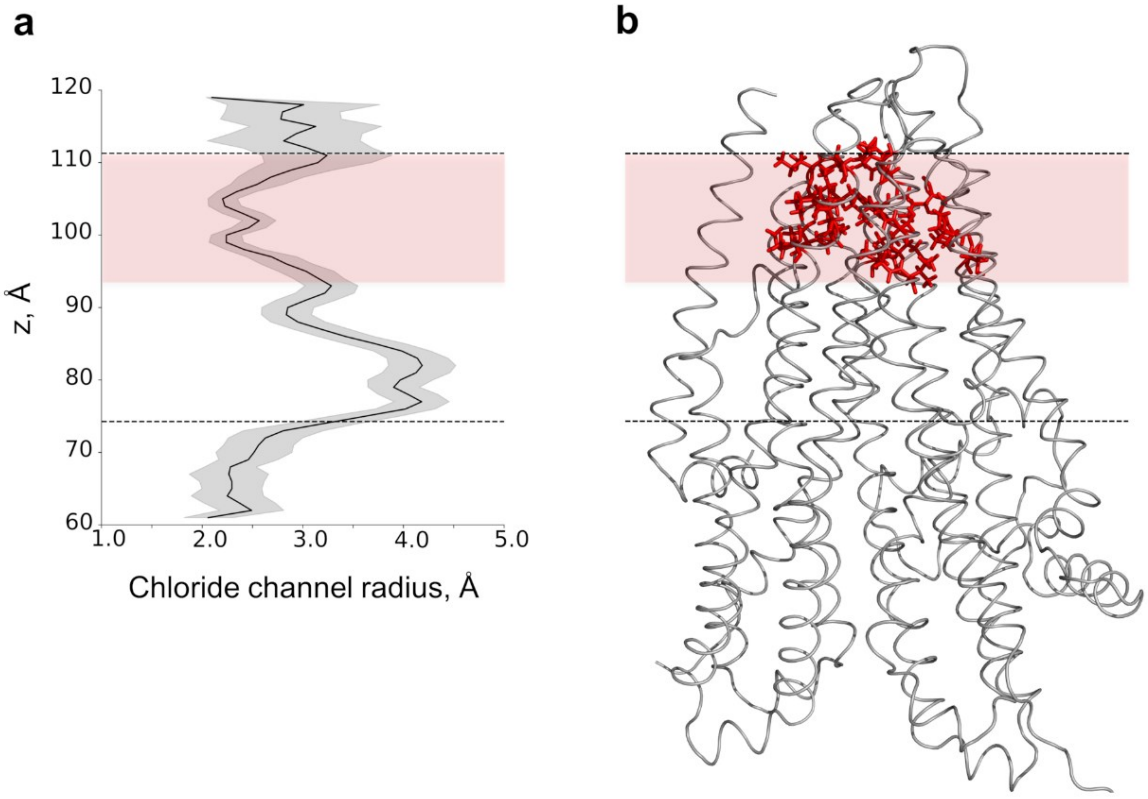

Figure 2. – The chloride channel narrows at two locations along the observed chloride pathway. *(a) The radius of the Caver defined spheres along the identified pathways from all open ion channels was averaged along the z-coordinate. The cavity located in the wide transmembrane region is characterized by large values (up to r=4 Å), while the cytosolic region and the narrow cross-sectional region (bottleneck) at the height of the extracellular side membrane layer show radius values between 2-2.5 Å. The narrow cross-sectional region located in the extracellular membrane leaflet is marked by the colored box (light red). (b) The 3D structures of the transmembrane domains are shown in the context of the ion channel profile, the channel-lining residues located at the height of the extracellular membrane leaflet are marked by red sticks.*

Thesis I.3. The potential limiting factor for chloride conductance is the so-called bottleneck region of the chloride pathway.

We observed that mainly hydrophobic amino acids are located in the bottleneck region of the chloride pathway. Additionally, the diameter of the ion channel of this region in the open conformations only slightly exceeds the size of the chloride ion, supporting the hypothesis that this section plays a role in limiting and regulating ion permeability. Furthermore, no spontaneous ion transfer was observed in our equilibrium molecular dynamics simulations, and the chloride ions entering the inner vestibule of the protein did not pass through the bottleneck region.

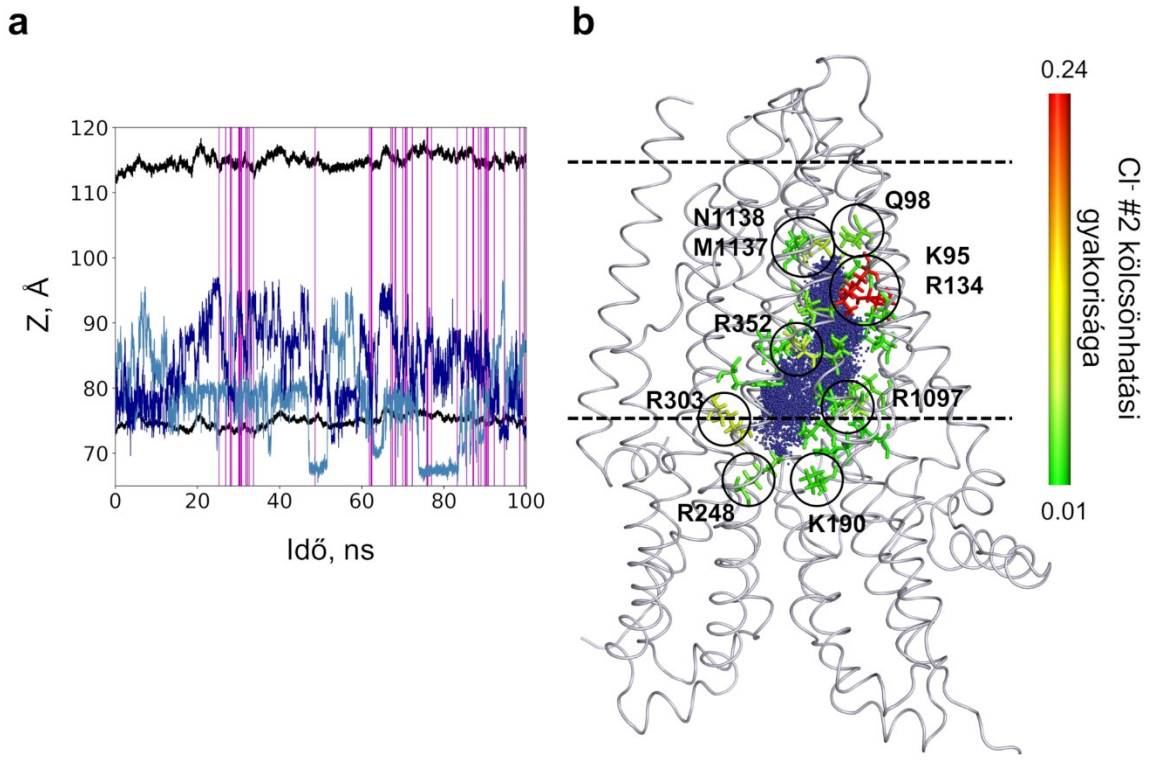

Figure 3. – In the equilibrium simulations chloride ions entered the ion channel but did not pass through the membrane region. *(a) The position of the two chloride ions entering the protein is shown by blue lines (#1 light blue and #2 blue). The black lines show the boundaries of the membrane bilayer, defined by the center of mass of the phosphorus atoms in the POPC molecules. Vertical lines (purple) marks open conformations along the simulation trajectory. (b) The positions of chloride #2 in the TM region, determined from all structures, is shown with blue dots in the context of the initial structure. The interaction frequency of this chloride ion and the protein is mapped onto the structure (amino acids in stick representation), and is color coded from green (low) to red (high). The amino acids with the highest interaction frequency are K95 and R134 (red).*

Thesis I.4. The kinked segment of transmembrane helix 8 (TM8) and the displacement of transmembrane helix 7 (TM7) observed in CFTR cryo-EM structures can inhibit ion channel function.

The equilibrium simulations with the zebrafish CFTR structure showed that the rate of ion channel openings was extremely low (0.047%). The kink of transmembrane helix 8 (TM8) in the membrane region caused lipid molecules to penetrate the ion channel, which may also contribute to the rare ion channel openings. This observation was supported by equilibrium simulations using a structural model obtained after straightening TM8 and moving TM7 to a position typical of homologous ABC proteins by homology modelling. In these simulations, we observed increased channel opening (2.038%), and in this case the lipid molecules could not penetrate the ion channel. These results suggest that the kink of TM8 may be a consequence of the structure determination procedure.

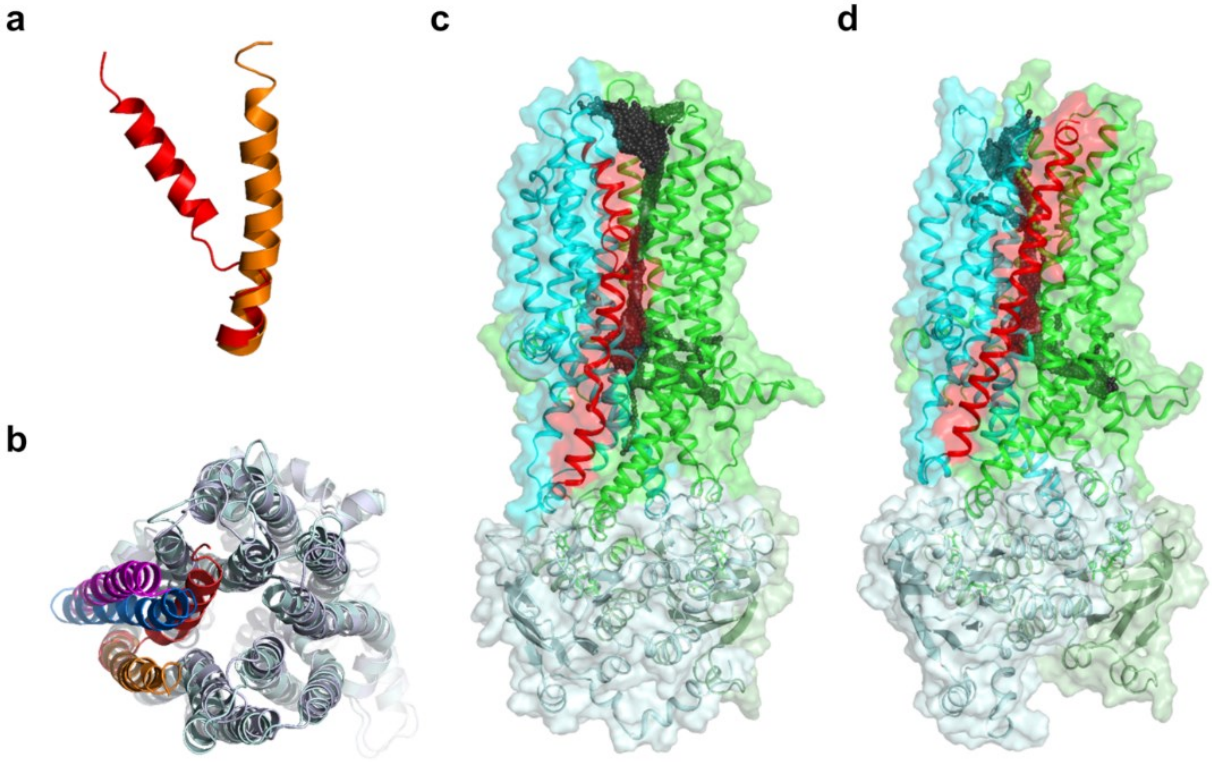

Figure 4. – The straightened TM8 results in more frequent openings in the simulations, and the ion channel remained isolated from lipid molecules in the membrane region. *(a) The kinked TM8 in the zebrafish structure (red) and the straight TM8 in the homology model based on the MRP1 structure (orange). (c-d) Cartoon and surface representation of the zebrafish CFTR structure. Black beads indicate the observed chloride pathways. (b) Extracellular side view of the aligned CFTR (pale cyan) and homology modeled (pale purple) structures (TMD1-2). CFTR TM8 (red) and TM7 (magenta), homology model TM8 (orange) and TM7 (blue).*

Corresponding publications: J1, J2

2. Study of the mechanical unfolding of the NBD1 domain of the CFTR chloride channel

2.1. Motivation and aims of the study

Several mutations causing cystic fibrosis (CF) affect the first nucleotide-binding domain (NBD1) of the CFTR protein. The most common mutation is the deletion of phenylalanine 508 (ΔF508/F508del). The ΔF508 mutation leads to the misfolding of the NBD1 during protein maturation resulting in the degradation of a significant amount of the produced CFTR protein [10], [11]. In addition to the misfolding problem, although a small fraction of the synthetized proteins can reach the plasma membrane, this mutation inhibits the function and reduces the stability of CFTR compared to the normal variant [12]. The long-term efficacy of currently used compounds that restore functional expression of CFTR is still not known [13]–[15]. The compounds used are mostly aimed to stabilize the folded proteins that are already in the plasma membrane, but these are inefficient for misfolded proteins. For this type of mutations, therapeutic efficacy could be enhanced to the greatest extent by adjusting the protein folding, which can help the protein to achieve its native structure. For the design of efficient CFTR correctors, it is essential to understand the CFTR structure and dynamics at atomic level, as well as a detailed description on the NBD1 folding process.

Previous studies have investigated the folding of NBD1 [16]–[19], but these experiments are low resolution. Detailed studies of isolated wild type and ΔF508 mutant NBD1 revealed a significantly reduced melting temperature of the mutant domain [20], [21]. The deletion of the amino acid F508 may cause a difference in the stability of NBD1 compared to the wild type, which may be reflected in the altered mechanical resistance of the protein. Therefore, one of our objectives was to investigate the resistance of NBD1 to mechanical unfolding at higher resolution in the wild type and the ΔF508 mutant NBD1.

Force spectroscopy methods have been successfully used to investigate the mechanical unfolding of proteins, their resistance to unfolding forces [22]–[25]. Therefore, our group studied the mechanical unfolding of NBD1 by atomic force microscopy (AFM) experiments. Although force spectroscopy experimental results provide close-to-reality data, they are often difficult to interpret and have low resolution. To overcome these challenges force-probe molecular dynamics simulations can be used. The advantage of this computational method is that it can be used similar to the force spectroscopy experiments to investigate the mechanical unfolding of proteins, but it also allows the characterization of

unfolding events at the atomic level [26], [27]. This method allows a more accurate description of the detachment of the secondary structural units and the unfolding pathways compared to experiments.

Taking advantage of the simulation method, we aimed to study the unfolding pathways determined by the experiments in more detail. To this end, we used force-probe (pulling) molecular dynamics simulations with a simplified Gō model (heavy-atom Gō, HA-Gō) [28] to investigate the mechanical unfolding of NBD1. The simplified model was required due to the size of NBD1 (~205 amino acids).

It has been shown using biophysical experiments that the ΔF508 mutation thermodynamically and kinetically destabilizes NBD1 and presumably contributes to the formation of kinetic traps with locally favorable energetics. Experimental studies have also investigated the folding of NBD1. It has been found that the amino acid F508 can create important interactions during folding that may play a role in directing the folding process. It was also shown that the ΔF508 mutation affects the speed of folding, which may cause the inappropriately timed folding of different segments or the inappropriate insertion of segments within the domain [16]–[18], [29], [30]. These critical elements are the β-strands S8 and S7, which are synthetized after the synthesis of the α-subunit, but which must be inserted between the N-terminal region of the catalytic domain and the α-subunit. This proper insertion could be hindered by the rapid folding of the α-subunit, and its inappropriate interactions [17], [18]. The β-strands S8, S7 and S6 and the α-subunit form a single unit in the structure, which are sequentially continuous as well. The amino acid F508 is located in this region, whose mutation prevents the correct folding of the protein. This region is referred to in the dissertation as the S6-α-S8 core (α-subunit including the α-helix H6 + the three β-strands of the β-subunit: S6, S7, S8).

In order to characterize more precisely the effect of F508 amino acid on the S6-α-S8 core region, we investigated the S6-α-S8 core unfolding pathways using fully solvated all-atom force-probe (pulling) unfolding simulations (AA-MD), the intermediate structural states and the non-native interactions during the mechanical unfolding for both the wild type (WT) and the S6-α-S8 core region carrying the ΔF508 mutation. Furthermore, we compared the steps of the unfolding process by analyzing the force-extension curves obtained from simulations and AFM experiments.

BIA (5-bromoindole-3-acetic acid) is a small molecule that has the potential to affect the folding of the NBD1. Since BIA stabilized NBD1 against thermal unfolding [31], it is likely to remain bound to the domain after NBD1 folding. Understanding the binding site of BIA could be of therapeutic importance. On the one hand, this knowledge may help to redesign the BIA molecule to increase its efficiency. On the other hand, it may contribute to the design of new molecules that bind to this site. Therefore, the objective of our group was to elucidate the corrector mechanism and binding site of BIA. In our experiments, we investigated whether the resistance of NBD1 to mechanical unfolding is altered by the addition of BIA.

2.2. New scientific results

In the second part of my thesis, I described our results related to the mechanical unfolding of the CFTR NBD1 domain. We investigated the mechanical unfolding of NBD1 similarly to the force spectroscopy (AFM) experiments of our group, but at higher resolution. We performed pulling molecular dynamics simulations using a simplified HA-Gō (heavy-atom Gō) model. From these simulation data, we determined the unfolding sequence of the secondary structural elements of NBD1, the unfolding pathways. These results contributed significantly to the interpretation of the unfolding patterns obtained from our force spectroscopy (AFM) experiments.

The S6-α-S8 core is a structurally distinct and a sequentially continuous region in the NBD1. The amino acid F508 is located in this region. To characterize more accurately the effect of F508 amino acid on the S6-α-S8 core region, we modelled the mechanical unfolding of the S6-α-S8 core using heavy atom molecular dynamics (HA-MD) simulations in the presence of explicit solvent. In addition, we investigated the intermediate structural states of the unfolding, the non-native interactions during unfolding, and the role of the amino acid F508 in stabilizing the NBD1 unfolding intermediates. Furthermore, similar to the force spectroscopy experimental results, we determined the contour length and contour length increment values from the simulations of the S6-α-S8 core unfolding. This allowed the direct comparison of experimental and simulation data and the interpretation of experimental results.

Thesis II.1. The absence of amino acid F508 in the ΔF508 mutant S6-α-S8 core causes a slight weakening of the S6 β-strand in force-probe molecular dynamics simulations.

We determined the unfolding pathways of the S6-α-S8 core region based on the fraction of native interactions of the secondary structural units during the unfolding simulations. The analyzes of the pathways showed that in the case of the ΔF508 mutant the unfolding of α-helix H6 occurred more frequently before the unfolding of β-strand S6 (ΔF508: 40%, WT: 22%). This was also shown by clustering the unfolding structures. Two structural groups were identified at the initial stage of the ΔF508 mutant's unfolding, one of which was characterized by unfolding of the α-helix H6 before the β-helix S6 (ΔF508 centroid 1, 2). In contrast, the wild type had only one initial group of structures with unfolding of the β-helix S6 before the α-helix H6 (WT centroid 1). We examined the distribution of the detachment time points of the secondary structural units and found that the ΔF508 mutant is characterized by an earlier unfolding of the β-strand S6. The examination of the distribution of rupture forces measured at the detachment time points of secondary structural units showed that the β-strand S6 unfolded at lower force values in the case of the ΔF508 mutant. In the case of the α-helix H6, there was no significant difference in the distribution of detachment time points and rupture forces between the wild type and the ΔF508

mutant. These results suggest that the weaker binding of the β-strand S6, rather than the stronger binding of the α-helix H6 characterizes the ΔF508 mutant during mechanical unfolding.

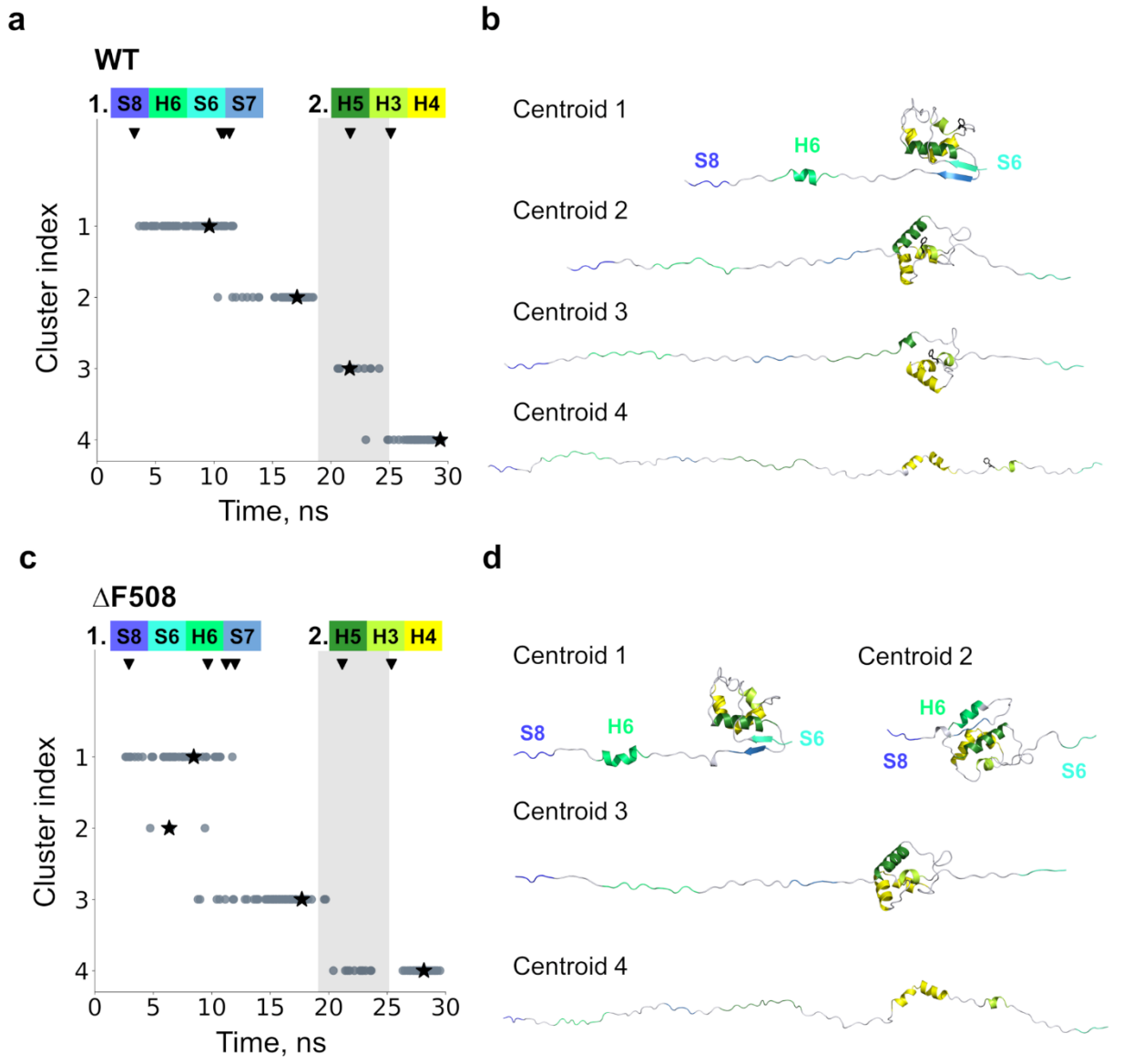

Figure 5. – Different structural clusters are characteristic to the wild type and to the ΔF508 mutant S6-α-S8 core region. *Intermediate structures from simulations of (a) wild type (WT) and (c) ΔF508 mutant were clustered using a contact based RMSD as a pairwise distance metric. The centroid elements of the resulting clusters are marked by asterisks, and the corresponding structures are shown on the right panel (b, d). The grey area highlights the structural group present in the wild type S6-α-S8 core region but absent from the ΔF508 mutant. Arrowheads mark the averaged detachment times of each secondary structural unit calculated from the fraction of native contacts.*

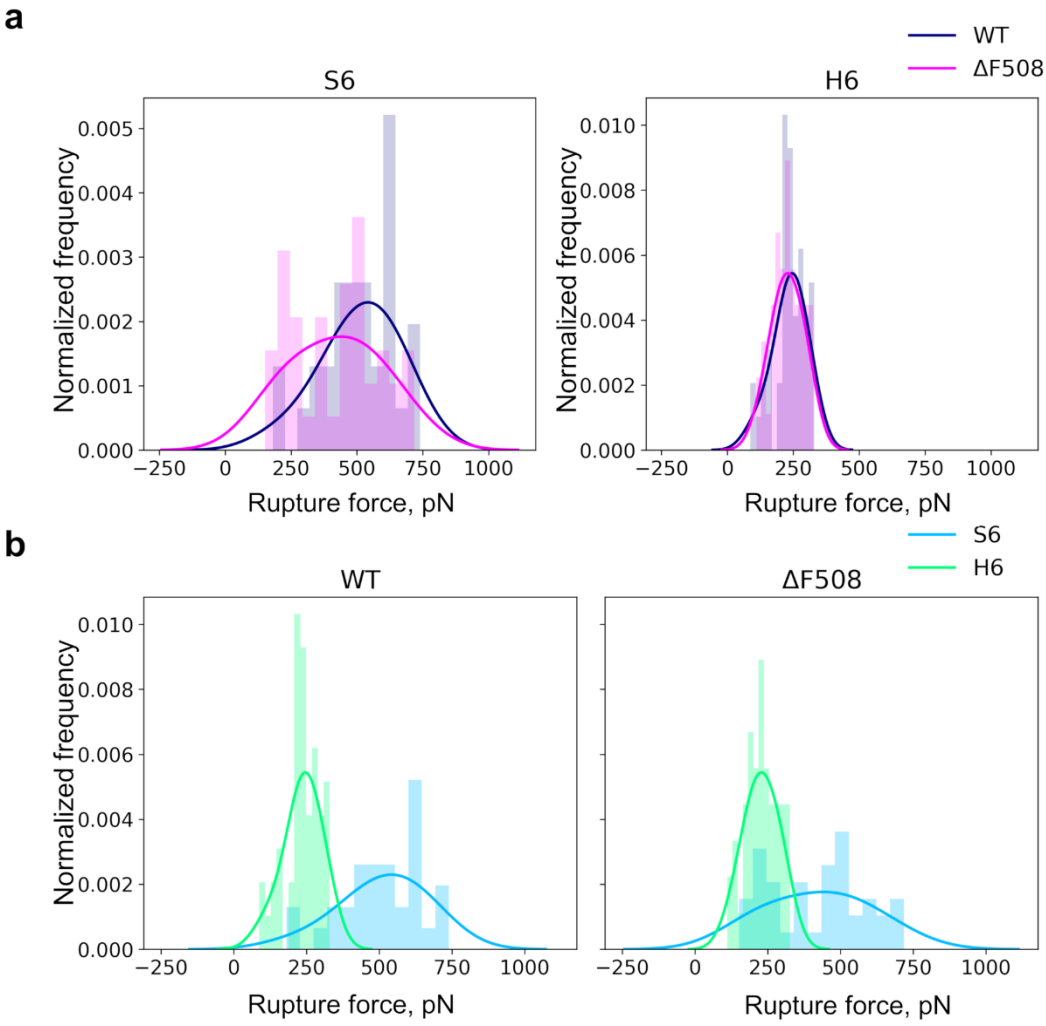

Figure 6. – In the case of ΔF508 mutant a greater proportion of β-strand S6 unfolded at lower forces, while the distribution of α-helix H6 rupture forces were unchanged. *Distribution of the β-strand S6 and the α-helix H6 rupture forces calculated from the AA-MD simulations with the wild type (WT) and the mutant (ΔF508) S6-α-S8 core region. (a) Distribution of S6 and H6 rupture forces calculated for WT (dark blue) and ΔF508 (magenta) plotted on the same graph. (b) Distribution of S6 (light blue) and H6 (green) rupture forces calculated for WT and ΔF508 plotted on separate graphs.*

Thesis II.2. The amino acid F508 and its environment form non-native interactions that may play an important role in initiating, directing, and supporting NBD1 folding during CFTR maturation.

Analysis of the unfolding of the S6-α-S8 core region showed that the α-subunit containing the amino acid F508 unfolded in the last step of the unfolding process in all simulations. Thus, during the unfolding of the wild type S6-α-S8 core, the F508 amino acid is part of the globular peptide segment. By clustering the conformations, we observed that the wild type also adopts a more persistent intermediate state before the last step of unfolding compared to the ΔF508 mutant. We observed that the F508 and its environment are involved in the formation of non-native interactions, which are not included in the native structure, during the unfolding process. The number of these interactions increased before the last step of the unfolding process, and this increase is significantly greater in the wild type. This intermediate structure formed during the unfolding process may be stabilized by non-native interactions. Many of the amino acids involved in these non-native interactions are hydrophobic and may play an important role in the early stages of folding, in triggering the initial local folding and in steering the process in the right direction. This observation was supported by folding simulations in which F508 and its environment established several interactions. The folding simulations suggested that the absence of F508 leads to the formation of a non-native interaction (Y515-Y512) that is unfavorable for correct folding, whereas the F508-Y512 interaction may be important for the formation of a correct helical orientation.

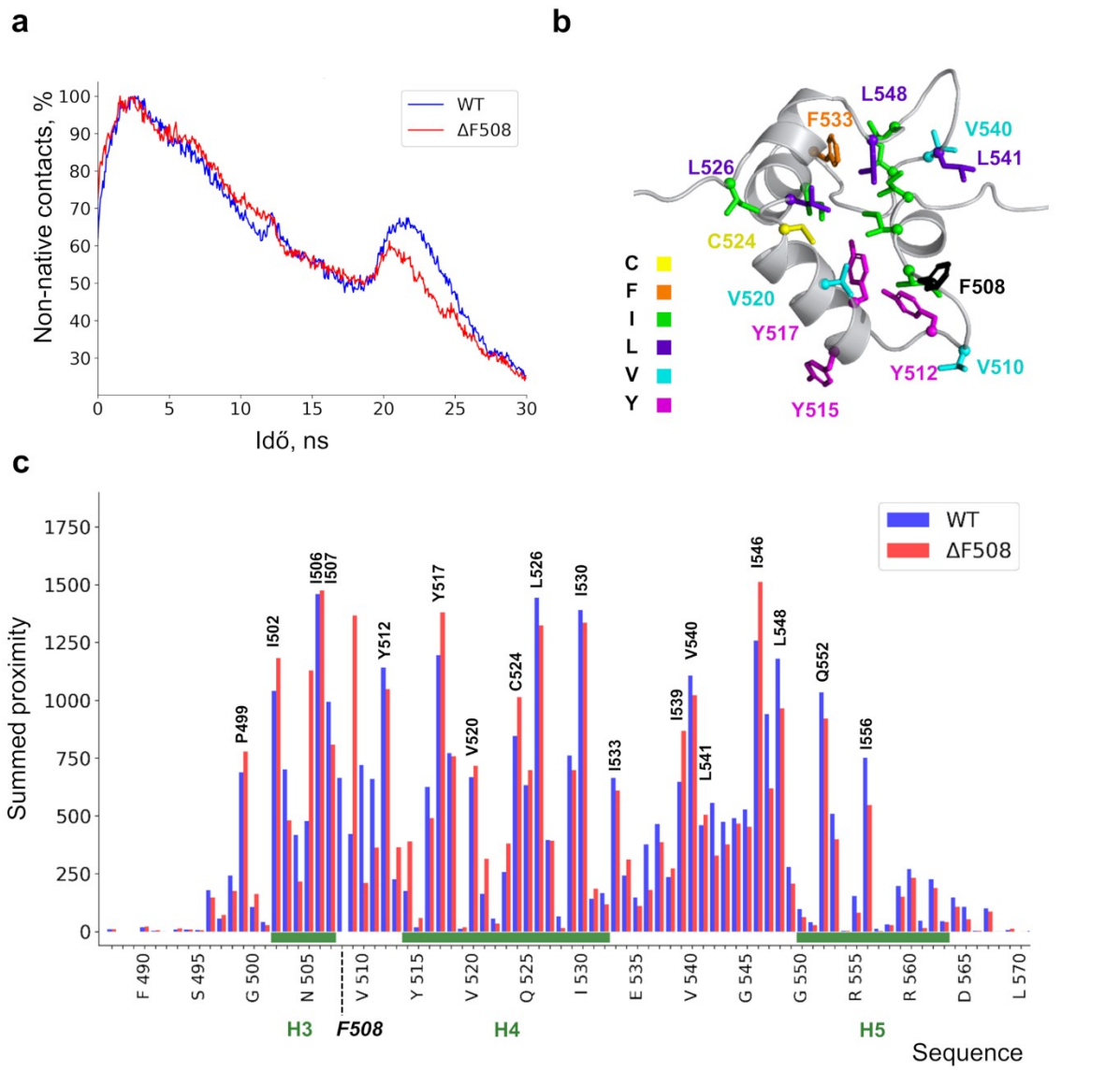

Figure 7. – An elevated number of non-native interactions is characteristic to the wild type S6-α-S8 core unfolding. *Percentage of non-native contacts obtained in simulations with (a) the wild type (WT) and the ΔF508 mutant. The proportion of non-native contacts in the 18-25 ns time window is increased for both peptide constructs but is more significant for WT. (b) Stick representation indicates amino acids forming the hydrophobic core in the structural group observed in WT before the last step of the unfolding. (c) Proximity values summarized per amino acid were used to determine the extent to which each amino acid was involved in non-native interactions in the last unfolding intermediate structure.*

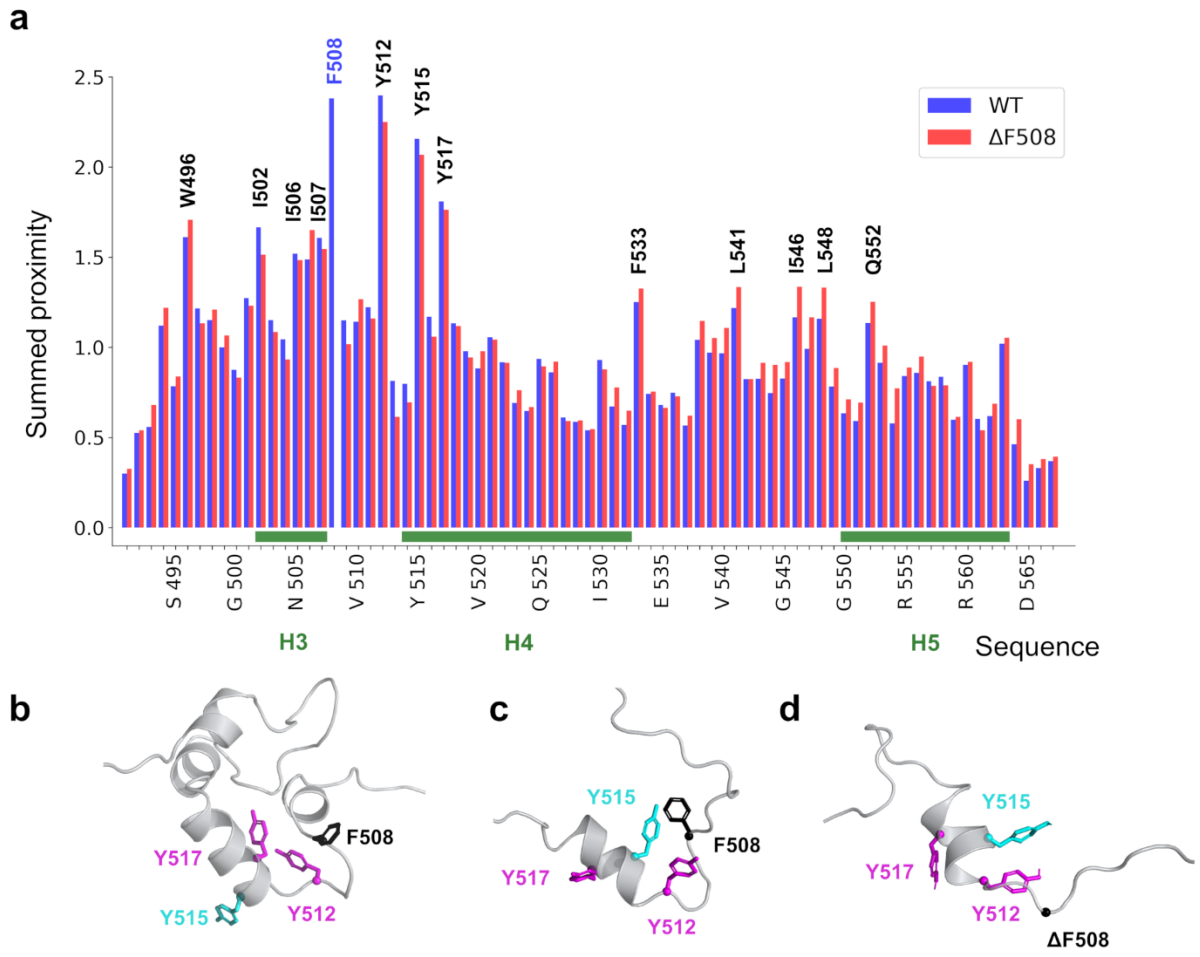

Figure 8. – F508 and its environment form local interactions during REX-DMD folding simulations. *(a) In REX-DMD (replica exchange discrete molecular dynamics) folding simulations, the pairwise amino acid contact frequency observed in the highest temperature conformations was calculated and the total contact values for each amino acid were plotted. (b) Conformations of hydrophobic amino acids in natively folded NBD1. (c) F508-Y512 and F508-Y515 are common non-native interactions in folding simulations. (d) The absence of amino acid F508 results in a lower probability of loop formation in the 508-512 region, which may negatively affect NBD1 folding.*

Thesis II.3. The ΔF508 mutation predominantly prevents the correct folding of NBD1.

Results of the experiments performed by our group showed that the unfolding patterns of the wild type and the ΔF508 mutant differ greatly. This may indicate that ΔF508-NBD1 is characterized by distinct mechanical resistance or that the protein production process resulted in a significant misfolding of the ΔF508 mutant. To resolve this, we performed force-probe (pulling) molecular dynamics simulations with the wild type and ΔF508 mutant S6-α-S8 core regions and determined the distribution of unfolding events similar to the experiments. The comparison of the results obtained from the two methods showed that the wild type experimental unfolding pattern matched well with the wild type unfolding pattern determined from the simulation data. However, the ΔF508 mutant experimental unfolding pattern differed significantly from the ΔF508 mutant simulation unfolding pattern. The difference between the results obtained by the two methods for the ΔF508 mutant derived from the fact that in the case of the mutant the simulations were based on the correctly folded native NBD1 structure. However, the protein produced *in vitro* during the experiments could contain both correctly folded and different populations of misfolded proteins as well. By the comparison, we showed that the difference observed in the experiments is not caused by the reduced mechanical resistance of the ΔF508 mutant, but by the reduced proportion of correctly folded proteins. This observation is supported by the data calculated from our experimental results, showing that the experimental force-extension curves in the case of the native S6-α-S8 core unfolding were twice as high for the wild type as for the ΔF508 mutant (WT: 56%, ΔF508: 28%). Furthermore, our AFM experiments in the presence of the compound BIA show that BIA increases the resistance of ΔF508-NBD1 to mechanical unfolding and stabilizes the α-subunit to a large extent.

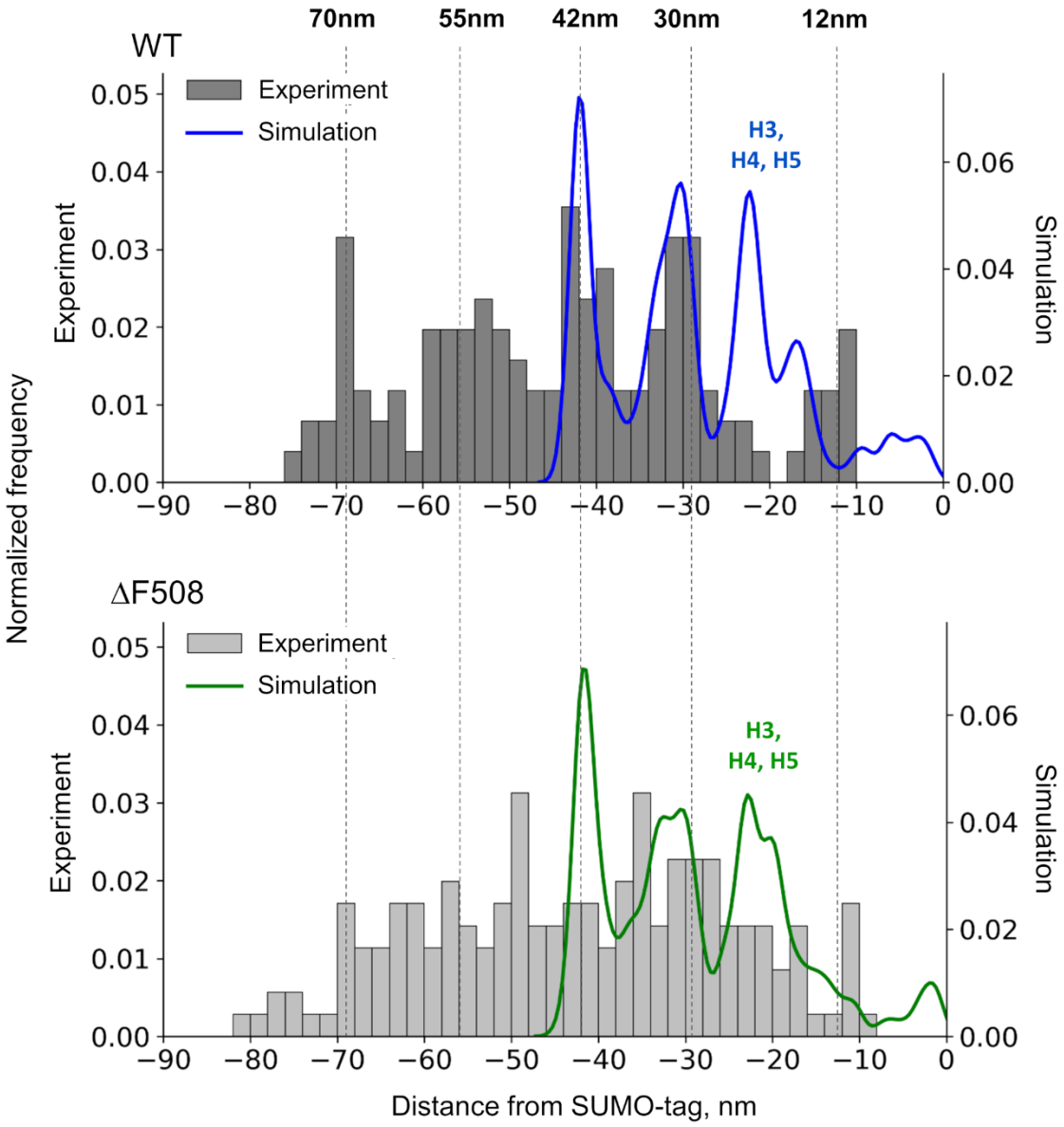

Figure 9. – There is a high degree of similarity between the distribution of contour length increments (Δk) calculated from the simulation and experimental force-extension curves of the wild type S6-α-S8 core. *Histograms show the experimental contour length (Δk) increments, measured from the reference point, from the unfolding event of the SUMO-tag (WT, n = 126; ΔF508, n = 178). The Δk values corresponding to the expected length increments based on the structure and simulations are indicated by dashed lines. The distribution of Δk values calculated from the AA-MD pulling simulations (kernel density estimate, kde) is shown by the colored lines (blue -WT, green - ΔF508).*

Corresponding publication: J3

3. Summary

Our molecular dynamics simulations allowed the characterization of the CFTR chloride pathway and the investigation of the CFTR NBD1 domain unfolding at the atomic level, which are difficult to study experimentally at high resolution. The methods and results presented in my thesis demonstrate that the combination and comparison of simulations and experimental methods can contribute to the understanding of biological processes at the molecular level.

The atomic level results obtained from the analyses of the chloride pathway can contribute to the interpretation of experimental results on the ion channel function of CFTR. In addition, they can provide useful data for understanding the mechanism of action of mutations that negatively affect the function and for the development of novel compounds aiming the restoration of ion channel function.

Examination of NBD1 using a combination of force spectroscopy experiments and simulation methods has allowed us to investigate the effect of ΔF508 mutation at the atomic level. Our results highlighted that the ΔF508 mutation likely creates unfavorable intramolecular interactions during the folding process resulting in the reduced stability of the α-subunit. This observation can contribute to the development of novel corrector molecules. We also discovered that the α-subunit can be a potential target for small molecules acting allosterically to counteract the effects of ΔF508 and other destabilizing NBD1 mutations.

List of publications

Journal publications related to the theses

[J1] B. Farkas, H. Tordai, R. Padányi, A. Tordai, J. Gera, G. Paragi, T. Hegedűs. (2019) Discovering the chloride pathway in the CFTR channel. *Cell. Mol. Life Sci.* 77:765–778. https://doi.org/10.1007/s00018-019-03211-4

[J2] T. Hegedűs, M. Geisler, G.L. Lukács, B. Farkas. (2022) Ins and outs of AlphaFold2 transmembrane protein structure predictions. *Cell. Mol. Life Sci*. 79:73. https://doi.org/10.1007/s00018-021-04112-1

[J3] R. Padányi*, B. Farkas*, H. Tordai, B. Kiss, H. Grubmüller, N. Soya, G. L. Lukács, M. Kellermayer, T. Hegedűs. (2022) Nanomechanics combined with HDX reveals allosteric drug binding sites of CFTR NBD1. *Computational and structural biotechnology journal*, 20, 2587-2599, *equally contributed https://doi.org/10.1016/j.csbj.2022.05.036

Conference publications related to the theses

[C1] B. Farkas, R. Padányi, H. Tordai, B. Kiss, M. Kellermayer, T. Hegedűs. Nanomechanics of the CFTR NBD1. *ABC2020 - 8th FEBS Special Meeting, ATP-Binding Cassette (ABC) Proteins: From Multidrug Resistance to Genetic Diseases*, 2020, Innsbruck, Austria

[C2] B. Farkas, H. Tordai, R. Padányi, A. Tordai, J. Gera, G. Paragi, T. Hegedűs. Discovering the conducting pathway of the CFTR channel using *in silico* methods. *FEBS Advanced Lecture Course, Biochemistry of Membrane Proteins – Structure, Trafficking, Regulation*, August 2019, Budapest, Hungary

[C3] B. Farkas, H. Tordai, R. Padányi, A. Tordai, J. Gera, G. Paragi, T. Hegedűs. Discovering the conducting pathway of the CFTR channel using *in silico* methods. *Gordon Research Conference, Mechanisms of Membrane Transport*, June 2019, New London, NH, United States

[C4] B. Farkas, H. Tordai, J. Gera, G. Paragi, T. Hegedűs. Characterization of the CFTR chloride channel. *ITTS Inaugural Conference*, September 2018, Vienna, Austria

[C5] B. Farkas. Investigation of intramolecular interactions of the disordered CFTR regulatory domain. *EFOP-3.6.3-VEKOP-16-2017-00002 Workshop – In memoriam Hámori József*, 2021, Budapest, Hungary

[C6] B. Farkas, R. Padányi, H. Tordai, B. Kiss, M. Kellermayer, T. Hegedűs. Characterisation of the mechanical unfolding of CFTR NBD1 using computational and experimental methods. (in Hungarian) *MaBiT - Bioinformatika 2020*, 2020, Budapest, Hungary

[C7] B. Farkas, H. Tordai, R. Padányi, A. Tordai, J. Gera, G. Paragi, T. Hegedűs. Discovering the conducting pathway of the CFTR channel using *in silico* methods. *FEBS Advanced Lecture Course, Biochemistry of Membrane Proteins – Structure, Trafficking, Regulation*, August 2019, Budapest, Hungary

[C8] B. Farkas, H. Tordai, R. Padányi, A. Tordai, J. Gera, G. Paragi, T. Hegedűs. Discovering the chloride conducting pathway of the CFTR channel using *in silico* methods. *GPU Day 2019 – The future of Computing, Graphics and Data Analysis*, July 2019, Budapest, Hungary

Other journal paper publications

[J4] B. Farkas and G. Csizmadia, E. Katona, G. Tusnády, T. Hegedűs. (2019) MemBlob database and server for identifying transmembrane regions using cryo-EM maps. *Bioinformatics* 36:2595–2598. https://doi.org/10.1093/bioinformatics/btz539

[J5] G. Csizmadia, B. Farkas, Z. Spagina, H. Tordai, T. Hegedűs. (2018) Quantitative comparison of ABC membrane protein type I exporter structures in a standardized way. *Computational and Structural Biotechnology Journal*, 16, 396–403. https://doi.org/10.1016/j.csbj.2018.10.008

PhD Proceedings – Annual Issues of the Doctoral School Faculty of Information Technology and Bionics

[P1] B. Farkas. Characterization of the CFTR chloride channel using in silico methods. *PhD Proceedings Annual Issues of the Doctoral School Faculty of Information Technology and Bionics* vol. 13 : 2018 p. 21 (2018)

[P2] B. Farkas. Describing the ion passage through the CFTR ion channel. *PhD Proceedings Annual Issues of the Doctoral School Faculty of Information Technology and Bionics* vol. 14 : 2019 p. 31 (2019)

[P3] B. Farkas. Mechanical unfolding of the CFTR chloride channel NBD1 domain using force probe simulations. *PhD Proceedings Annual Issues of the Doctoral School Faculty of Information Technology and Bionics* vol. 15 : 2020 p. 33 (2020)

[P4] B. Farkas. Effect of F508 deletion on the CFTR NBD1 nanomechanics. *PhD Proceedings Annual Issues of the Doctoral School Faculty of Information Technology and Bionics* vol. 16 : 2021 p. 33 (2021)

References

[1] H. Tordai, I. Leveles, and T. Hegedűs, "Molecular dynamics of the cryo-EM CFTR structure," *Biochemical and Biophysical Research Communications*, vol. 491, no. 4, pp. 986–993, Sep. 2017, doi: 10.1016/j.bbrc.2017.07.165.

[2] J. Das, A. A. Aleksandrov, L. Cui, L. He, J. R. Riordan, and N. V. Dokholyan, "Transmembrane helical interactions in the CFTR channel pore," *PLOS Computational Biology*, vol. 13, no. 6, p. e1005594, Jun. 2017, doi: 10.1371/journal.pcbi.1005594.

[3] G. Csizmadia, B. Farkas, Z. Spagina, H. Tordai, and T. Hegedűs, "Quantitative comparison of ABC membrane protein type I exporter structures in a standardized way," *Computational and Structural Biotechnology Journal*, vol. 16, pp. 396–403, 2018, doi: 10.1016/j.csbj.2018.10.008.

[4] B. Hoffmann *et al.*, "Combining theoretical and experimental data to decipher CFTR 3D structures and functions," *Cellular and Molecular Life Sciences*, vol. 75, no. 20, pp. 3829–3855, Oct. 2018, doi: 10.1007/s00018-018-2835-7.

[5] S. Chin *et al.*, "Lipid interactions enhance activation and potentiation of cystic fibrosis transmembrane conductance regulator (CFTR)," *bioRxiv*, Dec. 2018, doi: 10.1101/495010.

[6] Z. Zhang, F. Liu, and J. Chen, "Conformational Changes of CFTR upon Phosphorylation and ATP Binding," *Cell*, vol. 170, no. 3, pp. 483-491.e8, Jul. 2017, doi: 10.1016/j.cell.2017.06.041.

[7] Z. Zhang and J. Chen, "Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator," *Cell*, vol. 167, no. 6, pp. 1586-1597.e9, Dec. 2016, doi: 10.1016/j.cell.2016.11.014.

[8] F. Liu, Z. Zhang, L. Csanády, D. C. Gadsby, and J. Chen, "Molecular Structure of the Human CFTR Ion Channel," *Cell*, vol. 169, no. 1, pp. 85-95.e8, 23 2017, doi: 10.1016/j.cell.2017.02.024.

[9] T.-C. Hwang, J.-T. Yeh, J. Zhang, Y.-C. Yu, H.-I. Yeh, and S. Destefano, "Structural mechanisms of CFTR function and dysfunction," *Journal of General Physiology*, vol. 150, no. 4, pp. 539–570, Apr. 2018, doi: 10.1085/jgp.201711946.

[10] C. L. Ward and R. R. Kopito, "Intracellular turnover of cystic fibrosis transmembrane conductance regulator. Inefficient processing and rapid degradation of wild-type and mutant proteins," *J Biol Chem*, vol. 269, no. 41, pp. 25710–25718, Oct. 1994.

[11] G. L. Lukacs, A. Mohamed, N. Kartner, X. B. Chang, J. R. Riordan, and S. Grinstein, "Conformational maturation of CFTR but not its mutant counterpart (delta F508) occurs in the endoplasmic reticulum and requires ATP.," *The EMBO Journal*, vol. 13, no. 24, pp. 6076–6086, Dec. 1994, doi: 10.1002/j.1460-2075.1994.tb06954.x.

[12] T. Hegedűs, A. Aleksandrov, L. Cui, M. Gentzsch, X.-B. Chang, and J. R. Riordan, "F508del CFTR with two altered RXR motifs escapes from ER quality control but its channel activity is thermally sensitive," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1758, no. 5, pp. 565–572, May 2006, doi: 10.1016/j.bbamem.2006.03.006.

[13] H. G. M. Heijerman *et al.*, "Efficacy and safety of the elexacaftor plus tezacaftor plus ivacaftor combination regimen in people with cystic fibrosis homozygous for the F508del mutation: a double-blind, randomised, phase 3 trial," *Lancet*, vol. 394, no. 10212, pp. 1940–1948, Nov. 2019, doi: 10.1016/S0140-6736(19)32597-8.

[14] L. Kirwan *et al.*, "Longitudinal Trends in Real-World Outcomes after Initiation of Ivacaftor. A Cohort Study from the Cystic Fibrosis Registry of Ireland," *Ann Am Thorac Soc*, vol. 16, no. 2, pp. 209–216, Feb. 2019, doi: 10.1513/AnnalsATS.201802-149OC.

[15] K. Keown *et al.*, "Airway Inflammation and Host Responses in the Era of CFTR Modulators," *Int J Mol Sci*, vol. 21, no. 17, p. E6379, Sep. 2020, doi: 10.3390/ijms21176379.

[16] A. Khushoo, Z. Yang, A. E. Johnson, and W. R. Skach, "Ligand-driven vectorial folding of ribosome-bound human CFTR NBD1," *Mol Cell*, vol. 41, no. 6, pp. 682–692, Mar. 2011, doi: 10.1016/j.molcel.2011.02.027.

[17] S. J. Kim *et al.*, "Translational tuning optimizes nascent protein folding in cells," *Science*, vol. 348, no. 6233, pp. 444–448, Apr. 2015, doi: 10.1126/science.aaa3974.

[18] H. Shishido, J. S. Yoon, Z. Yang, and W. R. Skach, "CFTR trafficking mutations disrupt cotranslational protein folding by targeting biosynthetic intermediates," *Nat Commun*, vol. 11, no. 1, p. 4258, Aug. 2020, doi: 10.1038/s41467-020-18101-8.

[19] I. Protasevich *et al.*, "Thermal unfolding studies show the disease causing F508del mutation in CFTR thermodynamically destabilizes nucleotide-binding domain 1: Thermal Unfolding Mechanism for (F508del)NBD1," *Protein Science*, vol. 19, no. 10, pp. 1917–1931, Oct. 2010, doi: 10.1002/pro.479.

[20] W. M. Rabeh *et al.*, "Correction of both NBD1 energetics and domain interface is required to restore ΔF508 CFTR folding and function," *Cell*, vol. 148, no. 1–2, pp. 150–163, Jan. 2012, doi: 10.1016/j.cell.2011.11.024.

[21] C. Wang *et al.*, "Integrated biophysical studies implicate partial unfolding of NBD1 of CFTR in the molecular pathogenesis of F508del cystic fibrosis," *Protein Sci*, vol. 19, no. 10, pp. 1932–1947, Oct. 2010, doi: 10.1002/pro.480.

[22] R. H. Pires, M. J. Saraiva, A. M. Damas, and M. S. Z. Kellermayer, "Force spectroscopy reveals the presence of structurally modified dimers in transthyretin amyloid annular oligomers: Force spectroscopy of misfolded dimers in amyloid TTR oligomers," *J Mol Recognit*, vol. 30, no. 3, p. e2587, Mar. 2017, doi: 10.1002/jmr.2587.

[23] R. B. Best *et al.*, "Force mode atomic force microscopy as a tool for protein folding studies," *Analytica Chimica Acta*, vol. 479, no. 1, pp. 87–105, Mar. 2003, doi: 10.1016/S0003-2670(02)01572-6.

[24] S. P. Ng *et al.*, "Mechanical Unfolding of TNfn3: The Unfolding Pathway of a fnIII Domain Probed by Protein Engineering, AFM and MD Simulation," *Journal of Molecular Biology*, vol. 350, no. 4, pp. 776–789, Jul. 2005, doi: 10.1016/j.jmb.2005.04.070.

[25] L. Li, H. H.-L. Huang, C. L. Badilla, and J. M. Fernandez, "Mechanical Unfolding Intermediates Observed by Single-molecule Force Spectroscopy in a Fibronectin Type III Module," *Journal of Molecular Biology*, vol. 345, no. 4, pp. 817–826, Jan. 2005, doi: 10.1016/j.jmb.2004.11.021.

[26] J. Strzelecki, "AFM Force Spectroscopy and Steered Molecular Dynamics Simulation of Protein Contactin 4," *Acta Phys. Pol. A*, vol. 116, no. Supplement, p. S-156-S-159, Dec. 2009, doi: 10.12693/APhysPolA.116.S-156.

[27] E. Paci and M. Karplus, "Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations," *Journal of Molecular Biology*, vol. 288, no. 3, pp. 441–459, May 1999, doi: 10.1006/jmbi.1999.2670.

[28] P. C. Whitford, J. K. Noel, S. Gosavi, A. Schug, K. Y. Sanbonmatsu, and J. N. Onuchic, "An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields," *Proteins*, vol. 75, no. 2, pp. 430–441, May 2009, doi: 10.1002/prot.22253.

[29] R. A. Bartoszewski *et al.*, "A synonymous single nucleotide polymorphism in DeltaF508 CFTR alters the secondary structure of the mRNA and the expression of the mutant protein," *J Biol Chem*, vol. 285, no. 37, pp. 28741–28748, Sep. 2010, doi: 10.1074/jbc.M110.154575.

[30] V. Bali, A. Lazrak, P. Guroji, L. Fu, S. Matalon, and Z. Bebok, "A synonymous codon change alters the drug sensitivity of ΔF508 cystic fibrosis transmembrane conductance regulator," *FASEB J*, vol. 30, no. 1, pp. 201–213, Jan. 2016, doi: 10.1096/fj.15-273714.

[31] G. Veit *et al.*, "Structure-guided combination therapy to potently improve the function of mutant CFTRs," *Nat Med*, vol. 24, no. 11, pp. 1732–1742, Nov. 2018, doi: 10.1038/s41591-018-0200-x.