Bacterial Resistance via Multidrug Efflux Pumps
A Computational Study

by

Robert Schulz

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Physics

Thesis Committee:

Prof. Dr. Ulrich Kleinekathöfer (Jacobs University Bremen, Germany)

Prof. Dr. Mathias Winterhalter (Jacobs University Bremen, Germany)

Prof. Dr. Martin Zacharias (Technische Universität München, Germany)

Prof. Paolo Ruggerone, PhD (University of Cagliari, Italy)

Dr. Uwe Mamat (Forschungszentrum Borstel, Germany)

Date of Defense: 26 May 2011

Jacobs University Bremen School of Engineering and Science
Acknowledgment

First, I would like to thank my supervisor Prof. Dr. Ulrich Kleinekathöfer for giving me the opportunity to research this interesting and revealing subject as well as for the inspiring and insightful discussions. I have been glad to study in his group, where help was always available whenever needed. Furthermore, I would like to express my gratitude to Prof. Dr. Mathias Winterhalter for the ongoing support and discussions about the experimental background. Moreover, I would like to thank Prof. Dr. Martin Zacharias, Prof. Paolo Ruggerone, PhD, and Dr. Uwe Mamat for being in my thesis committee, reading my thesis, and their valuable suggestions. I am thankful to Prof. Roland Benz, Prof. Stephan Frickenhaus, Dr. Helge Weingart, Dr. Que-Tien Tran, Dr. Kozhinjampara R. Mahendran, and Dr. Attilio V. Vargiu for continuous help and valuable discussions on both the experimental as well as the theoretical part.

Special thanks to the groups of Prof. Kleinekathöfer, Prof. Heine, Prof. Materny, and Prof. Wagner as well as Astrid K. & Alex, Marina, Dr. Torge Schmidt, Katja, Eva & Niko, Astrid G., Moritz, Annemarie, Tien, Mahendran, Nadine, Thomas, Martin, Barbara, and Gabriel for all kinds of help and interesting discussions during conferences, coffee breaks, or barbecues.

I am especially grateful for the proofreaders of my thesis: Steffi, Katja, and Helge. Special thanks again go to Professor Winterhalter for the neverending, outstanding coffee supply. Last but not least, I am thankful to my family and my friends, sorry I could not mention all of you!
Abstract

Multidrug resistance against antibiotics is a common problem of treating bacterial infections. Bacteria are able to employ different resistance mechanisms to regulate the concentration of antibiotics in the cell, or disable them by degradation or target modification. One of the major contributors are efflux pumps, that reduce the periplasmic concentration of antibiotics and other toxic compounds by pumping them out into the extracellular space. This thesis comprises several computational approaches to investigate the functions of the individual proteins of Escherichia coli’s major efflux pump AcrAB-TolC. In the first part, the opening of the periplasmic tip of TolC, an outer-membrane channel, was investigated by mutating specific residues in this region as it was shown in X-ray crystal structures. Upon docking to one of several possible inner-membrane transporters, TolC is supposed to open iris-likely. In the present simulations, this opening was accomplished to a similar extent as in experiments. Interestingly, the cations in the system significantly influenced the outcome of the simulations.

In the second part, the dynamics of AcrB’s pumping process have been investigated by mimicking conformational changes which have been proposed by X-ray crystallographers. The outcome has been described by using a drug molecule as an unbiased measure. This drug was previously found in the binding pocket of a cocrystallized structure. During this investigation, the binding pocket has been found to shrink peristaltically, which forced the drug to leave the binding pocket. In the binding pocket and along the channel toward the exit gate, strongly interacting amino acids have been pointed out as well. In a subsequent work, the influence of its environment has been examined. Especially the interaction with the surrounding water and amino acids revealed several essential features of the extrusion by AcrB. One interesting aspect was the stream of water which slightly assisted the extrusion of the drug and also shielded the electrostatic interaction of the protein and the substrate. This disfavored a strong binding of the drug to the amino acids of the exit channel. Furthermore, the inter- and intramonomeric interactions have been examined and crucial conformational changes of the transition along the cycle have been highlighted. The effect of one particular mutation on the extrusion has been investigated in an additional work. The information presented in this thesis enhances the atomistic and functional understanding of E. coli’s major efflux pump. By combining these data with experimental results, the mechanism of extrusion can be interpreted more clearly, which might also improve the chances to develop new antibiotics and inhibitors with a higher efficacy in the future.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abstract</strong></td>
<td>v</td>
</tr>
<tr>
<td><strong>1. Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1. Transport of Antibiotics through the Bacterial Cell Wall</td>
<td>2</td>
</tr>
<tr>
<td>1.2. Thesis Overview</td>
<td>5</td>
</tr>
<tr>
<td><strong>2. Opening of the TolC Aperture</strong></td>
<td>9</td>
</tr>
<tr>
<td>2.1. Transition between Conformations of TolC</td>
<td>10</td>
</tr>
<tr>
<td><strong>3. Mimicking the Extrusion Process of AcrB</strong></td>
<td>29</td>
</tr>
<tr>
<td>3.1. Drug Extrusion via Functional Rotation in AcrB</td>
<td>30</td>
</tr>
<tr>
<td>3.2. Role of Water during the Extrusion in AcrB</td>
<td>46</td>
</tr>
<tr>
<td>3.3. Partial Targeted MD of AcrB</td>
<td>67</td>
</tr>
<tr>
<td>3.4. Effect of F610A Mutation on Substrate Extrusion</td>
<td>79</td>
</tr>
<tr>
<td><strong>4. Conclusion &amp; Outlook</strong></td>
<td>81</td>
</tr>
<tr>
<td><strong>M. Methods and Materials</strong></td>
<td>85</td>
</tr>
<tr>
<td>M.1. Molecular Dynamics Simulations</td>
<td>85</td>
</tr>
<tr>
<td>M.2. Techniques for Simulation &amp; Analysis</td>
<td>87</td>
</tr>
<tr>
<td><strong>S. Supporting Information</strong></td>
<td>95</td>
</tr>
<tr>
<td>S.1. Drug Extrusion via Functional Rotation in AcrB</td>
<td>95</td>
</tr>
<tr>
<td>S.2. Role of Water during the Extrusion in AcrB</td>
<td>105</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>111</td>
</tr>
</tbody>
</table>
In nature, bacteria can be found in a multitude of extreme “situations” including geysers in Yellowstone National Park, black smokers in the Atlantic ocean as well as in the human gastrointestinal tract. A popular example is Escherichia coli (E. coli), a commensal inhabitant of our intestines [1]. Therein, it is threatened by bile acids of the digestion system as well as by antibiotics used to treat severe bacterial infections. To dispose of these toxic compounds, bacteria are able to employ various resistance mechanisms [2–5]. One direct approach is to regulate the transport of material into and out of the cell.

To separate the cell interior from the extracellular space, bacteria institute so-called membranes, which are composites of two lipid layers, usually referred to as bilayer [1,6]. In conjunction with a mesh of peptidoglycan, they establish the cell wall of bacteria. This descriptive scheme is rather coarse because there are two different major setups of this cell wall, which are experimentally determined using the Gram method. In so-called Gram-positive bacteria, such as Staphylococcus aureus, Streptococcus pneumoniae, and Corynebacterium diphtheriae, there is only one cell membrane which is fortified by a thick layer of peptidoglycan on the outside. The peptidoglycan layer of Gram-negative bacteria, e.g., Escherichia coli, Pseudomonas aeruginosa, and Salmonella enterica, is much thinner than in the first type, being between two lipid bilayers in the so-called periplasmic space. In this case, the outer membrane is asymmetric containing lipopolysacharides in the outer layer and phospholipids in the inner leaflet (see Fig. [1.1]).

Despite common belief, not all bacteria are pathogenic to human beings. But those which are can often lead to severe diseases that require treatment with antibiotics. Since the first pioneering efforts in the 1930s and 40s, the importance of antibiotic treatment has risen steadily [8]. This is due to the fact that many different types of bacteria, parasites, and fungi can be treated with the various kinds of antibiotics. While the
first antibiotics, such as Penicillin, where of natural origin, there are also several kinds of synthetic antibiotics on the market now [9]. In general, different antibiotics act at specific action sites: structure and synthesis of DNA, RNA, proteins, as well as the cell wall. Despite scientific advancements, the antibiotic resistance of bacteria has increased especially within the last few decades [4, 10–14]. The major reasons are intensive usage of antibiotics and locally high selection pressure, e.g., in hospitals.

1.1. Transport of Antibiotics through the Bacterial Cell Wall

Gram-negative bacteria such as E. coli developed a network of transport mechanisms to regulate the in- and efflux of molecules, e.g., nutrients and toxins [6, 15]. On the one hand, the outer membrane is populated by outer membrane channels, so-called porins, responsible for the uptake of material, e.g., OmpF [6]. This particular kind of porins barely filters the passing molecules, often allowing harmful substances such as antibiotics to enter the periplasmic space. Although the transport through porins is passive, there are other types that are highly selective, allowing only passage of one or few molecules [6]. On the other hand, bacteria employ different proteins and protein complexes to keep the

Figure 1.1: Sketch of the cell wall of Gram-negative bacteria (adapted from [7]).
concentration of noxious compounds at a subtoxic level \cite{14}. These export complexes evolved naturally according to the environment of the bacteria, e.g., the human intestines with toxic bile acids \cite{12}.

While porins are mainly found in the outer membrane (OM) of Gram-negative bacteria, the inner membrane (IM) is occupied by transporters which either use ATP hydrolysis or a concentration gradient across this membrane to drive their function. This enables the bacterium to regulate the internal concentration of all kind of substances. The most important gradient used by IM transporters is the transmembrane proton gradient generated by proton pumps usually driven by ATP hydrolysis. This gradient is also the major driving force behind the extrusion of antibiotics in E. coli and similar kinds of bacteria.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{coordinated_network.pdf}
\caption{Scheme of the “coordinated network of efflux transporters” in Gram-negative bacteria to dispose of toxic compounds. (Figure used with permission of Prof. S. Schuldiner \cite{16}).}
\end{figure}

This thesis focuses on efflux transport of toxic molecules from Gram-negative bacterial cells, especially via the protein complex called efflux pump \cite{17,23}. This complex is part of a machinery employed by bacteria such as E. coli to facilitate the extrusion of noxious
molecules and thereby the survival of the organism. Schuldiner et al. [24] investigated a “coordinated network of transporters” to describe the path of efflux across both membranes (see Fig. 1.2). The article illustrates a two-step scheme which means one step per membrane to be passed. In the inner membrane, transporters of the major facilitator superfamily (MFS) and small multidrug resistance (SMR) family use the transmembrane proton gradient to pump substrates from the cytoplasm into the periplasm. There they are taken up by so-called efflux pumps and pumped through the outer membrane into the extracellular space.

In E. coli, the major efflux pump is AcrAB-TolC, a association of three proteins taking up toxins from the periplasmic space and pumping them into the extracellular space. Embedded in the inner membrane, a transporter of the resistance-nodulation-division (RND) family – AcrB – acts as the active part of this construct [25][28]. It is driven by the transmembrane proton gradient that induces changes of protonation states in the protein’s transmembrane domain. This domain is mainly composed of α-helices. These chemical changes lead to conformational changes in the neighboring helices which ultimately result in the pumping of substrates within the pore domain. In order to achieve this, the mechanical energy has to be transduced toward the latter domain by kinking of specific helices, supposedly transmembrane helix 8. These conformational changes promote the expel of the bound substrates from the binding pocket and extrusion toward the TolC docking domain which is docked to the periplasmic tip of TolC.

This second protein, which is basically a 140 Å long channel embedded in the outer membrane, is only open toward the periplasmic space when connected to an IM transporter [29]. Interestingly, TolC is expressed separately to AcrA as well as AcrB and is able to team up with several transporters, e.g., the α-hemolysin ATP-driven transporter HlyB as well [30]. In comparison to other OM channels such as OmpF, TolC is not only composed of a β-barrel embedded in the outer membrane but has an α-helical channel protruding from the membrane-embedded barrel into the periplasm. Thereby, it allows to cover the path toward its partnering IM transporters. The last partner of this efflux pump is AcrA [31]. The function of this periplasmic protein, which is probably linked to the inner membrane [32], is only partially understood yet. On the one hand, it is supposed to stabilize the complex by fusing TolC and AcrB. During the docking of all three proteins, AcrA might facilitate the opening of the TolC aperture [33]. On the other hand, it has been hypothesized that AcrA might transmit conformational information between the other proteins of this efflux pump [34].

The AcrAB-TolC complex is formed by homotrimers of AcrB and TolC. Despite the fact that this pump and its components have been studied for many years, the actual number of AcrA monomers is not yet proven. In general, a 3-3-3 stochiometry is considered most probable meaning that three monomers per protein are required to constitute
one efflux pump. In a recent crosslinking study, Symmons et al. [32] illustrated a probable setup by combining a symmetric structure of AcrB [26], the partially open structure of TolC [33], and a homology model of AcrA derived from a refined structure of MexA, which was introduced in this paper as well. This mentioned homologue of AcrA is part of an entire homologous efflux pump in P. aeruginosa: MexAB-OprM [35–41]. The combined structure of such an efflux pump describes a continuous pathway from the periplasm close to the inner membrane via AcrB and TolC into the extracellular space. In general, the protein walls of this path are impenetrable for pumped substrates, although there are reports about vancomycin being able to use the TolC-HlyB complex as an entrance into the bacterial periplasm [42,43].

Efflux pumps are experimentally investigated using and combining different techniques such as site-directed mutagenesis, measurement of minimum inhibitory concentration (MIC) or fluorescence intensities, crosslinking experiments, X-ray crystallography, and electrophysiology. This thesis was however performed to complement the experimental results on a more detailed, atomistic level. Therefore, all-atom, classical molecular-dynamics (MD) simulations are employed as the major tool. Despite the lack of information about subatomic, electronic states, this approach can help to understand the conformational changes and basic mechanisms that lead to the extrusion of toxic substrates, such as antibiotics, from the bacterium. Recent computational studies highlighted the leverage of combining computational approaches with experimental ones. During the analyses, the obtained data are to be connected to the available experimental results. Therefore, the spatial and temporal gaps between the approaches have to be covered. Using the MD program NAMD [64], the individual proteins are examined separately allowing to apply adapted simulation techniques dependent on the kind of protein in focus. The individual methods used in this thesis are described in more detail in Ch. M.

1.2. Thesis Overview

This thesis constitutes a collection of computational studies of the protein components of E. coli’s major efflux pump AcrAB-TolC. Initially, this chapter gives a general introduction into this thesis’ topic. More specifically, the focus of this thesis is described including a brief collation of the experimental and computational approaches in this field of research. The next two chapters contain both subprojects separately examining the individual proteins of the complex, i.e., TolC (Ch. 2) and AcrB (Ch. 3). AcrA was not examined intensively here as its function was believed to become more evident when studied in the complex with AcrB and TolC (compare Ref. [58] for structural informa-
Figure 1.3: Proposed setup of the complex of E. coli’s major efflux pump using the individual crystal structures (based on Ref. 32). The displayed proteins are TolC (brown), AcrA (green), and AcrB (blue, yellow, and red) which are dissolved in water and partially embedded (TolC & AcrB) in lipid membranes. The scale on the right hand side indicates the approximate sizes of the major segments of the complex, i.e., transmembrane domain of AcrB, periplasmic part of AcrB, α-helical channel of TolC, and β-barrel of TolC, from bottom to top respectively. The arrows describe the flow direction of protons/cations (cyan) and of substrates (orange).
presented articles is added at the end of this thesis (Ch. 8).

The first protein to be exhibited is TolC in chapter 2. It presents a study about the opening of the periplasmic aperture wherein mutations have been applied similarly to previous crystal structures [33]. During these simulations, cations hindered the opening by reestablishing salt bridges, that had been removed by the mutations, in so-called ion pockets in the aperture. To get rid of these bound cations at a time scale accessible by MD simulations, an electric field was applied. Thereby, the occupation frequency of these cation pockets was reduced. Several simulation parameters were varied in this work, such as KCl concentration, field strength, and field direction.

Afterwards, an investigation comprising several subprojects on the extrusion process of the transporter AcrB is described (Ch. 3). This part is based on a functional rotation hypothesis which was deduced from an asymmetric structure obtained by X-ray crystallography. Using these pieces of information, the conformational changes between two states have been induced in simulations to observe their effects on an unsteered drug molecule. This substrate was found during a cocrystallization in the binding pocket of one monomer. The major method of choice for this investigation is targeted MD, which is described in more detail in Sec. M.2.3. The presented studies mainly focus on this particular monomer. In the first subproject, the actual expel from the binding pocket and extrusion toward the funnel to TolC was the main subject. Furthermore, the proposed peristaltic pumping was observed between the subdomains constituting the binding pocket. Additionally, trial simulations have been performed to mimic the inverse direction of the rotation cycle and test the hypothesis about the cycle direction.

In a related work presented afterwards, more details about the extrusion process are described, such as interaction of drug with environment and influence of the surrounding water. During this study, the components of the nonbonded interaction energy are evaluated and related to the environment of the drug. Moreover, the overall water movement is measured to characterize the dynamics of water molecules, especially in the vicinity of the substrate. At last, a new approach is describe to use targeted MD more selectively by only steering some domains or monomers of the transporter. This allows to mimic the conformational changes of the transmembrane domain, where the chemical energy from the proton gradient is converted into mechanical energy, and observe the influence on the pore domain, where the actual pumping occurs. Furthermore, the neighboring monomers of the occupied one can be steered to better understand the cooperative effects between the monomers. In the last chapter, the presented results are summarized and possible next steps are discussed.
CHAPTER TWO

OPENING OF THE TOLC APERTURE

Summary of “Transitions between Closed and Open Conformations of TolC: The Effects of Ions in Simulations”

The article presented in the following section comprises a study of how the periplasmic tip of TolC opens to allow the passage of substrates from a corresponding transporter toward the extracellular space. To simulate TolC in a lipid bilayer, unbiased MD simulations have been carried out initially, which verified the stable closed conformation. A first idea to apply a voltage across the membrane to measure the current of ions through this long channel had to be discarded, due to the very low conductance of 80 pS compared to 2-4 nS of OM porins like OmpF at room temperature (compare Ref. 54 for TolC and Refs. 62, 66, 67 for OmpF). Hence, mutant structures have been compared to the wildtype as it was done in a crystallographic study before [33]. In this publication, Bavro et al. presented a partially open structure of a mutant TolC protein. This structure was considered as reference structure for the following investigation, because there is a lack of information about the natural open state. In vivo, this transition from closed to open is believed to occur in an iris-like fashion of the helices forming the aperture upon docking of TolC to one of several possible IM transporters, e.g., AcrB or HlyB. Here, the concept of MD simulations was enhanced by using an electric field to speed up the dynamics especially of the ions. This was considered necessary, because ions got stuck in the vicinity of the mutation site at the periplasmic tip of TolC. These sites are at the boundary between two neighboring monomers with negatively charged amino acids
attracting cations into so-called ion pockets. Due to the ion pockets being occupied by ions, the aperture remained closed during unbiased MD simulations. During this study, different KCl concentrations, field strengths, and field directions have been examined to identify the setup which leads to similar results as the crystal structures obtained by Bavro et al. [33]. By reducing the salt concentration, the probability of occupation of these ion pockets was decreased significantly. Furthermore, the direction of the applied electric field was crucial as well. For a positive field direction, the cations had to travel through the $\beta$-barrel and the $\alpha$-helical channel. Thereby, they were constrained to pass the narrow aperture with the ion pockets which increased the probability for cations to be bound in one of these pockets. By inverting the field direction, the cations did not necessarily pass through this narrow periplasmic entrance. In this case, it was more likely for them to not enter the TolC channel, but rather pass by and move on toward the lipid bilayer. This study emphasized the possible influence of ions on the protein structure, especially in association with the limited time scale of MD simulations.

Transitions between Closed and Open Conformations of TolC: The Effects of Ions in Simulations [61]

Authors: Schulz, R. and Kleinekathöfer, U.


2.1.1. Introduction

During the last few decades, the problem of antibiotic resistance rose to a serious problem within the field of antibacterial treatment [8]. One reason is the evolutionary overexpression of multidrug efflux pumps within the bacteria’s cell envelope [14][68]. In Gram-negative bacteria, e.g., Escherichia coli (E. coli) or Pseudomonas aeruginosa (P. aeruginosa), the cell envelope consists of two lipid bilayers, so-called membranes, with the periplasm in between. A large number of proteins have been found to be embedded in both the inner and the outer membrane of E. coli [1][69]. There are also systems of proteins that span over the whole cell envelope, creating a pathway from the cytoplasm or
the inner membrane’s outer leaflet to the extracellular space. One important transport class therein is the multidrug efflux pumps. These efflux systems are employed to export antibacterial drugs such as antibiotics as well as protein toxins out of the cell. This mechanism can diminish the effect of certain classes of antibiotics. Hence, new antibiotics have to be found which are able to overcome this mechanism.

Embedded within the outer membrane, proteins belonging to the outer membrane factor family, like the channel-tunnel protein TolC, act as exit ducts for different efflux pumps. It is believed that TolC is closed as long as it is not connected to any compatible inner-membrane transporter, e.g., HlyB or AcrB of E. coli. These examples describe the versatility, because in contrast to AcrB which is a RND transporter driven by proton-motive forces, HlyB is an ABC transporter which uses ATP to export its substrates. The closed structure of TolC shall ensure that no substrate is able to flow from the extracellular space into the periplasm. Furthermore, a membrane fusion protein is required to stabilize the fusion of transporter and tunnel; in the two cases above, these are HlyD and AcrA, respectively. The number of membrane fusion proteins needed to form a functioning efflux pump of, e.g., AcrB and TolC, is so far not well known, but in Zgurskaya and Nikaido, AcrA was proposed to be a trimer. While connecting to a transporter, the periplasmic aperture of TolC shall open iris-like, but the final aperture radius might depend on the corresponding transporter. Additionally, it was shown that, under certain circumstances, the antibiotic vancomycin is able to use the TolC-HlyB complex as entrance into the periplasm.

To understand the mechanism of TolC docking to a transporter, it has to be investigated in which way the TolC aperture opens upon docking. Furthermore, it has to be examined which residues are responsible for keeping the protein closed and how the residues belonging to the aperture move into an open conformation. By understanding this transition and the final state, it is possible to study the transport of ions and translocation of antibiotics through TolC in its open state. Finally, one can get a better insight into how TolC docks to transporters like AcrB or HlyB.

The crystal structure of TolC has been published in 2000 and shows a mono-barrel homotrimer (see Fig. 2.1B). The 100 Å long, periplasmic tunnel of coiled coils consists of 12 α-helices with two pairs from each monomer (Fig. 2.1A). One pair of each monomer is turned toward the center of the tunnel forming the aperture at the lower periplasmic end. In the middle of this periplasmic part, the equatorial domain is situated around the helices, which contains both short α-helices and short β-strands and probably functions as a hinge for the opening motions of the other helices. The other part of TolC, which is, in vivo, embedded in the outer membrane, consists of 12 β-sheets that form a 40 Å long β-barrel-like a channel through the membrane. The diameter of this channel is roughly 35 Å.
Figure 2.1: TolC protein structure in cartoon representation of the secondary structure (50) with (A) the β-barrel above and α-helical coiled coils below. (B) The monomers in different colors. (C) The protein in a simulation box with the orientation of the positive applied voltage (see Sec. 2.1.2). (D) Zoom to the aperture with the constricting ring in licorice representation of the side chains (dark shaded, T152 and D153; light shaded, Y362 and R367). (E) Side view of panel D looking from the periplasm toward the membrane part (figures created with VMD [77]).

Most previous ideas and hypotheses that describe the open state and/or the opening mechanism when getting into contact with its counterparts to form an efflux pump were based on rigid body assumptions [76,78–80]. In one very recent study [57], all-atom and coarse-grained molecular dynamics (MD) have been employed to investigate dynamical properties of wild-type (Wt) and one mutated TolC. These simulations as well as this one greatly differ from the rigid-body studies, because the flexibility of the protein structure is taken into account, enabling us to observe uncoiling of the periplasmic-coiled coils. Basically, springs and structural angle constraints are used to describe the quantum-mechanical nature of the chemical bonds that ensure the stability of the protein. In Vaccaro et al. [57], only equilibrium simulations have been performed and analyzed, whereas in this article, external forces are applied in most of the simulations.

Motivated by a recent crystal structure [33] of a TolC double mutant showing a par-
tially open conformation, we simulated two comparable double mutants and the Wt structure to investigate the dynamic differences that lead to the opening. As in the experiment, the point mutations have been symmetrically applied to two residues of all three monomers in the constricting ring (Fig. 2.1E, blue), which in vivo, keeps the aperture closed via salt bridges and H-bonds. This allows us to compare the Wt to two double mutants, derived from the Wt, as well as to the open crystal structure. Additionally, we also simulated the corresponding single mutants to one double mutant, which helped us to examine the importance of each point mutation for the effects to be observed in the double mutant.

Since our mutants are derived from the Wt crystal structure, they are in a closed conformation that allows us to dynamically observe how TolC might open due to the removal of the bonds in the constricting ring. In this article, the following questions will be addressed, among other topics:

- Is it possible to observe opening of the TolC aperture during ~10-ns-long MD simulations?
- Will both double mutants (or any single mutant) relax into conformations that are similar to the open crystal structure?
- Is it possible to speed up the simulations by applying an electric (steering) field in such a way that the closed TolC structure transits into a partially open structure?

As will be shown below, applying an electric steering field for speeding up the simulations can be very important in obtaining the opening dynamics in the short simulation time available. The steering field enables us to overcome certain energy barriers that typically require much more time to be passed by thermal fluctuations. On the one hand, studying the effect of electric fields was motivated by recent investigations in MD simulations of membranes and membrane proteins [62, 81–85], where their properties, including stability of the lipid bilayer and the protein structure, were investigated. (Using this information, we could ensure that no artificial effects due to the electric field occurred.) On the other hand, the electric steering field was motivated by the so-called grid-steered MD [80], which has been developed to steer molecules through membrane pores. In this study as well as in our investigations, the artificially applied electric field is used to speed up the MD simulations and is not meant to be applied in experiments. Additionally, several experimental studies have been performed on TolC [52–54], which helped us to focus on the important segments of the protein, i.e., basically the aperture region with a ring of residues that stabilize the constriction.
2.1.2. Simulation Setup

As stated above, the crystal structure of the wild-type (PDB code 1EK9) is simulated and compared to mutated versions of the same structure as well as a partially opened crystal structure. In contrast to the mutations in the literature \[52-54\] and to the mutation in the recent MD study \[57\], the mutations for the partially open crystal structure by Bavro et al. \[33\], Y362F and R367E, have been applied (i.e., tyrosine 362 is replaced by phenylalanine and arginine 367 by glutamate). For this mutant, two crystal structures with the symmetry variants C2 and P212121, PDB codes 2VDD and 2VDE, respectively, were obtained. The two symmetry variants are due to variations in the crystallization setup. The simulations denoted "MCryst" below were started from the PDB 2VDE and compared to simulations starting from the closed crystal structure derived from Wt, but with the mutations equal to MCryst denoted "MGlu". Thus, the sequence of the residues in the mutated crystal structure MCryst and our similar computer-generated mutant MGlu are identical. To investigate whether the effects on the mutant are due to the change in charge and/or length of the side chain, we also applied a double mutation Y362F and R367D, i.e., residue 367 is mutated into an aspartate which has a shorter side chain than the mutated residue above. This mutant is denoted "MAsp" in the simulations according to the amino acid into which the second residue is mutated (see Tab. 2.1).

Using VMD \[77\], the crystal structure of TolC was aligned to the z axis with the periplasmic helices pointing in positive z direction, placed into a preequilibrated POPE lipid bilayer in the x-y plane and solvated in TIP3P water \[87\]. At the end of the buildup phase, all lipid and water atoms that overlapped with the protein have been removed. Furthermore, KCl ions have been added at random positions in the water to neutralize the charge of the protein and to generate specific salt concentrations (0.1 M and 1 M). This setup leads to a periodic box size of $\sim(107 \times 107 \times 173) \text{ Å}^3$ after the equilibration of the Wt, with slight deviations for the mutants shown in Table 1. These mutants have been obtained using the Mutator plug-in of VMD.

For the MD simulations, the MD program NAMD \[64\] was employed together with the CHARMM27 force field \[88\] and periodic boundary conditions. The particle-mesh Ewald method was used with a grid spacing of maximal 1 Å per grid point in each dimension. After minimizing the system for up to 50,000 steps, the atom velocities were randomly chosen to resemble a Maxwell distribution. The temperature was kept at 310 K by applying Langevin forces to all heavy atoms with the Langevin damping constant set to 5 ps$^{-1}$. The integration time step was chosen to be 1 fs, to ensure the accuracy of the simulations under applied voltages. During the equilibrations, the pressure was kept at 1.01325 bar for 10 ns in the NpT ensemble using the Nosé-Hoover Langevin
Table 2.1: Details of the different simulations: The index carries a character '0' at the end for a neutralized system with just the required number of counter ions (24 potassium), a '1' for a 0.1 M concentration, and a '2' for a 1 M concentration. Furthermore, the voltage can be read from the last letter of the index: a for +1 V, b for -1 V.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Index</th>
<th>KCl Conc. [M]</th>
<th>(N_{\text{atoms}})</th>
<th>Voltage [V]</th>
<th>Time [ns]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wt1a</td>
<td>0.1</td>
<td>203,907</td>
<td>+1.0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Wt1b</td>
<td>-1.0</td>
<td>20</td>
<td>+1.0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Wt2a</td>
<td>1</td>
<td>200,547</td>
<td>+1.0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Y362F+R367D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAsp1a</td>
<td>0.1</td>
<td>203,868</td>
<td>+1.0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>MAsp1b</td>
<td>-1.0</td>
<td>20</td>
<td>+1.0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>MAsp2a</td>
<td>1</td>
<td>200,508</td>
<td>+1.0</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Y362F+R367E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGlu0a</td>
<td>0</td>
<td>204,067</td>
<td>+1.0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>MGlu0b</td>
<td>-1.0</td>
<td>20</td>
<td>+1.0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>MGlu1a</td>
<td>0.1</td>
<td>203,877</td>
<td>+1.0</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>MGlu1b</td>
<td>-1.0</td>
<td>20</td>
<td>+1.0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>MGlu2a</td>
<td>1</td>
<td>200,517</td>
<td>+1.0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Y362F+R367E</td>
<td>crystallized</td>
<td>0.1</td>
<td>198,139</td>
<td>+1.0</td>
<td>20</td>
</tr>
<tr>
<td>MCryst1a</td>
<td>-1.0</td>
<td>20</td>
<td>+1.0</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>MCryst1b</td>
<td>0.1</td>
<td>198,139</td>
<td>+1.0</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

piston pressure control. The van-der-Waals energies were calculated using a smooth cutoff (switching radius 10 Å, cutoff radius 12 Å). All subsequent simulations with an external electric voltage were carried out in the NVT ensemble with the direction of the field parallel to the z axis (see Fig. 2.1C).

Although the simulations with electric steering fields have been run for at least 20 ns, only the later 10 ns have been used for deep analysis (see Table 1). During the simulations with applied voltage, the \(\beta\)-barrel of TolC has been constrained to the central position of the simulation box. The constraints were imposed by harmonic forces with the force constants set to 1 kcal/M·Å². We want to emphasize once more that the external electric field is used as steering force to speed up the MD simulations similar to the grid-steered MD [86]. For example, within a simulation with an applied electric voltage of 1 V, the additional force experienced by monovalent charges is of magnitude of \(F \sim 9\) pN parallel to the z axis, which would correspond to an energy \(W \sim 1k_BT\) if an monovalent ion is moved 5 Å against the field direction. This sample calculation is based on the field strength, an approximation for the electrostatic force and corresponding work. The proposed ion movement, which will be discussed below, would also take place by thermal motion on a timescale larger than milliseconds. This is far beyond the time
scales accessible by MD simulations.

2.1.3. Results

In the following, we analyze the simulations described in Table 1. Before going into the details of the individual simulations, we show in Fig. 2 that the simulation results of the Wt structure and the MAsp as well as MGlub mutants are qualitatively different when applying an external steering voltage. As can be clearly seen, the Wt structure remains in its closed conformation while the mutated versions open partially. Furthermore, Fig. 2.2 visualizes that the mutant MAsp opens asymmetrically, which appears to be qualitatively similar to the asymmetric C2 crystal structure reported in Bavro et al. [33]. This asymmetry makes it inappropriate to use circular area calculations for the opening area. To quantify the effects of opening, we calculated the distances between representative aperture residues for which we chose G365 (Fig. 2.2). Again, the asymmetric opening of the MAsp mutant becomes apparent. Additionally, the last row shall visualize the paths of the aperture tips (C\text{\textalpha} atoms of G365), which also indicates that they do not move equally.

To be able to compare the degrees of opening using one number, we calculated the triangular area between the C\text{\textalpha} atoms of the residues G365 to get an average over the three monomers. This is certainly not the only way to calculate the area of the aperture that is actually overestimated, but it is easy to perform, and works in asymmetric structures for which other measures might not. Using methods to calculate the area of the channel more accurately at the position of residue G365 lead to certain problems, because the protein structure can no longer be described by a closed surface as one approaches conformational states that are more open. The residue G365 was chosen in accordance with Bavro et al. [33], and shall indicate the motion of the helix pairs, but not the change of the internal area at the constriction, which is difficult to calculate because of the asymmetry.

For comparison, the real internal area at the constriction of the Wt crystal structure can be estimated with $A \sim 16.7 \text{\AA}^2$ (or radius $r = 2.3$ \text{\AA}) using the program CAVER [89]. As mentioned above, this program calculates only a circular area.

An alternative approach to visualize and quantify the motion of the residues belonging to the aperture is called the porcupine plot, used before, e.g., in Barrett et al. [90] and Törnroth-Horsefield et al. [91]. It is employed in Fig. 2.3 to represent the direction and magnitude of the deviation per residue below the equatorial domain. Here, the deviation for the final structure of MGlub with respect to the partially open crystal structure (2VDE) is shown using a color scale indicating the magnitude and the maximum deviation (values < 2 \text{\AA} are not shown). As can be seen easily, two monomers moved further away from the reference while the third one remained close to it. Looking at the
Figure 2.2: Results of simulations with applied voltage: In the first row, snapshots of the protein with backbone (cartoon) and all-atom (transparent surface) representation of the last frame are drawn using the same perspective as in Fig. 2.1E. In the second line, the distances (in Å) between the C_α atoms (van-der-Waals representation) of G365 are displayed with the protein backbone in transparent cartoon representation. In the third line, the traces of the residue G365 C_α atoms are shown with time-dependent coloring, starting from the equilibration.
trajectory of MGlul0b, the aperture is approximately passing through the state of the partially open crystal structure. We want to note in passing that the aperture is not moving closer to the membrane, indicated by the upper image. This can be also verified during the whole trajectory (data not shown).

Concerning the stability of the system, the standard analysis including RMSD calculations has been performed. There are no apparent distortions of the structure, but the obtained data is not expressive since the direction of the motion is not taken into consideration. Most important, all protein parts below the equatorial domain (see Fig. 2.1D) are moving. The closer the residues are to the aperture, the more they tend to move (see upper image in Fig. 2.3). This indicates that the equatorial domain functions as a hinge.

Figs. 2.4 and 2.5 show the triangular area for the simulations with an applied voltage of +1 V and -1 V, respectively. As mentioned above, previous to all simulations there is a 10-ns equilibration phase without a steering field. The triangular area did not
change much in these equilibration runs (data not shown), except for MGlul0. In the latter case, the protein already begins to open during the equilibration. Starting at the equilibrated crystal structures, Wt and MCryst, and applying a steering field, the areas also remain rather constant for the two simulations as can be seen in Figs. 2.4 and 2.5. This is not a surprise, because the voltage is not strong enough to significantly influence the protein itself. For the computer-generated mutants, the situation is different. For positive applied voltage, the mutant MGlu remains closed in the presence of a 100 mM KCl solution. Additionally, MGlu0a tends to close again toward an aperture size similar to MGlu1a, which indicates that the presence/absence of ions is of some importance. However, the simulations of MAsp do start to open and the triangular areas rise to values that are slightly larger than for MCryst. This is true for all investigated concentrations. When prolonging the trajectory for the MAsp2a case to 30 ns (data not shown), the area declines again slightly to approximately the value found for MCryst. Additional simulations with +0.5 V (data not shown) showed similar results, which indicates that the actual field strength is not of high importance, but instead, the presence of the field is. The reasons for this will be discussed below.

In the case of a negative applied voltage, MGlu and MAsp both show an opening of
the protein. In contrast, prolonging the field-free equilibration runs of the mutants to 20 ns (data not shown) does not lead to a partially open conformation. The data for MGlu2 is not shown in Figs. 2.4 and 2.5 because the results are similar to those of MGlu1. For a voltage of -1 V, slightly more time is needed until the aperture begins to open. For the wild-type, there are no qualitative differences between Wt1 and Wt2. Comparing the values of MAsp1a and MGlu1b, the latter one reaches greater values in the triangular area than the former one, due to a higher asymmetry of the aperture region which does not significantly increase the area values. Furthermore, the helices in MAsp1a (see Fig. 2.2, second column), which are rather close to the initial conformation, move slightly inward. This can be interpreted as natural constraints of the coiled-coils versus breaking apart of the structure, especially the α-helical barrel.

Concluding, the electric steering voltage has an influence on the opening or at least on the speed of opening for some of the described setups, but not for all of them. The side-chain length of the mutated residue 367, which is the difference between the mutants MGlu and MAsp, seems to have some influence on the strength of the constricting ring. Of course, our findings are always restricted to the simulated time spans that were up to 40 ns in total (10 ns field-free equilibration plus 30 ns including external field).

To analyze the findings, we determined the electrostatic potential maps, which show

Figure 2.5: The same as in Fig. 2.4 but for a voltage of -1 V.
Figure 2.6: Slices through the potential map for different simulations: (A) Wt1a, (B) MGlul1a, (C) MAsp1a, and (D) MAsp1b.

The global distribution of all charges over the whole trajectory [81]. Some of these maps are shown in Fig. 2.6 as slices through the middle of the protein \((y = 0)\) which are averaged over all frames of the trajectory from 10 ns to 20 ns with applied voltage [81,82]. Within this figure, the protein oriented along the \(z\) axis, and the lipid bilayer with the highest potential values along the \(x\) axis, are well distinguishable from the solution. Using this method, one can see the open aperture for both MAsp1a and MAsp1b (Fig. 2.6 C and D). In contrast to this mutant, MGlul does open in MGlul1b (similar to Fig. 2.6 D), but not in MGlul1a (Fig. 2.6 B). This cannot be understood readily just by looking at the secondary structure and global motions.

Because this more global analysis of the charge distribution does not lead to a clear explanation of the electric field effects on the opening of the tunnel, a more local analysis was performed to understand why the aperture remains closed, e.g., in the simulation MGlul1a. To this end, we investigated the region close to the mutated residues (Fig. 2.1 E) considering the amino-acid conformations and charge locations. As described earlier [33], mutating R367 inhibits the salt bridge toward D153. At the same time, though,
this mutation leads to a cation (affinity) pocket formed by the negatively charged residues D153 and D367/E367 as well as the polar T152, which can be occupied by potassium ions in this case. A similar residue configuration was described in Blaustein et al. [92], where Ca$^{2+}$-binding domains of a Na/Ca exchanger protein have been investigated.

![Figure 2.7: The MGlue aperture (cartoon representation) with the residue side chains (licorice representation colored by amino-acid type) of the cation pockets (transparent surface representation) occupied with potassium ions (van-der-Waals representation). Same perspective as in Fig. 2.1 E. Each cation pocket consists of the residues T152, D153, and E367.](image)

While analyzing the pathways of the ions, especially of the potassium ions, through the aperture and close to the three cation pockets between the monomers (Fig. 2.7), one can easily observe the attraction of potassium ions toward these pockets, which leads to new formations of ionic bonds mediated by these potassium ions. Especially, one can easily distinguish between simulations with positive and negative voltages applied. In the positive case, the cations flow through the channel and therefore pass closely by the cations’ pockets. In the negative case, the cations instead accumulate in the region beside the protein and therefore comparatively seldom come close to the cation pockets. Particularly for the case of MGlue0b, i.e., a simulation with negative applied voltage and minimal ion concentration, there are basically no cations close to the cation pockets (see also discussion below).

The occupation of the cation pockets over time during the last 10 ns of some of the applied voltage simulations is shown in Fig. 2.8. In these occupation numbers, a clear difference between the simulation MGlue1a (which does not show opening) and MAaSp1a (which does show opening) can be seen. Fig. 2.8 indicates that as long as at least two cation pockets are occupied frequently, no opening of the channel occurs. This can be
Figure 2.8: Occupation for each of the three cation pockets with ion selection range of 3 Å. A black line represents that the respective pocket is occupied.

understood directly, since each monomer is connected to two of the pockets and both have to be empty to allow opening motions of this particular monomer.

Furthermore, the average occupation frequency was calculated over a time span of 10 ns for all three pockets in the different simulations and stages of the simulation, i.e., the equilibration; the first 10 ns; and the subsequent 10 ns with applied steering field (Fig. 2.9). Additionally, some simulations have been run with a voltage of +0.5 V using the state after the first 10 ns of the +1 V simulations. This was found to be applicable,
because the structure did not change much within the first 10 ns. Despite the decreased voltage, the results are similar with a slightly higher occupation frequency of the cation pockets.

Comparing these occupation frequencies with the conformations of the aperture, the correlation is quite striking. If the occupation frequency is < 0.2, the protein definitely opens. For values between 0.2 and 0.24, the opening depends on the concentration and the observed stage, e.g., both values for MAsp2a are similar although only the latter state is opened, which describes the transition into the partially open conformation. Hence, the correlation between occupation frequency and final conformation could be confirmed for all simulations performed with the double mutants that built up cation pockets. It has to be stated clearly that this observation is within the timescale currently accessible by MD simulations.

In addition, simulations with the single mutants from Y362F as well as R367D have
been performed (data not shown) to examine the contribution of each residue to the strength of the constricting ring. Regarding the triangular area values (as in Fig. 2.4), it can be seen that the values remain lower than the values of MGlu for both single mutants. For R367D, the occupation of the cation pockets has been investigated, which resulted in an occupation frequency < 0.2, although the protein did not open during the simulation. This verifies that, although the cation pockets are relatively empty, the protein is not able to open — which means that the occupation frequency threshold is only applicable for the double mutants. In general, it can be stated that both mutations have to be applied to enable the opening of the aperture without external stimuli.

2.1.4. Conclusion

The aim of this study was to investigate the dynamic effects of two point mutations, Y362F and R367E, on the aperture stability of TolC within MD simulations. In recent experiments [33], double mutations of this kind lead to a partially open structure. Interestingly enough, using equilibration runs at physiological salt concentrations did not lead to a major change in the conformation on a 10-20 ns timescale. This is also seen in Vaccaro et al. [57], with the mutations Y362F and R367S. By minimizing the salt concentration toward a solely neutralized system MGlu0, the protein opened slightly during the equilibration, which could finally be explained by the corresponding low ion-pocket occupation frequency (Fig. 2.9). In this simulation, the helices rotate iris-like toward an open conformation (Figs. 2.2 and 2.3), as also indicated in Bavro et al. [33]. This leads to the suggestion that it might be favorable to use a minimal ion concentration when equilibrating mutants. Otherwise, ions might establish temporary ionic bonds that could increase the stiffness of the protein; see also Laine et al. [93]. Furthermore, it was checked that it is not enough to simply mutate residues Y362 or R367 separately; a double mutation is needed to induce an opening of TolC.

The simulations showed that applying an electric steering field is an alternative approach to speed up simulations and to investigate the aperture stability of TolC. But in contrast to the crystal structure from Bavro et al. [33], the mutated protein does not open symmetrically under physiological conditions within the simulated time. Furthermore, the opening depends on the side-chain length and the charge/polarity of the residues within the constricting ring. One reason is that the ions are able to influence the opening by formation of ionic bonds in place of the wild-type salt bridges that have been deleted by the mutations. Although only potassium has been investigated here, it seems unlikely that the observed effects of the ions are restricted to this ion type. Because of two negative side chains in the pocket, other cations like Na⁺ or Ca²⁺ should have similar or stronger effects due to the charge, which is indicated in Blaustein

25
et al. [92]. Furthermore, the method employed for the initial ion placement might be a critical point. In this study, the ions were distributed randomly in the water layer. If a Poisson-Boltzmann solver would be used to determine the positions for the ions at the electrostatic extrema, the cations might be placed directly into the cation pockets - which would be rather unfavorable for this and similar investigations.

In this study, the effect of bond reformation is likely to be only a problem of the timescale within MD simulations. In addition to the results specific to TolC, this investigation also shows how the ions of the solution can affect large domain motions. This is especially the case for spontaneous motions, and might be applicable to entirely different systems as well. For example, in Laine et al. [93], it was revealed that the structural flexibility of calmodulin depends on the number of bound calcium ions. Another example is given in Sotormayor and Schulten [94], where calcium ions protect cadherin from unfolding.

The obtained data can be used to investigate the electric conductance of the open state of the protein, which can be compared to experiments similar to studies of OmpF [62]. In that study, the temperature dependent ion conductance of OmpF has been investigated. Therefore, it has to be examined how stable the opened structures are and which influence the mutations have on the measurements. Moreover, the rather symmetrically opened structure from MGlu0b can be used to analyze the docking of TolC with in vivo partners like AcrB and AcrA. In several hypothetical studies [33, 79, 80], it has been investigated how TolC might connect to AcrA. While opening the aperture of TolC, the α-helices turn to open a groove that is favorable to dock with the α-helical domain of AcrA. In Bavro et al. [33], it was pointed out that the connection to AcrB is required before the groove is accessible. Furthermore, it was stated that the bonds between the residues of the constricting ring are substituted by new bonds toward the TolC docking domain of AcrB. As can be seen from these simulations, the influence of the electrostatics and possible bound ions on the opening of TolC is rather important. In future studies, one could try to put charges such as those from AcrB close to the TolC Wt structure and analyze possible opening motions. As a starting point one could use the results of docking studies between the partially open crystal structure of TolC and AcrB [33]. The obtained information from these next steps can be used to understand the mechanism of efflux pumps and antibiotic resistance in more detail, which is important for designing new antibiotics.

2.1.5. Acknowledgments

We thank Ben Luisi and Vassiliy Bavro for making Bavro et al. [33] and the crystal structures available before publication as well as for interesting discussions. We also
thank Mathias Winterhalter, Paolo Ruggerone, and Matteo Cecarelli, together with their group members, for valuable comments.
CHAPTER
THREE

MIMICKING THE EXTRUSION PROCESS OF ACRB

Summary of “Functional Rotation of the Transporter AcrB: Insights into Drug Extrusion from Simulations”

The study described subsequently was a first step to understand the relationship between the structure and the dynamics of AcrB. This work was based on the knowledge of the asymmetric X-ray crystal structure which was used to develop a hypothesis of a transition cycle, describing conformational changes between the states found in the crystal structure. This idea has been taken here to examine the effects of enforced conformational changes. Therefore, a drug molecule, which was found in the binding pocket of a cocrystallized structure [25], was used as indirect measure of these transitions. It has been shown that by inducing the transition along the proposed cycle direction, the drug, not being steered by the MD program in any way, has been squeezed out of this region and toward the exit. This expulsion has been facilitated by a zipper-like closure of the binding pocket which was initiated by the neighboring subdomains. This so-called peristaltic pumping is believed to be part of the essentials of the extrusion. It drives the substrate molecule out of the binding pocket and into the exit channel, which has been observed in the decline of interaction energy between drug and binding pocket as well. Furthermore, an opening of the exit channel is required to allow the extrusion, which is only available in the proposed cycle direction. Trials of the inverse direction lead to an ejection from the binding pocket as well, but did not show any significant drug movement toward the exit gate. The path from the binding pocket to the exit gate was closed, because the conformational changes of the subdomains framing this so-called exit channel did not lead to an opening of it along this direction of the cycle.

Authors: Schulz, R.; Vargiu, A.V.; Collu, F.; Kleinekathöfer, U.; and Ruggerone, P.


3.1.1. Introduction

The acquisition of multidrug resistance (MDR) by bacteria, both in hospitals and in the community, has become one of the most serious impediments to improved healthcare [4, 10–14]. Unfortunately, MDR is not restricted to antimicrobials, being common to antimalarials, herbicides, and anticancer agents as well [13]. A key role in MDR is played by efflux pumps, which feature some characteristics with respect to common membrane transport systems [6, 15]. Indeed, while the latter typically are highly specific for their substrates, efflux pumps possess a broad specificity for a wide range of chemically unrelated molecules and drugs [4, 11, 13, 46, 95].

MDR is of particular concern in Gram-negative bacteria, since this class includes several human pathogens, e.g., Pseudomonas aeruginosa and Enterobacter aerogenes [5, 11, 13, 96]. Genetic and biochemical data [5, 40, 97–100] have shown that the major efflux systems in these bacteria constitute a tripartite complex spanning the periplasmic space across both the inner and the outer membrane [101, 102]. These efflux systems contain an inner-membrane transporter of the Resistance-Nodulation-Division (RND) superfamily [14, 18, 24] and extrude a large variety of toxic compounds, including novel experimental antimicrobials [103]. In E. coli, the system is composed of the outer membrane protein TolC [29], the periplasmic membrane fusion protein AcrA [31, 32], and the inner-membrane cation-drug antiporter AcrB [104].

The active part of the efflux pump - AcrB - (see Fig. 3.1A) is primarily responsible for the uptake and selectivity of the substrate as well as for the energy transduction. Its structure has first been solved as a symmetric homotrimer [26]. Three main domains have been identified in each monomer: the transmembrane domain embedded in the inner membrane, which provides the energy using the transmembrane proton gradient; the pore/pumping domain in the periplasm, that is supposed to contain the gates from which
substrate uptake and extrusion toward TolC occur \cite{34,45}; and the TolC docking domain, containing a central funnel and presumably being in contact with TolC. More recently, AcrB has been crystallized as an asymmetric trimer with \cite{25} and without \cite{27,28} a substrate bound in the interior of the protein. Each monomer was found in a different conformation (hereafter Loose, Tight, and Open, or L, T, and O, respectively, following Ref. \cite{27}). In the structure of Murakami et al. \cite{25}, the drugs doxorubicin and minocycline were found inside a binding pocket in the T monomer. The three conformations of AcrB have been interpreted as states of a transport cycle, schematically represented in Fig. 3.1B, which occurs via a three-step functional (not physical) rotation \cite{25,27}. Following the hypotheses formulated in Refs. \cite{34,15,96}, the substrate should enter the pore domain of the transporter via the L and/or T monomer, either from open clefts.
in the periplasm or through grooves between helices at the interface between pore and transmembrane domain. Then, the substrate should accommodate into a binding pocket when the monomer assumes the T conformation and move out toward the TolC docking domain upon a subsequent change to the conformation O. The proposed mechanism is primarily based on the available crystal structures, and has been confirmed only indirectly [49, 51]. Recently, it has been subject of a critical review [34]. In particular, it is not known how conformational changes of the protein cause the extrusion of the drug, and to what extent diffusion is important in the process.

A direct and atomistic-level description of the interplay between structure and dynamics of the conformational changes might render the proposed pictures of the function less speculative and allow to acquire knowledge on the structure-dynamics-function relationship. Such insights will be of support for the analysis of the huge amount of genetic, mutagenetic, and other biochemical data on RND transporters [4, 11, 13, 14, 46, 55, 96]. Additionally, they will constitute valuable information for a structure-based design of efficient antibiotics and inhibitors, by identifying possible target and key domains in the different steps of the extrusion process. In this respect, molecular dynamics (MD) simulations already pinpointed important atomic-level details of the functioning of TolC [57, 61] and MexA [58], a homologue of AcrA. Despite the importance of AcrB, no computational studies of this transporter have been reported so far.

Here, we present a first attempt to address the relationship between functional rotation and extrusion of substrates: starting from the structural information available on the complex of AcrB with doxorubicin [25], we simulated the proposed final extrusion step of the functional rotation. This was done via targeted molecular dynamics (TMD) simulations [105], which enables to mimic the conformational changes of the protein without explicitly considering the proton transfer and the related energy transduction, which would require the introduction of quantum-mechanical calculations. TMD has been successfully applied to study conformational changes in large systems as F1ATPase [106] and MurD [107], and it has recently been shown to provide reliable transition paths as compared to other methods used to sample conformations of proteins [108]. Note that in this work we are not investigating the issue of substrate specificity of AcrB, which would require additional compounds to be considered. In the following we show that doxorubicin leaves the binding pocket upon induction of functional rotation, although its total extrusion into the TolC docking domain is not observed. The main aspects as well as the possible limiting factors of the process are discussed. Furthermore, we investigate the presence of a peristaltic-like mechanism and characterize its underlying atomic rearrangements.
3.1.2. Results

The quite modest success in the fight against bacteria endowed with the MDR machinery [10] is partly due to the lack of knowledge of the connections between structural and dynamical features, which determine the function of the efflux systems at the molecular level. In an effort to shed some light thereon, we investigated to what extent drug motion is related to the suggested extrusion step of the functional rotation. To this end, we started from the T state for the occupied monomer, wherein the anthracycline antibiotic doxorubicin was cocrystallized previously [25]. Then, the transition T→O, i.e., the final extrusion step of the functional rotation proposed in Refs. [25,27,28], is enforced via TMD. Note that such a short notation highlights the conformational changes of the T monomer, but all other monomers are also forced toward their corresponding conformations, i.e., T→O is equivalent to L-T-O→T-O-L (see Fig. 3.1B). The main results presented below stem from simulations in which all heavy atoms of the protein were steered. To estimate possible side effects of the TMD approach, we further discuss the results from TMD simulations where only Crets atoms were steered (see Table S1.1 for a list of all the performed TMD simulations). During all the TMD simulations, doxorubicin is not steered by the external bias that drives the conformational changes of the protein, but it is free to move. In the T conformation, the suggested entrance from the periplasm to the binding pocket is opened and the exit gate toward the central funnel is closed (see Fig. 3.1B). In the O conformation, the closing/opening configuration is toggled, hence the substrate should be able to move out of AcrB toward TolC [28].

Displacement of the Drug  Initially, doxorubicin is found within a binding pocket which is located between β-sheets of the subdomains PC1 and PN2 [25] and formed by the residues F136, Q176, F610, F615, F617, and F628 [55], as shown in the inset of Fig. 3.1A. During the transition, these subdomains undergo conformational changes, thereby displacing the drug. In general, the whole binding region, which contains the described binding pocket has a quite large internal volume, probably with more than one binding site [13].

Fig. 3.2 displays the calculated distance between the centers of mass (COMs) of the binding pocket and doxorubicin, dBP, as a function of the simulation time, along with the values of the interaction energy. At the end of the TMD simulations, the RMSD of the protein with respect to the target (T-O-L) is ~0.6 Å (Fig. S1.1), indicating that the transition has been accomplished. Furthermore, the substrate has moved away from the binding pocket by a total distance of ~8 Å toward the gate to the central funnel, formed by the residues Q124, Q125, and Y758 [28]. As shown in the inset of Fig. 3.2, the interaction energy increases significantly as the transition proceeds, thereby denoting an
Figure 3.2: Movement of doxorubicin within the protein. Plot of the distance between the COMs of doxorubicin and the binding pocket, $d_{\text{DOX-BP}}$ (solid curve), as a function of the simulation time during the T→O transition. The curve represents an average over simulations with and different initial velocities (see Methods). Also shown are the same distances during ensuing standard MD simulations with restraints on $C_\alpha$ atoms of the protein (dot-dashed), and without restraints (dashed). The interaction energy is given in the inset.

The initial and final positions of the drug of one of the TMD simulations are shown in Fig. 3.3 as well as the structural changes of the binding pocket and of the gate to the central funnel. The displacement of doxorubicin toward the gate is confirmed by the profile of the distance between the COMs of the three residues forming the gate and that of doxorubicin (data not shown). During the T→O transition, this distance decreases by $\sim 7$ Å, which is a clear indication of the movement of the drug along the path that was identified by Sennhauser et al. 28. Note that the magnitude of displacement is essentially independent of the initial orientation of doxorubicin within the binding pocket (see Fig. S1.2). Additionally, the obtained displacements are almost insensitive to randomly reinitializing the initial velocities of all atoms or to extending the simulation time to 5 and 10 ns (Fig. S1.3). In these longer simulations, the major movement of doxorubicin occurs at a different relative time, with respect to the total TMD simulation time, although the final displacements are very similar to those seen in the shorter runs. This indicates a minor dependence of our general results on the simulation time and fortifies the reliability of our calculations.

To assess the stability of doxorubicin in the final position at the end of the TMD runs, we performed two sets of four standard MD simulations starting from the final TMD
**Figure 3.3**: Drug displacement and conformational changes of the protein along the transition T→O. A) Movement of the drug from the binding pocket toward the gate (yellow and red represent the initial and final configuration, respectively). It can be seen from the comparison of the two conformations that residues F136 and F628 in the bottom of the binding pocket and F615 and Q176 in the upper part, are mostly involved in the squeezing of the binding pocket; interestingly residue F610, which mutagenesis experiments have shown to be important for the activity of the pump, seems not directly involved in pumping the substrate out of the binding pocket; B) Porcupine plot of the conformational changes of the subdomains in the pore domain (arrows represent the displacement, in Å, from the initial position of the Cα atoms of each amino acid, colored according to the color scale bar). Important subdomains are highlighted: PN2 (transparent blue), PN1 (transparent green), and PC2 (transparent red). As PC1 lacks major changes, it is omitted for clarity. The movement of the drug toward the central funnel is also shown (color code as in panel A); C) Opening of the gate (residues Q124, Q125, and Y758, yellow space-filling representation) and of the BP-Gate path (backbone representation, formed by residues 48-50, 85-89, 126, 163, 177-181, 273-276, and 767-772 of the occupied monomer T - gray - as well as residues 67-70 and 113-117 of the neighboring monomer O - orange, see also text). Transparent and solid representations refer to initial and final states of the T→O step, respectively. This allows to better appreciate the opening of a path towards the TolC docking domain. For the sake of clarity when compared to the other panels, the final conformations of doxorubicin (transparent stick representation) and the binding pocket (black-and-white sticks) are also shown.

configurations (see Figs. 3.2 and S1.4). In the first case, we restrained the Cα atoms of the protein, in order to keep the backbone in the final T-O-L conformation. In the second one, we left the system completely unrestrained and free to relax. In half of the unrestrained simulations, the drug moves further away from the binding pocket by 2 Å in the direction of the gate; in all the remaining runs (2 unrestrained and 4 restrained), it oscillates around its final TMD position. Nevertheless, doxorubicin does not move
back toward the binding pocket in any of these simulations.

From the visual inspection of the final position of doxorubicin as shown in Figs. 3.3A and B, it is clear that the drug did not enter the central funnel of the TolC docking domain during the $T \rightarrow O$ transition. Certainly, the real time scale of the process is out of reach by the computational tools used in the present work, and such a limitation might be a reason for the absence of the complete extrusion in our simulations. Indeed, diffusion could play a relevant role in driving out the drug from AcrB, but this process would occur on a time scale hardly approachable by our protocol. Apart from methodological issues, additional factors have been suggested to be necessary for the full extrusion of the substrate. For example, the necessity of cooperativity effects associated with the binding of a second substrate (absent in our simulations) to a neighboring monomer has been invoked to interpret kinetic data [56]. Furthermore, a more involved pattern of configurations assumed by the monomers and connected to the extrusion process has also been inferred from the analysis of crystallographic structures [34, 44]. A further possible reason might be the absence of the membrane fusion protein AcrA. Its contribution to the functionality of the efflux system seems to go beyond a simple structural linker between TolC and AcrB. Surely, a deeper understanding of this interplay will benefit from the simulations of the entire system exploiting the model recently proposed by Symmons et al. [32].

Note that upon induction of the conformational transition $T \rightarrow O$, the subdomain PC2 moves inward to close the entrance and is followed by PN1 which opens the exit [27] (Fig. 3.3B). The distance between the COMs of PN2 and PC1 declines (see Fig. S1.5) accompanied by a rotation of the two subdomains, thereby resulting in a shrinkage of the binding pocket. Thus, the motions of the subdomains appear to be the first requirement for the squeezing of the drug out of the binding pocket. However, the largest displacements between the COMs of the subdomains do occur in the first half of the TMD simulation, while most of the drug displacement is seen in the second half (see Figs. 3.2, S1.5, and Video S1.10). Interestingly, the RMSD of the residues of the binding pocket from the target does not drop much until almost half of the TMD simulation is over. Then, it starts to decrease in correlation with the movement of the drug (Fig. S1.1), indicating that more specific and local conformational changes are involved in the unbinding of doxorubicin. Thus, our attention focused on the action of specific groups of residues.

**Evidence for Peristaltic Motion** A peristaltic pumping was proposed as the extrusion mechanism by Pos and coworkers in 2006 [27]. To identify possible fingerprints of the peristaltic action and correlations between motions of residues and drug displacement,
Figure 3.4: Sequence of the peristaltic squeezing. Upper panel: Snapshots extracted from the MD simulations represent configurations just after the squeezing of certain residue pairs in comparison to the initial state (snapshot labeled “Zero” on the left, solid stick representation). For each snapshot only those residues which are mainly involved in the squeezing step are highlighted. Lower panel: Minimum distances between selected pairs of residues of the binding pocket, $d_{\text{min}}$, as a function of the simulation time. To better identify different behaviors, we reported running averages (step length equal to 10) of the raw data. The tiny-dotted line represents the distance between the COMs of the binding pocket and doxorubicin (as plotted in Fig. 3.2).

we compared the latter with the evolution of the minimum distances, $d_{\text{min}}$, between selected couples of residues in the binding pocket. In particular, we selected those pairs of residues whose distances decline predominantly during the T$\rightarrow$O transition, namely F136-F615, F136-F617, F136-F628, and Q176-F615. In Fig. 3.4, the evolution of their
average minimum distances over 5 TMD simulations (lower panel) is shown together with three representative configurations associated therewith (upper panel). Interestingly, the changes in the distances among the selected residues occur in a step-wise fashion, with residue pairs at the “bottom” of the binding pocket closing first, and those at the “top” last, producing a zipper-like motion.

The first reduction of $d_{\text{min}}$ affects the pair F136-F628, but the substrate essentially keeps its position in the binding pocket. Successively, the residues F136 and F617 approach each other, and $d_{\text{DOX-BP}}$ starts to increase (I in lower panel of Fig. 3.4). The configurations assumed by the three residues are displayed in snapshot I of the upper panel in the same figure. At about one third of the TMD simulation time, the distance of F136-F615 starts to decrease, and this reduction correlates with a large movement of the drug (II in Fig. 3.4). Note that this is the largest reduction (∼5 Å) in the distance among the observed pairs of residues. At approximately the same time, the squeezing motion between Q176 and F615 takes place, which is also related to a substantial displacement of the drug (III in Fig. 3.4). These two amino acids happen to act as a clamp for the planar rings of doxorubicin. The phenyl ring of F615 is atop of one of the rings belonging to the drug and the Q176 amide is on the other side. While the drug is moving toward the gate, the connection between the ring of F615 and the drug is changing from one planar ring of the drug to the next one in a stepwise fashion. Once the residues F615 and Q176 squeeze the substrate out of the binding pocket and thereby close the return path of doxorubicin (see snapshot III in the upper panel of Fig. 3.4), the drug is able to move further away from the binding pocket.

Concerning each individual TMD simulation, the connection between the zipper-like closure of the binding pocket and the drug displacement can be seen in 4 out of 5 different 1-ns-long simulations (Figs. S1.6A, C, D, and E). A three-step mechanism can be roughly recognized in the graphs, with a sequential closure of the pairs from the innermost (F136-F628) to the outermost one (Q176-F615). The remaining run of this set (Fig. S1.6B) could be viewed as a borderline case in which the last step is very short. Despite this, the closure of the binding pocket maintains a basically sequential character, where the outermost pairs (Q176-F615 and F136-F615) close after the innermost ones (F136-F617 and F136-F628). Additionally, three longer TMD simulations (two of 5 ns and one of 10 ns) were performed to assess the dependence of our results from the simulation time (Figs. S1.6F, G, and H, respectively).

As expected, the molecular details of the process (final displacement, side chain conformation and dynamics) are slightly sensitive to the simulation protocol (see also Fig. S1.3). Nevertheless, a sequential closure of the binding pocket is still detectable in all panels of Fig. S1.6. In addition to the four out of the five 1-ns-long simulations mentioned above, in one out of the two 5-ns-long ones evidences of a three-step mechanism
are recognizable (see caption of Fig. S1.6 for an extended discussion). Unfortunately, a meaningful statistics, needed for a thorough discussion of the possible limits inherently present in the TMD protocol, is out of range for these longer trajectories.

It is worthwhile to point out that the results of recent mutagenesis experiments have evidenced a significant impact of the mutation F610A on the minimum inhibitory concentration of doxorubicin, while other mutations, including those of the phenylalanines 136, 178, 615, 617, and 628 to alanine, showed smaller effects. According to our simulations, F610 is not prominently involved in the zipper-like action, but might act as binding partner when doxorubicin enters the pore domain, and/or might close the escape from the binding pocket toward the periplasm. Upon mutation to alanine, these actions might not be efficient anymore. Additional studies are required to gain insight into the effect of the F610A mutation.

**Further Essentials of Drug Displacement**

To enhance the understanding of the results presented above, we investigated the dynamical coupling between squeezing motions of the binding pocket and other specific residues located beyond it. In particular, we chose those residues lining the path from the binding pocket toward the exit gate. This path (hereafter called BP-Gate path and sketched in Fig. 3.3C) is formed, with reference to the initial conformation, by the residues 48-50, 85-89, 126, 163, 177-181, 273-276, and 767-772 of the occupied monomer T as well as residues 67-70 and 113-117 of the neighboring monomer O. A series of TMD simulations have been performed in which we kept the BP-Gate path of the T monomer unsteered and forced only the rest of the protein, thereby applying the same bias as in the previous TMD simulations. According to our results, doxorubicin leaves the binding pocket also in these simulations, but the overall displacement is smaller by $\sim 3\,\text{Å}$ with respect to the one shown in Fig. 3.2. Indeed, the BP-Gate path remains too narrow for doxorubicin to leave the binding region and to move toward the exit gate. Furthermore, the drug is tilted by $\sim 90^\circ$ with respect to the final position in Fig. 3.3A (see Fig. S1.7), which also hinders further motion toward the gate. This result emphasizes the importance of a concerted opening of the BP-Gate path in addition to the zipper-like closure of the binding pocket.

Since the position and the orientation of amino acids seem to be important for the displacement of the drug from the binding pocket, we further extended our set of simulations to shed more light on this aspect. In the dynamics described so far, all non-hydrogen atoms have been targeted, which corresponds to a forced movement of the side chains during the TMD simulation. To analyze the importance of these movements for the drug displacement in comparison to the influence of the backbone/subdomain, we performed a series of TMD simulations where only the C$_\alpha$ atoms were targeted. This also allowed
to test the influence of the biasing force on our results, as a large fraction of the protein is now free to move. We observed a significant displacement of the drug during this set of TMD simulations (see Fig. S1.8), in qualitative agreement with those obtained by targeting all heavy atoms, hinting at the importance of subdomain motions for the displacement of the substrate (see Fig. 3.5). On average, the displacement is reduced by $\sim 3.5$ Å with respect to the one reported in Fig. 3.2. Clearly, the number of possible paths explored by the drug is expected to increase when targeting only the C$_\alpha$ atoms, due to the larger flexibility of the protein. Consistently, a displacement comparable to that shown in Fig. 3.2 is observed only in 3 out of 10 TMD simulations (data not shown). In addition, in some of the 7 remaining runs, doxorubicin does not move straight toward the exit gate, but also turns slightly aside where the interior along the BP-Gate path leaves space to roam. These results indicate that the arrangement of the side chains is able to significantly influence the maximal displacement of the drug.

Additionally, we performed a simulation with a lower force constant (see Tab. S1.1). The aim was to obtain an indication of the minimal force required to accomplish the conformational changes in the protein, especially along the BP-Gate path. It turns out that, during the entire TMD simulation, the distance of doxorubicin from the binding pocket is lower by a couple of Å with respect to that in previous simulations with larger force constants (Fig. S1.9). This is related to a larger RMSD of the binding pocket from the target along the simulation (inset in Fig. S1.9), which, although very small, has an important effect on the displacement of the drug. In combination with the results of the TMD simulations where only C$_\alpha$ atoms have been targeted, these findings highlight the importance of individual residues including their side-chain conformations for the displacement and subsequent extrusion of doxorubicin.
Reversing the Direction of the Cycle

Analyzing the asymmetric crystal structures of AcrB [25,27,28], it is reasonable to suppose that drugs exit the transporter from the monomer in the O conformation. Therefore, we have considered the direct transition T→O at first. However, the possibility of a functional transition from T to O via the L conformation cannot be ruled out a priori. Thus, we carried out simulations for the two steps of the reverse cycle direction, i.e., T→L and L→O (Fig. 3.6). The investigation is important for two reasons. Firstly, the direction T→O has been suggested to be the functional one from analyses of structural data [25, 27, 28], but it lacks a direct proof. Secondly, the comparison between the two directions should give a better picture of the conformational changes and drug-amino-acid interactions which are mainly involved in the displacement of the drug.

Interestingly, the specific movements of PN2 and PC1, which have been described above as responsible for the shrinkage of the binding pocket during T→O, can also be observed during T→L (data not shown). Indeed, the substrate tends to leave the binding pocket in both cycle directions. However, in contrast to T→O, the drug displacement never exceeds ~2 Å for T→L, hence doxorubicin does not approach the exit gate. This can be attributed to the quite large internal volume of the binding region [13].
substrates may exploit their flexibility and change their orientation. Importantly, the drug does not move remarkably during the second step \( \text{L} \rightarrow \text{O} \) of the reverse cycle as well. Again, this points to the need of a concerted closure of the binding pocket and widening of the channel toward the exit gate.

### 3.1.3. Discussion

The molecular dynamics underlying the functioning of many active transporters, which include efflux transporters of the RND family, are not fully understood yet. Although the increasing number of crystal structures permits us to have a closer look at the atomic details of the structure, the dynamical aspects are not caught, and only hypotheses can be advanced concerning the functional mechanisms. MD simulations with atomistic detail are an appropriate tool to investigate structure-function-dynamics relationship in these systems. In this work, we performed TMD simulations to investigate the relations between supposed functional motions in AcrB \cite{25, 27, 28} and the extrusion of the antibiotic doxorubicin without explicitly considering the energy supply associated with the proton gradient across the inner membrane.

Our results show a detachment of the drug from its initial binding pocket within the T monomer. Moreover, during the \( \text{T} \rightarrow \text{O} \) step of the functional rotation, doxorubicin travels by \( \sim 8 \, \text{Å} \) and approaches the gate to the central funnel. Importantly, this movement is believed to be part of the suggested extrusion process in AcrB. Our data also support the proposed peristaltic pumping mechanism, and highlight the atomistic dynamics at its basis. In particular, there is evidence to suggest a zipper-like squeezing of the binding pocket, which leads to an unbinding of the substrate along the \( \text{T} \rightarrow \text{O} \) direction of the cycle. The closing of the binding pocket is initially caused by the movements of adjacent subdomains, whereas the rearrangements of individual residues lining the binding pocket strongly influence the detachment of doxorubicin in the end. The molecular details of the extrusion process depend slightly on the TMD simulation protocol (simulation length, targeted atoms), but the main features are robust against these changes.

While investigating the feasibility of the cycle in the reverse direction \( \text{T} \rightarrow \text{L} \rightarrow \text{O} \), additional simulations have shown a similar squeezing of the binding pocket during the \( \text{T} \rightarrow \text{L} \) transition. However, no substantial movement of the drug toward the gate has been seen. This could mainly be due to the lack of concerted widening of the BP-Gate path and the exit gate. Moreover, even if such movements do occur during the subsequent \( \text{L} \rightarrow \text{O} \) step, they are not coupled to squeezing of the binding pocket, and do not cause any significant movement of the substrate. Altogether, these results strongly points at \( \text{L} \rightarrow \text{T} \rightarrow \text{O} \) as the legit direction of the functional rotation.

Although a substantial movement of the substrate was seen in our TMD simulations
of the T→O transition, the drug never reached the central funnel of the TolC docking domain, which is a necessary step to achieve the full extrusion of the drug out of AcrB. One possibility to explain this is that further movement of the drug might just be directed diffusion within a confined geometry occurring on a time scale much larger than that captured in the simulations. In addition, the motion of the drug might further be enhanced by attractive interactions between the substrate and residues around the gate or even TolC, or by the presence of other substrates. Finally, the influence of the neighboring monomers as well as the other proteins constituting the efflux pump have to be understood. In the long run, it would be very important to model the whole tripartite efflux pump, i.e., AcrB together with TolC and AcrA. This could complete the picture of the protein-protein interactions involved and their cooperative effects on the drug extrusion. Nonetheless, using the present results it should be possible to obtain a better understanding of the structure-function relationship in RND transporters and its connection to dynamical aspects. Finally, molecular insights on the efflux mechanism in AcrB might be of help for the research on human RND transporter, e.g., the Niemann-Pick C1 disease protein and the hedgehog receptor Patched [34].

3.1.4. Methods

System Setup For our simulation setup, the crystal structure from Ref. 28 was chosen. After addition of hydrogen atoms, a restrained structural optimization was performed. The structure of doxorubicin was taken from Ref. 25 and placed into the system in the same relative position within the binding pocket as original. The latter structure was not used since several loop residues (499 to 512) of the pore domain were not resolved, and the resolution was lower with respect to that in Ref. 28. The combination of a crystal structure with a substrate from another structure was possible since the binding pocket of the protein accommodates the drug very well; indeed, doxorubicin keeps its position during the equilibration. Moreover, the RMSD between the C_α of the structures from Refs. 25, 28 is less than 1 Å. After the placement of the drug, a second relaxation was performed. The protein-substrate complex was inserted into a pre-equilibrated POPE lipid bilayer, which is parallel to the x-y plane, and solvated in TIP3P water with a physiological KCl concentration of 0.1 M. At the end of the buildup phase, all lipid and water atoms which overlapped with the protein were artificially removed; the total number of atoms of the system is 451,962. This setup leads to a periodic box size of \( \sim 172 \times 169 \times 184 \, \text{Å}^3 \).  

43
**Force Fields Parameters**  The AMBER force field parm99 \cite{109} was used for the protein, the TIP3P parameters for water \cite{87}, and Aqvist’s parameters for the ions \cite{110}. For doxorubicin several parameters were taken from the GAFF force field \cite{111} while the missing ones were generated using modules of the AMBER package \cite{112}. In particular, atomic restrained electrostatic potential (RESP) charges were derived using antechamber, after a structural optimization performed with Gaussian03 \cite{113}. The GAFF parameters for the POPE lipids were generated following the protocol in Ref. \cite{114}.

**Simulation Protocol**  The unbiased and the targeted MD simulations were both performed with the program NAMD 2.7b1 \cite{64}. After an initial energy minimization, the system was gradually heated up to 600 K and finally quenched to 310 K. All these simulations were performed in the presence of restraints on the phospholipids and the heavy atoms of the protein. A time step of 1 fs was used for the integration of equations of motion. Furthermore, periodic boundary conditions were employed, and electrostatic interactions were treated using the particle-mesh Ewald (PME) method, with a real space cutoff of 12 Å and a grid spacing of 1 Å per grid point in each dimension. The van-der-Waals energies were calculated using a smooth cutoff (switching radius 10 Å, cutoff radius 12 Å). Furthermore, the simulations were performed in the NpT ensemble and the temperature was kept at 310 K by applying Langevin forces to all heavy atoms with the Langevin damping constant set to 5 ps$^{-1}$. The pressure was kept at 1.013 bar using the Nosé-Hoover Langevin piston pressure control.

The functional rotation was simulated by means of TMD \cite{105} (built-in module of NAMD) which allows to induce conformation changes between two known states. To prevent any hindrance on the T monomer by the neighboring ones, we also steered those toward their next state. Note that the TMD algorithm has recently been demonstrated to produce reliable transition paths as compared to other methods \cite{108}. In this respect, to assess the influence of the biasing force on the dynamics of the system, we performed a series of TMD simulations using different values for the force constant per atom ($k = 2, 3, \text{and} 4$ kcal·mol$^{-1}$·Å$^{-2}$) and the simulation time (1, 5, and 10 ns). The results discussed in the main paper refer to simulations of 1 ns with $k = 3$ kcal·mol$^{-1}$·Å$^{-2}$, which are consistent with the literature \cite{106,107,115}. All TMD simulations performed are detailed in Tab. \ref{tab:1.1} together with the comparison among distances between COMs of doxorubicin and the binding pocket and the final positions of the drug (Fig. \ref{fig:1.2}). The setup, the analyses as well as the atomic-level figures, were performed using VMD \cite{77}. 

<table>
<thead>
<tr>
<th>Hyperlink</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>parm99</td>
<td>\cite{109}</td>
</tr>
<tr>
<td>TIP3P</td>
<td>\cite{87}</td>
</tr>
<tr>
<td>Aqvist’s</td>
<td>\cite{110}</td>
</tr>
<tr>
<td>GAFF</td>
<td>\cite{111}</td>
</tr>
<tr>
<td>antechamber</td>
<td>\cite{113}</td>
</tr>
<tr>
<td>POPE</td>
<td>\cite{114}</td>
</tr>
<tr>
<td>NAMD 2.7b1</td>
<td>\cite{64}</td>
</tr>
<tr>
<td>TMD</td>
<td>\cite{105}</td>
</tr>
<tr>
<td>VMD</td>
<td>\cite{77}</td>
</tr>
</tbody>
</table>
3.1.5. Acknowledgments

We are grateful to K. M. Pos (Frankfurt-Main), M. Winterhalter (Bremen), and M. Cec- carelli (Cagliari) for stimulating and fruitful discussions during the course of this study. We acknowledge computer time from DEISA, CINECA (Bologna), Cybersar (Cagliari), CASPUR (Rome), and CLAMV (Jacobs University Bremen).

3.1.6. Author Contributions

Conceived and designed the experiments: RS AVV UK PR. Performed the experiments: RS AVV FC. Analyzed the data: RS AVV UK PR. Contributed reagents/materials/analysis tools: RS AVV. Wrote the paper: RS AVV UK PR. Contributed materials: FC.
Summary of “Role of Water during the Extrusion of Substrates by the Efflux Transporter AcrB”

The following manuscript continues the train of thought of the previous one and illuminates more details of the extrusion mechanism including water dynamics and cooperative effects. Therefore, the vicinity of the drug and the entire extrusion pathway, from the entrance along the binding pocket toward the exit gate, have been investigated for more general quantities. This was rendered possible as the data analyzed was derived from more than ten equivalent simulations. During this study, the data was separated into two clusters which allowed to distinguish necessary and optional properties of the extrusion process. Moreover, the interaction energies between the drug (the same as in the previous work) and its environment was examined as well. During the transition which is supposed to lead to the extrusion of substrates from the interior of the transporter, a slight but distinct stream of water has been found. Its direction was found to follow the extrusion pathway from the entrance close to the outer leaflet of the inner membrane along the binding pocket toward the exit gate to the TolC-docking domain. Furthermore, the electrostatic interaction increased along the path from the binding pocket to the exit gate. On the one hand, this was partially due to the fact that the substrate left a hydrophobic region during the simulations. On the other hand, the hydrophilic exit channel has drawn the drug toward the exit. Interestingly, this effect was weakened by the surrounding water molecules which screened the protein electrostatics and thereby disfavored a strong binding of the drug to any region of the exit channel. In this study, this particular effect has been described as the lubrication effect of water.

3.2. Journal of Physical Chemistry B (2011): Role of Water during the Extrusion of Substrates by the Efflux Transporter AcrB [65]

Authors: Schulz, R.; Vargiu, A.V.; Ruggerone, P.; and Kleinekathöfer, U.

Reproduced with permission from Journal of Physical Chemistry B, Schulz, R.; Vargiu, A.V.; Ruggerone, P.; and Kleinekathöfer, U., “Role of Water during the Extrusion of Substrates by the Efflux Transporter AcrB” (doi:10.1021/jp200996x), Copyright 2011 American Chemical Society.
3.2.1. Introduction

The transport of substrates through the cell envelope is an extremely important but also troubling subject for bacteria. On the one hand, nutrients are required as energy source for the cell cycle and building blocks for the construction of proteins, lipids, etc. On the other hand, toxic compounds such as bile acids in the human intestines, biocides, and antibiotics have to be kept at nontoxic concentrations in the cell [116, 117]. Therefore, nature developed several mechanisms to regulate the influx and efflux of toxic molecules [3, 5, 12–14, 47, 118]. Among these mechanisms, the active extrusion of antibiotics is a subject closely related to bacterial resistance, which has become a serious issue both in the clinics and in the community in the last decades [119, 120]. The situation is of particular concern for Gram-negative bacteria (such as *E. coli*, *P. aeruginosa*, *K. pneumoniae*, *A. baumannii*), among which major nosocomial pathogens exist worldwide [120, 122]. In these bacteria, efflux pumps containing transporters of the resistance-nodulation-cell-division (RND) family play a major role in the intrinsic and acquired resistance to multiple classes of structurally distinct antimicrobials, including clinically relevant ones [121, 123]. Indeed, it has been shown that the lack of functional components of these efflux pumps renders bacteria susceptible to antibiotics, while their overexpression has been associated with the appearance of multidrug resistance (MDR) in *E. coli*, *S. enterica*, and more pathogenic isolates [68, 124, 125].

One paradigmatic efflux pump is the AcrAB-TolC multidrug efflux pump of *E. coli* which is, to a large extent, responsible for antibiotic resistance in this bacterium [24, 32, 126, 127]. As such, it is a crucial target in the development of new antibiotics and inhibitors. The AcrAB-TolC system has been studied most intensively among the efflux pumps [44, 46]; it is a tripartite complex composed of an active RND transporter embedded in the inner membrane, AcrB [26, 60], an outer membrane exit duct in the outer membrane, TolC [29, 61], and a periplasmic membrane-fusion protein, AcrA [31]. In this system, the passageway of substrates is supposed to be from the periplasm and outer leaflet of the inner membrane through the so-called pore domain of AcrB toward TolC and out of the cell. The latter protein protrudes about 100 Å into the periplasmic space. In this setup, AcrA is believed to stabilize the complex and possibly transmit conformational information as well as time-dependent changes during the functional rotation of AcrB toward TolC [44].

In the system of interest, AcrB pumps toxic substrates from the periplasm into the TolC channel and thereby out of the cell (for a better overview of the entire extrusion pathway from the cytoplasm to the extracellular space including schematic figures, see refs [44, 46, 128]). Therefore, it is fueled by the proton gradient across the inner membrane. Recent structural studies revealed that AcrB assumes an asymmetric structure both in
the presence and in the absence of a substrate bound to a periplasmic pocket. Using minimum inhibitory concentration (MIC) assays, important residues have been identified within the binding pocket (BP) which are responsible for the proper work of the transporter. The substitutions of F136, Q176, F610, F615, F617, and F628 by the amino acid alanine affect the MICs of various substrates with different impact. This indicates a high flexibility of the BP in the interaction with its compounds. The effects of F610A, the substitution with the highest influence on the MIC of doxorubicin, have been addressed in a computational study as well. Moreover, the broad specificity of the BP has recently been confirmed by docking simulations.

On the basis of the asymmetric X-ray crystallography structures of AcrB, a three-step functional rotation hypothesis has been proposed. Toxic substrates are taken up from the periplasm between subdomains of the pore domain (Loose and/or Tight). Then, they are bound in the internal BP (Tight). At last, these molecules are extruded upon peristaltic conformational changes (Open). These steps are connected to the individual states of the individual monomer: open entrance, closed BP and exit in Loose; open BP in Tight; toward Open with an open exit, closed entrance and BP (see Fig. 3.7 for structure and functional rotation cycle). Here the states of the cycle are labeled as in ref 27, i.e., Loose (L), Tight (T), and Open (O) in comparison to Access, Binding, and Extrusion of ref 25, respectively.

The described conformational changes are initiated by the transmembrane proton gradient or more specifically by the protonation of one of the residues in the transmembrane proton-relay network. This leads to a conversion of chemical into mechanical energy which is supposed to be transduced toward the pore domain via the surrounding helices and, in particular, transmembrane helix 8 being elongated toward the subdomain PC2 in the pore domain. The key structural changes occur during the T→O transition (shown also in Fig. 3.8), which involves the closing of the entrance (in this article, the entrance is defined to be between the subdomains PC1 and PC2 as used in ref 28—similar to cleft in ref 25) and of the BP as well as the opening of the gate (represented by the amino acids Q124, Q125, and Y758 as in ref 28) to the funnel in the TolC-docking domain. This proposal is supported by recent disulfide cross-linking experiments and by experiments with covalently linked AcrB protomers, which proved the conformational interaction and interdependence of the monomers. More recently, fluorescence assay experiments strongly pointed to the existence of an extrusion path passing through the BP and the gate.

From a computational perspective, several simulations of the individual parts of the pump have been performed in the last years, i.e., for AcrB and TolC. Recently, all-atom targeted molecular dynamics (TMD) simulations on the AcrB transporter embedded in a lipid bilayer and in the presence of explicit solvent were
carried out by the present authors \cite{60} to mimic the functional rotation and connect it to drug translocation. Although the time scale of these TMD simulations was too short to observe complete extrusion in such a large system, a significant movement of the drug toward the gate to the central funnel was revealed. These computations support the earlier proposed peristaltic mechanism as well as the preferred direction of the pumping cycle, i.e., T→O. In general, TMD provides the opportunity to induce conformational changes between two known states, despite the limited simulation time available. This method has been used before in studies of different kinds of proteins such as F\textsubscript{1}-ATPase \cite{106} and MurD \cite{107}.

In this paper, we present results from a new set of TMD simulations of the transitions T→O and T→L of the proposed transport cycle (hereafter simply called TO and TL, respectively). Despite this short notation, which only describes the transition of the T monomer, the TMD forces are applied to all three monomers at the same time, i.e., T→O corresponds to L-T-O→T-O-L. In the current investigation, the simulation lengths have been prolonged by one order of magnitude with respect to those reported in ref \cite{60}. In comparison to the former study, we focus particularly on the role of water during the process by exploiting the full-atomic details of our simulations here.

There are several reports of a directed flow of water connected to the transport of substrates such as in cotransporters (see, for example, refs \cite{132}–\cite{137}), indicating the importance of joined water flux upon exchange of molecules with the cells’ environment. Although AcrB cannot be classified as a cotransporter, we observe that water closely follows doxorubicin along its path from the BP to the gate and expect this finding to persist during the entire pumping process along the extrusion path. In the next sections, we demonstrate that water has a significant role in the substrate extrusion. By almost totally enveloping the drug, water acts as lubricant and screens the electrostatic interactions between the drug and the channel toward the gate, which is mainly lined by hydrophilic amino acids. Furthermore, we analyze whether a distinctly directed flux of water occurs during the extrusion of a substrate. Finally, to quantitatively support the picture proposed above, we compared the dynamics of the drug-water interactions with those between the drug and the protein.

### 3.2.2. Methods

**System Setup** The equilibrated structure, comprising the AcrB-doxorubicin complex embedded in a POPE bilayer and 0.1 M KCl solution, was adopted from ref \cite{60}. The simulated system, shown in Fig. 3.7, contains \~450,000 atoms with a size of the simulation box of approximately 172×169×184 Å. For the protein, the AMBER force field parm99SB \cite{109,138} was employed together with the TIP3P parameters for the
Figure 3.7: Topside view of AcrB (monomers in yellow, brown, and green cartoon representation) in a lipid bilayer (gray surface) with doxorubicin (red) bound in the binding pocket. The water surrounding the drug is drawn as a light-blue surface. Note that the figure is tilted by 30° toward the viewing plane. Inset at the top: Zoom-in of the bound drug with surrounding water. Inset at the bottom: Scheme of the functional rotation L→T→O.

water 87 and Aqvist’s parameters for the ions 110. The POPE lipid parameters were computed using the protocol described in ref 114. Furthermore, the substrate molecule doxorubicin was structurally optimized first using Gaussian03 113. Its parameters are a combination of the GAFF force field 111 and newly computed ones using modules of the AMBER package 112. Especially, the program antechamber was used to obtain the atomic restrained electrostatic potential (RESP) charges. For all amino acids, their standard protonation states were considered, i.e., the states as for pH 7.
Simulation Protocol  The program NAMD 2.7 [64] was employed to perform the reported molecular dynamics simulations. Before performing any of the simulations described later on, the structural stability of the protein and the membrane was investigated in connection with ref [60] already. Therefore, the secondary structure was examined graphically along several trajectories, using the plugin “Timeline” of VMD [77]. This clearly showed that no problems with structural stability occurred. Starting from the equilibrated configuration as described in ref [60], production runs were performed using a 1 fs time step and periodic boundary conditions including the particle-mesh Ewald (PME) method for the electrostatics (cutoff distance 12 Å, switching distance 10 Å, grid spacing 1 Å). The temperature was kept at 310 K using a Langevin thermostat [139] (damping constant 5 ps\(^{-1}\)) and the pressure was kept at 1.013 bar via the Nosé-Hoover Langevin piston pressure control [140,141].

The functional rotation was simulated using the TMD method [105] as implemented in NAMD. In this procedure, the protein is steered from an initial to a target conformation. Though only one monomer of AcrB is analyzed below, all monomers were forced toward their subsequent state in the pumping cycle (see Fig. 3.7 for cycle description). This is done to avoid any intermonomer effects on the T monomer which is occupied by the substrate. In this study, forces have been applied only to the C\(_\alpha\) atoms of the protein. In a previous study [60], all heavy atoms were steered though some comparisons with steering only the C\(_\alpha\) atoms have been reported as well. In the latter case, the side chains have more freedom to move. Note that the TMD algorithm has recently been demonstrated to produce reliable transition paths as compared to other methods [108]. The TMD force constant per atom was set to \(k = 3\) kcal mol\(^{-1}\) Å\(^{-2}\) which did not lead to any artificially large values of the interaction energies or artificial states in the previous article as well [60]. Furthermore, the data analyses which are detailed in the Supporting Information as well as the creation of the AcrB figures have been performed using the program VMD [77].

3.2.3. Results

In this work, we performed 15 10-ns-long TMD simulations with the purpose of characterizing the translocation of doxorubicin along the TO step of the functional rotation cycle in detail with a special focus on the dynamics of the solvent. Moreover, we carried out 10 1-ns-long TMD simulations which semiquantitatively yield the same results as the former ones and are therefore not shown, but are used to connect to the short TL simulations at the end of this article. In agreement with our previous study [60], we observe a significant displacement of the drug as a consequence of the squeezing of the BP during TO. Importantly, the opening of the exit channel and of the gate [28] allows
Figure 3.8: Conformational changes (top panel, state T; lower panel, state O) in the pore domain of AcrB (looking from atop). The thin arrows emphasize the changes of subdomains PC2 and PN1 that mainly affect the putative extrusion pathway (green arrow). The drug (orange halo) is initially located in the BP, i.e., roughly in the middle of the upper graph. The BP is closing during this transition with the drug moving upward along the green path.

for a flow of water through these regions. However, the role of water in possibly lowering the steric and electrostatic hindrances associated with the passage of antibiotics has not been discussed in the literature before. Nonetheless, water may also act as a transporting fluid, but nothing is known about the direction of this flux of water, e.g.,
whether it is in the same direction as the substrate like in cotransporters \[132,133\] or in the opposite direction. To estimate the importance of the solvent in the extrusion process, the interactions between the drug and its environment, i.e., water and protein, are investigated thereafter. Finally, the results of the TO simulations are compared to another set of TMD simulations mimicking the opposite direction of the cycle (TL). We would like to emphasize that the statement of a closed entrance or a closed exit refers to being closed for a substrate molecule to pass. Even in these so-called closed states, water molecules can still flow through these channels. Furthermore, there are quite a few additional channels in the protein through which water can flow, but they are not accessible by molecules of the size of a substrate molecule.

**Conformational Changes and Drug Displacement** To transport substrate molecules, AcrB undergoes large conformational changes which are associated with the functional rotation. As part of this cycle, the putative extrusion pathway is particularly affected during the TO transition. These changes can be seen in Fig. 3.8 where the principal conformational changes of the pore domain are shown. The model pathway described in this figure is used in several sections of this article. In the following, we connect these conformational changes to the position of the drug as well as the structural and dynamical role of water in the process in more detail.

**Figure 3.9:** Plot of the drug-BP distance as a function of the relative simulation time. Results from all the simulations are shown as dots, and averages of clusters cl+ and cl- are represented by solid lines. For the definition of cl+ and cl-, see the main text.

Fig. 3.9 shows the temporal progress of the center-of-mass (COM) distances between
the drug and the BP extracted from 15 TO simulations. These results are in good agreement with simulations reported earlier \([60]\) using a very similar setup, namely only targeting the C\(_\alpha\) atoms as in the present investigation. The final values of the drug-BP distances vary significantly in this setup of the TO simulations (at its most advanced position in our simulations, the substrate has arrived in the neighborhood of the exit gate \((d_{\text{drug-BP}} \sim 15 \text{ Å})\), while the COM distances between BP and exit gate is 25 Å). Moreover, one has to keep in mind that the length of doxorubicin along its major axis passing through its structure of four rings is about 15 Å. Hence, describing the substrate as a point-like particle will certainly not suffice to represent all features of the extrusion process. Additionally, the data points are not connected per trajectory to focus on the global variations rather than on the individual progress. Note that the initial position of the drug is not necessarily in the COM of the BP (drug-BP distance = 0). This might be due to the polyspecificity of the transporter with possibly several binding sites in the BP for different substrates \([13, 59]\). All TO simulations started with the same initial structure, but the initial velocities of all atoms are different for each simulation. Furthermore, there are random forces from the pressure and temperature control methods \([139–141]\). Because of these random influences, different paths of the substrate extrusion have been sampled. To simplify the presentation of our results, the trajectories have been grouped into two clusters (see Fig. 3.9) distinguished by the drug displacement: one subset comprising those simulations with a final COM distance larger than 12 Å (denoted cl+ in this article), and the other one featuring less displacement of the drug (cl-). This clustering allows us to analyze the data more globally by differentiating the key elements of successful extrusion from the BP. The first cluster contains four and the second one eleven trajectories. Out of the latter one, one simulation behaves rather different from the rest, i.e., the actual distance to the entrance becomes even smaller than in the other simulations, and it has to be discussed separately.

**Substrate-Environment Interactions** To quantify the hydration of the substrate during the transition, those water molecules have been counted that are within a range of 3.5 Å. This ensures that the first and possibly second hydration shell are included. In Fig. 3.10, the number of waters surrounding the substrate is shown as a function of the drug-BP distance (see also Fig. S2.11; the simulation time is only considered indirectly in this so-called parametric plot). For lower distances, the number of waters increases in both clusters, as expected, since the drug is leaving the hydrophobic BP where its tail is in direct contact with phenylalanines. Out of the BP and along the channel toward the funnel in the TolC-docking domain, the ratio of hydrophilic over hydrophobic residues increases with the distance from the BP (see Fig. S2.12). In the following, we will return
Figure 3.10: Number of water molecules surrounding the substrate (dots) as function of drug displacement $d_{\text{drug-BP}}$ for the two clusters defined in the text: $cl^+$ and $cl^-$. The lines represent the running averages of the dots over 20 data points. The red line indicates the initial state of the drug.

to the importance of the electrostatic topology of the channel several times.

The hydration shows a monotonic nearly linear increase as a function of drug-BP distance; i.e., the further the substrate is apart from the BP, the more water molecules surround it. Nevertheless, one should be aware that using the drug-BP distance as the one-and-only parameter to estimate the progress of extrusion is just an approximation. The orientation of the drug during the extrusion process is very important as well (more the corresponding subsection), in particular in connection with the water dynamics. In Fig. 3.11, doxorubicin is shown with the surrounding waters at the beginning and at the end of a representative simulation. In the initial conformation (Fig. 3.11 upper panel), water molecules are mainly found around the more hydrophilic head of the substrate (see Fig. S2.14 for the number of water molecules separated into those close to the head and those close to the tail of the drug). At the end of the simulation, the drug is much more enveloped by waters (Fig. 3.11 lower panel), and only a few amino acids are in direct contact with the substrate, namely S48, T87, Q125, Q176, N274, and R620. Regarding the hydrogen bonds, there is a slight increase in the average number from three to five along the exit pathway (Fig. S2.13). Although the data show heavy fluctuations, there is a recognizable increase until $d_{\text{drug-BP}} \sim 10 \, \text{Å}$; afterward, it remains almost constant. To get a better understanding of the underlying interactions, the corresponding energies will be analyzed below. Among the surrounding waters, some establish H-bond contacts,
eventually connecting the substrate to the exit channel. The latter ones might have an important role in the translocation of the substrate since these lower the mechanical stress and increase the screening of the electrostatic interactions. The number of water
bridges instead drops on average from five to three with likewise heavy fluctuations (data not shown).

Figure 3.12: Partial interaction of the drug with its environment versus the COM distance. While the upper panel shows the interaction energies of substrate with the entire environment, the two lower panels only show the interactions with protein and water, respectively. The vertical lines are drawn at the initial state to easily compare these data.

As stated above, the environment around the substrate changes significantly along the translocation path from the BP toward the gate. Especially the hydration of the drug (shown in Fig. 3.10) and the ratio of hydrophilic residues along the channel increase during the extrusion (see Fig. S2.12). These features suggest that the electrostatic
interactions may have an important role in assisting the unbinding of the substrate from the BP and the translocation process in general. In Fig. 3.12, the interaction energies between the substrate and its environment (entire environment, only protein, and only water; all capped at 12 Å distance from the drug) are shown. Interestingly, the total nonbonded energy related to the entire environment (upper panel) decreases steadily until a distance of 10 Å, after which it remains almost constant. This behavior is almost equal to the change of the electrostatic interaction which confirms the importance of such contributions in the first part of the unbinding. On the other hand, the VDW interactions slightly increase along the entire path sampled in the present simulations. Along this path, the interaction energy with water is stronger than with the protein (compare lower two panels of Fig. 3.12). As expected, this contribution is almost completely electrostatic in nature as can be seen in the lower panel of Fig. 3.12. Considering the total nonbonded interaction energy with water, a linear decrease can be observed starting from a drug-BP distance of \( \sim 3 \) Å. This initial drop can be traced back to the increasing number of waters surrounding the drug. At around 10 Å, the decrease ceases and the interaction remains rather constant apart from the ongoing fluctuations.

In contrast to this, the larger interaction of the substrate with the protein is due to VDW forces that rise while \( d_{\text{drug-BP}} \) augments (Fig. 3.12 middle panel). Especially close to the BP, the presence of water prevents even more unfavorable drug-protein VDW interactions associated with the squeezing of BP residues. In the same region, the electrostatic interaction is rather weak mainly due to the substrate tail residing in the hydrophobic binding pocket. Along the extrusion path, this interaction rises because of the increasingly hydrophilic environment (compare to Fig. S2.12). Interestingly, the average electrostatic interaction remains rather constant after \( \sim 10 \) Å. Additionally, the VDW interaction between protein and drug weakens steadily when the drug has left the BP and entered the channel toward the gate. This reduction might be due to the increasing size of the channel associated with the TO transition and the waters enveloping the drug (see also Fig. 3.14).

To summarize this part, the graphs in Fig. 3.12 shed some light on the possible role of water in facilitating the diffusion of the substrate along the exit channel. Without the presence of water, the minimum of the nonbonded interaction energy (protein-drug) would be at \( \sim 10 \) Å of the drug-BP distance (center panel) if we would neglect the values at very small distances. Afterward, the interaction energy rises steadily to a maximum at \( \sim 12 \) Å, i.e., when the substrate faces the exit gate. Looking at the lower panel, there is a decrease of the negative interaction energy of about the same amount from 8 to 14 Å. In conjunction with the discussed increase, the combination of the two terms flattens the profile of the total interaction energy (upper panel). Furthermore, taking into account that the entropic contribution to the process is likely to be very small,
these results point at a relevant role of water in assisting the diffusion of substrates along the channel. A similar feature can be seen when the drug is very close to the BP preventing too strong binding. Such a strong binding would also hinder an effective transport. A word of caution needs to be made concerning these short distances since this is a one-dimensional description of a movement in three-dimensional space. The drug is not just moving on a straight line and for these short distances does not move further into the BP but moves to a position alongside it. An increasing electrostatic attraction along a reaction pathway is, in a way, similar to the findings of a study on an ADP/ATP carrier [142]. Therein, a substrate was initially placed close to the entrance of the protein. During these unbiased simulations, the substrate was electrostatically pulled inward by hydrophilic residues toward the center of the transporter. In the present simulations, the electrostatics seem to play an important role as well, though the conformational changes of the protein are of course the most crucial factors. Despite grouping trajectories into $c_{l^+}$ and $c_{l^-}$, we have mainly discussed the first one until now. Nonetheless, another question is still open: What structural and energetic factors exist that determine these different outcomes and what is the role of the water therein? These questions will be addressed in the next subsection.

**Water Flux within the Channel** To describe the solvent dynamics within the putative extrusion path, the average directions of motion of selected waters have been calculated. Therefore, a grid path has been superimposed on the structure of the protein (see Fig. 3.8 and Fig. 3.15). The grid ranges from the entrance between subdomains PC1 and PC2 through the interior of the pore domain along the BP and via the exit channel toward the gate (for details, please refer to the Supporting Information). Along this path, the average direction of motion has been calculated by considering those waters around a grid point at any time (see Fig. 3.13).

It can be seen in Fig. 3.13 that the overall water flow is directed from the BP and toward the gate. The length of an arrow describes the average magnitude of directed movement of one water molecule per 50 ps in the vicinity of a particular grid point. The shortest arrows are found in regions of larger conformational changes, e.g., closure of the entrance. Furthermore, some of the water molecules originally in the entrance region are squeezed out into the periplasm upon its closing along the TO transition. Note that these directions only indicate the average direction of movement over the entire trajectory and not the starting or end point of actual movements. While, for example, the overall direction in the exit channel is toward the exit gate, the water molecules arriving there in the final state do come from different directions. This emphasizes the existence of many small holes in the protein through which water but no drug molecules
Figure 3.13: Average direction of movement of waters near grid points along the model extrusion path. The length of an arrow is the average magnitude of directed movement of one water molecule per 50 ps in the vicinity of a particular grid point. In addition, the paths of two representative TO simulations are depicted in which the straight elements of the path indicate the drug movement within the same time step, i.e., 50 ps.

can pass. Additionally, the arrows at \( y \sim 0 \, \text{Å} \) indicate an additional pathway of water molecules toward the interdomain region. In this region, the protein contains several loops leading to holes which allow the passage of water. An important point to note from Fig. 3.13 is that the movement of the water molecules is much faster than that of the substrate indicated by the length of path travelled within 50 ps.

To analyze in more detail why the substrate seems to have no big effect on the flux of waters, the number of water molecules along the grid path is shown in Fig. 3.14. This number of water molecules strongly depends on the simulation time due to the TMD setup. The steering forces change the protein conformation and with this the water accessible volume along the extrusion pathway. In the T state of AcrB the entrance is open and the exit closed. Consequently, the entrance can accommodate many water molecules and the exit relatively few. In contrast in the O state, few water molecules are found in the closed entrance region, while more are present in the now open exit channel space. Hence, the number of water molecules can be used as a rough estimate of the channel volume along the extrusion pathway. Using CAVER [89], the radius of the exit channel has been estimated to be less than 2 Å in the closed L and T states as well as approximately 6-7 Å in the O conformation. This has to be compared to a rough approximation of 5-6 Å for the radius of doxorubicin. Additionally, it can be seen in the
region where the substrate stays during all simulations (near the BP and along the exit channel), that the number of waters does not change much compared to its neighboring grid points. While this is an average over all simulations, it still implies an additional disturbance. Moreover, the BP does not get filled with waters during the transition, but shrinks at the same time as the drug leaves the BP. While this shrinkage is induced by the TMD method here, it might be *in vivo* a simply favorable transition of the BP to remain closed during the absence of a substrate. Unfortunately, to test this is beyond the scope and capabilities of the present work.

![Figure 3.14: Average number of waters near grid points along the model extrusion path, calculated for different simulation times. The average is computed considering all 10-ns-TO simulations.](image)

**Orientation of the Substrate during Extrusion** During the translocation, the substrate changes its orientation which is partially due to the varying environment. To clarify this issue, we have to abandon the point-like description of the drug as used before. A principal axis has been associated with the previously introduced plane of doxorubicin to estimate these reorientations, defined by its four rather planar rings. The geometric setup is required to analyze the orientation of the drug with respect to the protein using the definition of spherical coordinates. The orientation of the substrate is given in terms of the spherical angles $\phi$ and $\theta$ of its principal axis (see Fig. [S2.15]). The angle $\phi$ characterizes the rotation in the $x$-$y$ plane which is parallel to the membrane. Accordingly, the angle $\theta$ describes the rotation perpendicular to this plane corresponding to a direction downward from the $z$ axis.
Figure 3.15: Orientation of the drug in the AcrB pore domain during the simulations TO (the vertical line represents the initial state). Both clusters are distinguished here and a running average has been calculated for each. Possible substrate orientations are shown in the inset. The red molecule corresponds to the initial conformation, the blue one to an orientation of $\phi \sim 40^\circ$, and the green one to an orientation of $\phi \sim 150^\circ$ toward the end of a $cl+$ trajectory.

The variation of both angles, $\phi$ and $\theta$, with the movement of the drug is depicted in Fig. 3.15. These graphs include running averages for the two clusters $cl+$ and $cl-$ which have been introduced before. For $cl-$, $\phi$ varies by as much as $150^\circ$ along the trajectories. In contrast to this, the changes in $\phi$ for cluster $cl+$ are less pronounced and show a nearly linear decrease. The largest difference between the two clusters is visible for
drug-BP distances between 7 and 9 Å. In this region, the trajectories of \( cl^- \) rotate much more than those of \( cl^+ \) which might actually be the reason for the hindrance in the displacement along the exit channel. An unfavorable orientation might hinder an easy movement toward the exit. The angle \( \theta \) increases almost linearly for both clusters but with large fluctuations. For \( cl^- \), the maximum change of the running average is about 35° since the drug movement is smaller compared to \( cl^+ \) with about 40° change. Also in the behavior of this angle, there is a large spread of values for drug-BP distances between 7 and 9 Å which might be related to the fact that the drug spends more time in this region, i.e., has more time to explore different orientations. Despite the chance of exploring a larger variety of orientations described by \( \theta \), the drug is not able to advance in the trajectories of \( cl^- \), pointing out a probable larger influence of \( \phi \) compared to \( \theta \) in determining the movement of the drug.

**Reverse Functional Rotation: Qualitative Behavior of TL Simulations** To understand the effect of the changing environment on the drug as well as the internal waters, ten simulations have been performed in the reverse direction, inducing the conformational changes of the protein according to TL. Because it was already shown in ref.\textsuperscript{60} that TL does not allow extrusion, the simulation length was kept at 1 ns and is only compared to the more general aspects of the TO simulations, which hold for both 1 ns and 10 ns time scales. In comparison to TO, the only major change in the pore domain is the closure of the BP during this transition. This means that the entrance stays open all the time, and the exit gate remains closed. Therefore, the drug is not able to travel far along the exit channel.

Fig. S2.16 shows the distribution of the ten TL trajectories. It is surprising that one of the trajectories shows a rather large movement of the drug to a distance from the BP of 12 Å. This means that the drug travels as far as the upper limit of the TO \( cl^- \) simulations. However, a closer look reveals that the results are rather different. In Fig. S2.17, the orientation of the drug during these TL simulations is shown. Obviously, the orientation is rather different from those of the TO direction for larger drug-BP distances (compare to Fig. 3.15). Actually, the drug rotates in a direction more or less perpendicular to the extrusion pathway. This can also be seen while investigating the number of protein contacts. In Fig. S2.18, the ratio between hydrophilic and hydrophobic residues indicates an increasing number of hydrophilic residues compared to the TO results. Actually, the conformational changes going from T to O or from T to L are drastically different. This can also nicely be seen looking at the number of water molecules along the extrusion pathway shown in Fig. 3.16. During the TL transition, the entrance stays open while the exit keeps its closed conformation. This is in clear contrast to the TO transition in
which the entrance closes and the exit opens. Not only is the exit closed, but also the extrusion pathway close to the exit is much narrower than in the TO direction. This can be inferred from the number of waters close to the exit in Fig. 3.16. The area in which the water population actually really changes in the TL direction is the binding pocket. There the number of water molecules reduces quite drastically, which means that the binding pocket is closing during the transition. As mentioned already, the statement that the entrance or exit is closed refers to being closed for a substrate molecule, and water molecules can still move through these and other holes in the protein. Thereby, leaving waters from one region are often followed by others entering through these narrow passageways (compare to Fig. 3.13).

Figure 3.16: Number of waters along the extrusion path for TL simulations.

3.2.4. Conclusion

Though it might not be the primary function of many transporters, lots of them additionally transport water at the same time as their substrates [133, 134]. In the present study, we have used 15 TMD simulations with a length of 10 ns each to demonstrate that water plays an important role also in the efflux transporter AcrB. Several features extracted from our simulations support such involvement. First of all, there is a clear directional water flow from the BP toward the exit. The average velocity of the water was shown to be larger than that of the substrate. Moreover, there are also water molecules following the substrate; i.e., these are more or less pushing the drug molecule. This
strongly indicates that the substrate is actually moving in a stream of water which also helps to determine the directionality of the substrate movement, possibly leading to a directed diffusion in vivo. Second and maybe even more importantly, the water flattens the electrostatic interaction energy profile between the drug and its protein environment, and it lowers the contribution of the VDW interactions. Along the exit channel, the interaction energy between substrate and protein alone shows a certain pattern, which might get the substrate stuck at a certain position though it is supposed to leave through the gate. The influence of water counteracts this effect and the electrostatic interaction energy profile is flattened, thereby allowing a rather unrestricted passage. Additionally, the water competes for hydrogen bonding which would otherwise bind the substrates to the protein more strongly. In this respect, water really acts as a lubricant for the substrate even in this quite narrow channel.

This shielding of the water molecules might also be related to the polyspecificity of AcrB, an important characteristic of these transporters [59]. If the transporter would be very specific with respect to a certain substrate, the protein structure would have to show some specific corresponding feature to be able to screen and to react to the particular property the substrate possesses. Substrates of AcrB need to have a rather hydrophobic part to be bound in the hydrophobic BP due to stacking interactions between the rings of the substrate and the protein’s side chains [60]. Choe et al. [132] already described that water molecules assist the unbinding of substrate molecules from the direct interactions with the protein residues. Without the water molecules, the interactions between substrates and the AcrB transporter would become less likely to break upon the transition T→O.

Our results are, of course, limited by the feasible time scale of molecular dynamics simulations as well as the technique used here. In vivo, the extrusion lasts micro- to milliseconds [39], which is much longer than the simulated time here. However, the meaning of time in TMD simulations is limited since the transitions are enforced which would otherwise take much longer. Furthermore, the in vivo time scale describes the turnover rate per efflux pump, which includes the time traveling into the binding pocket of an AcrB monomer as well as through the ∼140 Å long channel of TolC. Therefore, a possible directed diffusion of the substrate can only be imagined from the farthest position observed in this study into the funnel toward TolC. Another limitation of TMD is that conformational changes are induced simultaneously in all the selected regions of the protein. Although the general dynamics are Newtonian with TMD “kicks” being applied in addition, this limitation does not allow us to study any kind of sequence or dependence of these induced changes. Finally, as this study focused on AcrB only, the interaction between the protein partners of the efflux pump cannot be examined but only inferred from existing experimental data such as the hypothetical model of ref [32].
In conclusion, the present TMD simulations show a correlated movement of the substrate and the water molecules in the protein which is significantly directed toward the exit channel and gate. This direction is defined by the conformational changes of the pore domain, which are induced \textit{in vivo} by the converted mechanical energy from the transmembrane gradient and mimicked in the present study by means of the TMD method. Specific aspects of the interactions and functionality of the transporter can be better understood at a molecular level, i.e., the requirement of water being present to allow an unrestrained flow of substrates through the protein and out of the cell. This knowledge helps to comprehend the full functional cycle of the transporter by adding another important ingredient for the successful pumping: the lubrication effect of water.

### 3.2.5. Acknowledgments

We are grateful to Niraj Modi (Bremen), M. Ceccarelli, and A. Kumar (Cagliari) for stimulating and fruitful discussions during the course of this study. We acknowledge computer time from DEISA, CINECA (Bologna), Cybersar (Cagliari), CASPUR (Rome), and CLAMV (Jacobs University Bremen). A. V. Vargiu acknowledges financial support from “Regione Autonoma della Sardegna” through a Research Fellow on fundings of the Project “PO Sardegna FSE 2007-2013, L.R.7/2007 Promozione della ricerca scientifica e dell’innovazione tecnologica in Sardegna”.
3.3. Partial Targeted MD Simulations of AcrB

After familiarizing with the chances and drawbacks of the TMD method \[105\], it was applied more selectively in another study. Herein, we intended to improve the understanding of the interactions between the domains and monomers, e.g., the transduction of mechanical energy from the transmembrane domain into the pore domain. Therefore, the same approach was applied as in Refs. \[60,65\] i.e., targeted molecular dynamics (TMD) simulations, but with different selections of steered protein domains. Thereby, the focus was on identifying regions which move in correlation with the steered sections. Furthermore, we investigated the possible interdependence of the monomers by estimating the hindrance or support of neighboring monomers to each other. During this work, the temporal limitations of MD simulations became more obvious.

3.3.1. Results

Table 3.1: Description of the TMD simulation setups used in this study, defined by selections of residues which are steered by the TMD method. Some selections are a sum or difference of two selections, described accordingly. Note that each selection only refers to the C\(_\alpha\) atoms of the amino acids.

<table>
<thead>
<tr>
<th>Index</th>
<th>TMD selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>freeDyn</td>
<td>none</td>
</tr>
<tr>
<td>tmDom</td>
<td>transmembrane domain</td>
</tr>
<tr>
<td>NBmons</td>
<td>neighboring monomers</td>
</tr>
<tr>
<td>freePP</td>
<td>tmDom and NBmons</td>
</tr>
<tr>
<td>fullTMD</td>
<td>entire protein</td>
</tr>
</tbody>
</table>

As the transporter AcrB can be clearly divided into three domains, i.e., transmembrane domain, pore domain, and TolC-docking domain as well as three monomers, the selections of steered residues have been defined analogously. Tab. 3.1 contains all crucial information of these different settings. In general, only the C\(_\alpha\) atoms have been steered in the selected regions. All TMD simulations are compared to an unbiased control simulation freeDyn of 200 ns simulation length and the results of TMD simulations with all C\(_\alpha\) atoms steered - fullTMD (the same setup as in Sec. 3.2). To render these selections more comprehensible, the particular regions of interest are highlighted in Fig. 3.17.

In general, the focus is on the same monomer as in the previous Secs. 3.1 and 3.2 because the same system setup was used here as well. Furthermore, this is in agreement with the importance of the energy-dependent transition T\(\rightarrow\)O (see Secs. 3.1.1 or 3.2.1 for clarification of the abbreviations) which this monomer experiences in our setup. Initially, the steering of the transmembrane domain is investigated to find possible clues considering the transduction of mechanical energy from the interior of the transmembrane domain toward transmembrane helix 8 (abbreviated as tmH8) and toward the
pore domain. This helix was suggested to be responsible for the closure of the entrance in the pore domain\textsuperscript{[25, 27, 28]}. Therefore, its conformational changes and effects on the pore domain are examined more closely. Afterwards, the neighbors are first steered exclusively and then in combination with the previously steered transmembrane domain. This is done in order to clarify the influence of the neighboring monomers on the transition of the observed, unbiased monomer domains.

### Possible Transduction of Mechanical Energy

The hypothesis described by crystallographers\textsuperscript{[25, 27, 28]} states that, while cycling through the obtained states of the individual monomers, the converted chemical energy from the proton gradient has to be transferred into the pore domain where the actual uptake-binding-extrusion is taking place. Ref.\textsuperscript{[28]} presented a schematic which highlights the essentials including the conformational changes of \textit{tmH8}. Although the pore and transmembrane domains are linked at four spots by peptide bonds, the interface between both mainly contains unstructured loops, rendering the connection quite flexible. The helix \textit{tmH8} is one of two extended
helices (the other being helix number 2) which protrude farther toward the pore domain. Looking at the conformational differences between the states of the cycle (derived from asymmetric structures), \( tmH8 \) moves more prominently while \( tmH2 \) keeps its position more or less. Hence, helix number 8 was taken into consideration for more detailed TMD selections.

Figure 3.18: Movement of the subdomains: The arrows of the porcupine plot indicate the linear deviation per \( C_\alpha \) atom between initial and final state of the transition \( T \to O \). The longer rod-like arrows describe the principal axis of each subdomain in the direction of the porcupine plot. The position of the drug molecule in the center is taken from a fullTMD simulation.

As stated above, the main focus is set on the pore domain of the occupied monomer. Hence, the dynamics of its four subdomains are examined more closely in the following. The results of a fullTMD simulation are displayed in Fig. 3.18, thereby indicating the expected maximum extent of conformational changes of these subdomains. To simplify the observations, the information of each subdomain’s movement is mainly reduced to the translational motion of the center of mass (COM) and the rotational movement of the major principal axis (PA). For the latter, the same spherical coordinate system will be used as already described in Sec. 3.2 and depicted in Fig. S2.15.

To estimate the influence of the time scale on the overall results, several simulations have been performed between one and fifty nanoseconds. As shown in Fig. 3.19, the COM displacement of the subdomains does fluctuate, but without a distinct direction of transition for three out of four of them. Only PC2 shows a significant decline of more than 1 Å for these \( tmDom \) simulations, with the 1-ns simulation being the lower limit. This brings into question the applicability of this measure. Hence, another observable has been introduced - a major principal axis per subdomain (see Fig. 3.18). This principal axis has been calculated the same way as in Sec. 3.2, defining the major orientation of a group of atoms, in this case a subdomain. The resulting directional vector has been used to calculate angles or spherical coordinates relative to the coordinate system of the
Figure 3.19: Translational movement of the four subdomains during simulations of \textit{tmDom} in the same order as Fig. 3.18. The described spherical angles are of the same type as described in Fig. S2.15. All data have been smoothened by a running average applied on at least 20 data points.

protein. Thereby, the positional transition of the subdomains can be described more accurately by incorporating their rotational movement.

While \textit{fullTMD} follows a distinct path guided by the TMD forces, the \textit{tmDom} simulations show much more angular flexibility in general (Fig. 3.20). Interestingly, the longest simulation of 50 ns does not necessarily lead to results closest to \textit{fullTMD}. This hints at a possible lack of other, intermonomeric contributions, facilitating the transition between the states of the functional rotation cycle. Apart from PN2, only the $\phi$ angle of PC2 shows a distinct direction of transition in all trajectories which is probably due to extended \textit{tmH8} steered by TMD and and being close to this subdomain. The transition
of PN2 is weaker than of PC2, but seems to be more proportional to the time scale than the motion of PC2. The comparison of \textit{tmDom} and \textit{fullTMD} simulations indicates a correlation between expel of the substrate from the binding pocket and the degree of transition (compared to \textit{fullTMD}) of PN2. As described in previous studies already, PC1 is the least moving part showing a slight translational shift but almost no rotation compared to the other subdomains. Nevertheless, some crucial elements of the examined cycle transition are missing, such as for PN1 which is supposed to open the exit channel and gate.

To quantify the importance of the helix \textit{tmH8}, it has been virtually separated into its “head”, comprising its 6 uppermost residues toward the pore domain, and the lower part. This separation has been used to setup additional TMD simulations with the \textit{tmDom}
selection altered by deselecting the head or the entire helix. During these simulations, the measured angles of PC2 did not change significantly. Hence, the obtained results indicated that the conformational change of this helix, especially its “head” are crucial for the transition of PC2, which is limited by the simulation length (data not shown). After a sufficient long TMD simulation without steering the “head” of tmH8 or the entire helix, the transition will probably occur as expected from the crystal structure, but this time scale was not accessible for this thesis.

Despite the less pronounced transitions of some of the subdomains compared to the reference trajectory of fullTMD, the major conformational changes, being the subdomains PC2 and PN2, have been observed to a certain extent. While for PC2, a distinct COM displacement and rotation of the $\theta$ angle has been measured, PN2 did show rotational movement for both angles. Note that the latter subdomain did not show any significant COM deviation during fullTMD, but the $\phi$ angle of PC2 did change by 20° in fullTMD compared to 10° in tmDom independent to the simulation time length. Furthermore, the data of the $\theta$ angle of PC2 did also not correlate with the time scale. Additionally, the transition of PN1 did not occur at all during tmDom, neglecting the minor change of its $\phi$ angle. Hence, the collection of selectively steered residues is going to be extended hereafter.

Influence of the Neighbors  To estimate the influence of the neighboring monomers on the conformational state of the monomer in focus, we compared the previous simulations of tmDom to setups where either the entire neighboring monomers have been selected for TMD (NBmons) or another one which combines the selections of transmembrane domain and neighbors, called freePP hereafter. This last setup only allows the periplasmic domain of the occupied monomer to move unbiased by the TMD method. Due to the lack of induced changes in the transmembrane domain during the NBmons simulations, the results of this setup have only been used to assist the discussion of freePP.

Because the COM measurement did not yield any representative results (Fig. 3.19), the examination is focused at the angular movements here. Fig. 3.21 displays these rotational changes per subdomain. The first, obvious aspect of these results is that they are closer to the reference data from fullTMD simulations. While in tmDom only the $\phi$ angle of PC2 changed along the proposed transition path (Fig. 3.20), the longer simulations here (5 & 50 ns) do follow quite closely the fullTMD trajectory. This is not the case for all the presented angle data as can be seen for PN2. The trajectory of this subdomain follows less closely the reference compared to tmDom (Fig. 3.20). The results for this might be related to the position of the drug in the pore domain, which, on average, does not leave the binding pocket as far as in the tmDom simulations. Although this
is surprising because a greater fraction of the protein is biased using TMD, it can help to comprehend the flexibility of this subdomain and the reason for the closing of the binding pocket without a substrate being present. Interestingly, the asymmetric crystal structures without a substrate did not show any significant conformational difference to the cocrystal considering the subdomains surrounding the binding pocket, i.e., PC1 and PN2. Nevertheless, there is no information of possible triggers – such as transmembrane helices – which would open or close the binding pocket. Furthermore, an open, unoccupied binding pocket would be energetically rather unfavorable, which was already stated in Sec. 3.2. An unexpected “improvement” over tmDom was the transition of PN1. For this subdomain, at least the $\theta$ angle goes along with the fullTMD data although the total deviation of the fullTMD is only 7°. This again points at the limited time scale of
the performed simulations.

Figure 3.22: Comparison of different TMD selection using the RMSD to measure the degree of accomplished transition to the next state of the function rotation cycle.

At last, the conformational transitions of both major setups, \textit{tmDom} and \textit{freePP}, are compared more quantitatively using the RMSD measure (see Eq. M.5). In Fig. 3.22, these are examined with two setups used as references, namely the 200-ns-long unbiased simulation \textit{freeDyn} and a sample trajectory of \textit{fullTMD}. While the general idea of longer simulation time and “the more parts you steer, the closer the results to \textit{fullTMD}” still seem to be true for most of the cases, some trajectories of Fig. 3.22 deviate from this expectation. For instance, the RMSD of PN2 does not seem to depend on the TMD selection neither strongly on the time scale. In fact, the 50-ns simulation of \textit{freePP} shows a slight increase of RMSD at the end of the trajectory (at \(\sim 80\%\)). Moreover, subdomain PC2 shows a greater decline for \textit{freeDyn} than for \textit{tmDom}. This was not entirely surprising, because unbiased MD simulations which were performed at the beginning of this thesis already depicted the asymmetric state of the protein as
extreme case. In general, it seems that PN2 tends to assume a position in the middle between both the closed and open states considering the entrance. Only the 50-ns freePP simulation follows the same decline as the fullTMD. This strongly points to the necessity of favorable intermonomeric interactions.

While tmDom already showed unbiased, partial transitions of PC2 and PN2, the results of freePP enhance these by adding a rotational movement of PN1. This TMD setup also hinted at the necessary intermonomeric interactions and cycle state dependencies described in Ref. [34]. Therein, Seeger et al. proposed a more elaborate scheme of the functional rotation cycle. This scheme also included certain restrictions of the states of the neighboring monomers. Interestingly, the transition of PN2 in freePP did not always lead to the same final φ and θ angles as it was shown in Fig. 3.20. This issue is still not fully understood, but points toward a possible correlation with the position of the drug, which was already assumed in Sec. 3.2.

**Correlated Movements** To investigate the intra- and intermonomeric interactions, correlation matrices have been calculated using the program g_covar from the Gromacs Package [143], which computes the Pearson correlation of a set of atoms, in this case all Cα atoms of the protein. Ref. [144] describes this approach as inapplicable to study three-dimensional protein systems, because the Pearson correlation does only consider colinearly correlated movements of two atoms. Hence, more elaborated correlations cannot be estimated using this method. Therefore, Lange et al. [144] developed a new method which they called “generalized correlation” and which is supposed to be able to cover these as well. In the following, both methods will be used to describe potential intra- and intermonomeric interactions.

Fig. 3.23 displays a set of Pearson correlation matrices calculated from simulations of fullTMD, tmDom, and freeDyn with the two groups of subdomains PN and PC highlighted. Note that only the residues of the occupied monomer are considered. To enhance the comprehension of the more important interactions, all correlation coefficients whose absolute values are lower than 0.5 have been neglected in this graph. What can be easily seen is the decreasing amount of data points from fullTMD via tmDom toward freeDyn. Unfortunately, no significant correlation between steered and unsteered regions can be seen here. Hence, this correlation matrix is compared to the results from the generalized correlation [144] in Fig. 3.24.

Examining the data shown in Fig. 3.24, a strong difference of the average value can be observed. Furthermore, the general correlation only distinguishes strong or weak correlation, while the Pearson correlation also considers negative values representing anti-correlation. Nevertheless, this information is not necessarily required here, although
some data points in Fig. 3.23 already pointed to anti-correlated motions. While the patterns in the upper triangle can easily be linked to the structure of AcrB, the stripes in the lower triangle are not that simple to be structurally connected. Additionally, attempts to subtract two related matrices did succeed for the Pearson correlation (data not shown), but failed for the generalized correlation due to strongly varying average values of these data from different simulations. In Ref. 146, a possible solution has been described by subtracting an average correlation over interatomic distances from all the matrix elements. Hence, a combination of both approaches seems to be a more suitable
3.3.2. Discussion

While the TMD method [105] was applied on all residues of the protein in Secs. 3.1 and [3.2] it is used more selectively on specific segments of AcrB in the present section. This allowed to search for “domino effects”, in which one part is pushed and another neighboring one reacts thereon. To a certain extent, this goal was accomplished in this work by combining and comparing several selections. Although the time scale is limited by the MD simulation technique, the comparison of several simulation lengths helped to identify these subdomain transitions which are limited by the simulation time. While the movement of PC2 induced by \( \text{tmH8} \) was already expected before the study, the conformational changes of PN2 were surprisingly unrelated to the actual TMD selection. This domain was stronger influenced by the time scale and the position of the drug. Moreover, PN1 seems to require a defined interaction with the neighboring monomer. In general, the majority of subdomain movement was rotational and not translational. Additionally, the initial state of PC2 which was derived from the crystal structure is obviously an extreme case, because in the unbiased MD simulation, used as control,
the RMSD of PC2 from the O state was reduced by almost 50 %. The application of two different correlation methods did not directly yield any new information. Only by combining them with the results obtained with other methods, the data seem to help visualizing the correlation during certain transitions. Nevertheless, this selective version of the TMD method offers new possibilities to study protein transitions, especially if the general direction of energy transduction is known already which helps to define advantageous selections of steered residues.
3.4. Effects of the F610A Mutation on Substrate Extrusion in the AcrB Transporter

An additional study along this train of thought has been performed in the group of Professor Ruggerone in collaboration with our group. Therein, the focus was on the binding pocket and a particular mutation - F610A. This mutation was found to affect the minimum inhibitory concentration (MIC) values for various antibiotics. Interestingly, doxorubicin did not seem to strongly interact with this particular residue in previous computational studies. Hence, the results using the previous approach, i.e., targeted MD simulations, have been compared to docking calculations. In the case of the mutant, the drug was bound much stronger in the binding pocket than in the wildtype during these simulations. In some cases, the substrate left the binding pocket, but in another direction than observed before - more toward the center of the pore domain or the entrance between subdomains PC1 and PC2. This study indicates that a decreased binding affinity might not be the only reason for reduced MIC values.

In this work, wildtype simulations similar to those described in Secs. 3.1 and 3.2 are compared to two setups of a mutated transporter as well as to docking calculations. To ensure that the mutant was set up correctly, the mutation was first introduced into the equilibrated wildtype structure and briefly re-equilibrated again and then used for TMD simulations as performed in the previously mentioned sections. The obtained data have been compared to a new setup of the mutant protein which was prepared from scratch. Both procedures yielded very similar results which have been compared to the docking afterwards. These docking calculations have been carried out using the ATTRACT package of Ref. 147. This tool allowed for the flexibility of both ligand and receptor, which enhances the reliability of the results.

As can be seen in Fig. 3.25A, the mutant structure allowed the drug to slide deeper into the binding pocket where it is packed more closely by hydrophobic residues, e.g., phenylalanines. Applying the TMD method as in Sec. 3.1 (steering all heavy atoms), the substrate left the binding pocket, but did not enter the exit channel. In this setup, the molecule moved sideways to the direction of extrusion, toward the center of the pore domain or the entrance (called cleft in this manuscript). Fig. 3.25B displays the verification of the initial position of the drug (before the TMD simulations) in the binding pocket. The slight deviation between both molecules might simply be due to a certain degree of flexibility of the binding pocket to accommodate substrates. The outcome of this study was that the decrease in MIC upon mutating F610A is probably due to a stronger binding in the binding pocket which might interfere with the functional rotation.
cycle by not allowing the drug to easily leave this region. Thereby, the exit channel might already close when the drug is finally able to depart, although the issue of the limited simulation time scale mentioned several times in this thesis still holds.

Figure 3.25: (A) Comparison of the drug position resulting from unbiased simulations of wildtype and F610A mutant. The drug molecules of the wildtype and mutant simulations are colored magenta and cyan, respectively; the same for the amino-acid side chains of the binding pocket red and blue stick representations. The grey secondary structure representation in the background resembles the surrounding monomer. Residue 610 is drawn as thicker sticks. (B) Two configurations of the substrate are superimposed as found after unbiased MD simulations (blue) and via docking calculations (green). The grey transparent sticks resemble the residues’ side chains of the binding pocket. Figure reproduced with permission from JACS Communication, in press. Unpublished work copyright 2011 American Chemical Society.
While the emergence of multidrug resistance becomes a serious impediment to improve healthcare, the development of new antibiotics is barely able to keep up. Furthermore, the understanding of the underlying mechanisms advances only slowly, both with respect to the molecular effects of antibiotics on their target sites as well as the different resistance strategies. One of the major topics in current research of antibiotic resistance is the mechanism of active transport systems of bacteria, so-called efflux pumps, which pump various types of toxic compounds such as antibiotic molecules out of the bacterial cell. During the last years, a special interest arose in E. coli’s major efflux pump AcrAB-TolC which has already been extensively investigated in experiments. Due to the limitations of the laboratory equipment, information with molecular details has been difficult to obtain. Therefore, computer simulations of such protein systems can help to bridge this gap.

To understand the functioning of E. coli’s major efflux pump on an atomistic level, its protein components have been studied individually in this thesis. This was an indispensable step before trying to perform investigations of the entire complex. Otherwise, the dynamics of each individual protein cannot be distinguished from the cooperative interactions. The presented results help to improve the understanding of the actual efflux transport by pumping toxic compounds from the periplasm into the extracellular space. Furthermore, these data provide assistance to future systematic studies of the complex to overcome the multidrug resistance of pathogenic Gram-negative bacteria. At the beginning of this project, little atomistic and dynamic information was available. Hence, this work required to find suitable approaches to study the components of this efflux pump. While molecular dynamics simulation has been used as the general tool for most of the studies herein, this technique was often enhanced, by adding auxiliary
forces for example, dependent on the kind of protein to be studied.

In the first part, there is the outer-membrane channel TolC, that is more static apart from the periplasmic tip which opens upon docking to an inner-membrane transporter. To understand the process of the possibly iris-like opening of this aperture, mutants have been generated to break the network of salt bridges and hydrogen bonds which usually keep the protein closed while no transporter is present. This idea was described by X-ray crystallographers before who presented partially open crystal structures. During this investigation, cations hindered the opening of the aperture due to reconstitution of the previously mentioned salt bridges. By applying an electric field, these ions have been urged to leave the periplasmic tip and not get stuck in this region again.

In the second part, the inner-membrane transporter AcrB undergoes major conformational changes to accomplish the uptake, binding, and extrusion of antibiotics toward TolC and further into the extracellular space. In this thesis, several successive studies revealed important information on the extrusion cycle which was previously described as a functional rotation. Therefore, the structures of an asymmetric trimer and a cocrystal accommodating a molecule of doxorubicin have been used to generate a system setup to mimic the conformational changes in the transition cycle, especially the extrusion step. During this transition, the binding pocket is closed and the exit channel opens connecting this pocket to the exit gate toward TolC. This allows the passage of substrate molecules as a result of the peristaltic pumping mechanism.

First of all, the proposed direction of transitions in this cycle has been verified and compared to the inverse direction. While in the first case, the drug is squeezed out of the binding pocket toward the exit gate, it got stuck after leaving the binding pocket during simulations in the inverse direction. The major difference between both directions of transition was the lack of conformational changes which led to an opening of the exit channel. Therefore, the drug remained outside of the binding pocket but still in the center of the pore domain. Despite the open passageway toward the exit, the drug never passed the exit gate during any simulation of the proposed direction. A major issue is probably the limited simulation time compared to the in vivo process which is supposed to occur on the micro- to millisecond time scale. Nevertheless, this was a first important step to understand these conformational changes at an atomistic level.

During the second part of the investigation of AcrB, the influence of the surrounding water and protein environment was examined more closely. Performing sets of at least ten simulations each, it was possible to neglect the less important details of each individual simulation and to focus on the more general aspects of the data. Thereby, a distinct water stream was found to flow from the entrance into the interior of the domain and through the exit channel toward the exit gate. This is another facilitator to accomplish the extrusion of substrates. Furthermore, the water surrounding the drug shields the
drug molecules from the electrostatic attraction of the hydrophilic amino acids along the exit channel. This could bind the molecule to the protein surface without waters being present. While the binding pocket is mainly hydrophobic not allowing any electrostatic interaction, these residues of the exit channel draw the substrate toward the exit gate upon opening of this passageway. Both of these described effects are considered to be important ingredients of an effective extrusion, describing the contribution of water as lubrication effect.

In a last approach, the conformational changes of AcrB were induced more selectively by keeping part of the transporter unbiased, e.g., the transmembrane domain where the electro-chemical energy is converted into mechanical energy. The results describe the responsible subdomain movements for opening and closing of the entrance. The conformational changes of the binding pocket are discussed as well. Additionally, the importance of one particular transmembrane helix is highlighted by additional simulations with adapted selections of steered residues. Moreover, the intermonomeric interactions have been simulated by steering both the transmembrane domain and two of the three periplasmic monomer segments of the transporter. This enhanced the previous results by showing part of the crucial transition to open the exit channel. Here again, the limited simulation time scale constrained the positive outcome of this approach, with much longer simulations elucidating more details. Nevertheless, this more selective steering approach has been shown to induce conformational changes in unbiased regions of AcrB, which can help to enhance the understanding of the inter- and intramonomeric interactions, hindering or assisting the actual function of the transporter. In an additional study, the effect of a particular mutation on the extrusion has been investigated. Therein, molecular dynamics simulations have been compared to docking calculations to verify the outcome. The results indicated a stronger binding of the drug in the pocket which diminished the chances of extrusion.

Overall, this thesis comprises a collection of information of the individual dynamics of each associated protein in the complex of the efflux pump. With both the data that has been obtained about the opening transition of TolC and the extrusion mechanism of AcrB and its related conformational changes, this thesis improved the atomistic and functional understanding of E. coli’s major efflux pump. Nevertheless, the temporal limitation of MD simulations requires the combination of computational and experimental results to achieve the ultimate goal of fully understanding the pumping process with all its details, including the underlying interactions at atomistic level. The described studies in this thesis can be further extended by applying mutations in regions of interest, such as the binding pocket or the exit channel of AcrB as it was shown for one residue already. Additionally, the final extrusion step from the exit channel through the gate and toward TolC still needs to be examined. By application of other methods such as
metadynamics a more quantitative description of the extrusion might be possible. This kind of method might also be useful to quantify the transition between the closed and the open state of the aperture of TolC.

In the future, the obtained information can be used for the investigation of the whole efflux pump to be able to distinguish between movements due to interactions in one protein and due to corporative effects. Therefore, different approaches might be feasible. On the one hand, more elaborate docking techniques could be used to refine the current model of the complex. On the other hand, coarse-grained molecular dynamics simulation would allow to extend the feasible time scale of simulations of the entire efflux pump including two lipid bilayers and the solution. Currently, the actual function of AcrA beyond stabilizing the complex is only suspected, but not fully unraveled yet. The proposed transduction of conformational information from AcrB toward TolC is a reasonable option, but its necessity for TolC as an extended channel is not clear yet. Furthermore, the information of interactions between substrates and the surrounding amino acids can help to setup systematic studies to possibly improve the efficacy of current antibiotics and inhibitors or to create new ones. Therefore, atomistic simulations can assist in specifying the essential interactions for drugs to bind in the pore domain of AcrB. Additionally, the obtained information can assist in understanding the function of other similar transport system, such as human RND transporters, e.g., the Niemann-Pick C1 disease protein and the hedgehog receptor Patched.
M.1. Molecular Dynamics Simulations

As the term “classical” MD already points out, the method of choice here only considers Newtonian mechanics, omitting quantum-mechanical (QM) calculations. Any quantum effects are indirectly incorporated into the collection of interaction parameters, the so-called force field. By numerically solving Newton’s equation of motion for a protein-membrane system, which resembles the protein of interest within its modeled natural environment, we are able to gather atomistic details of each amino acid’s movement in relation to global protein motions. While the protein undergoes conformational changes, the residues have to adapt to the altered energetics in the surrounding, including arriving or departing molecules such as substrates of a transporter (see Ch. 3).

M.1.1. Theoretical Background

Solving Newton’s equation of motion practically means to numerically calculate the forces acting on each atom $i$, which are derived from the sum of several bonded and non-bonded potentials:

\[ \vec{F}(\vec{r}_i) = -\vec{\nabla}V_{\text{total}}(\vec{r}_i) \]  

\[ V_{\text{total}}(\vec{r}_i) = V_{\text{nmbnd}}(\vec{r}_i) + V_{\text{bnd}}(\vec{r}_i) \]  

(M.1)

(M.2)
Nonbonded potential (van-der-Waals and electrostatic):

\[ V_{\text{nonbonded}}(\vec{r}_i) = V_{\text{VDW}}(\vec{r}_i) + V_{\text{el-stat}}(\vec{r}_i) \]  

(M.3)

Bonded potential (bond, angle, proper, and improper dihedrals):

\[ V_{\text{bonded}}(\vec{r}_i) = V_{\text{bond}}(\vec{r}_i) + V_{\text{angle}}(\vec{r}_i) + V_{\text{p. dihedral}}(\vec{r}_i) + V_{\text{improper d.}}(\vec{r}_i) \]  

(M.4)

These equations describe the components of the total potential (Eq. M.2), i.e.:

- the nonbonded interactions including the van-der-Waals interaction, represented by the \textit{Lennard-Jones (12,6) potential}, and the electrostatic potential (Eq. M.3) as well as

- the bonded interactions between two (bond distance), three (angle), and four (torsion angle or proper, and improper dihedral) neighboring atoms (Eq. M.4).

The required parameters for the potential components are either experimentally obtained or approximated via electron-density calculations, e.g., CHARMM [88] or AMBER [109]). In the course of this thesis, different force fields have been used to study different aspects of the extrusion process. While in Sec. 2.1, CHARMM was applied in correspondence to Ref. 62, 66, 67, AMBER was used for the studies on AcrB (Ch. 3). This change was necessary due to the lack of parameters for the drug, which was placed into the AcrB transporter. In collaboration with A. V. Vargiu, F. Collu, and P. Ruggerone, the parameters have been derived from quantum-mechanical calculations using Gaussian [113] (for details see Sec. 3.1.4).

\textbf{Limitations to Simulations} Classical MD is applied here is to simulate the extrusion of antibiotics. But within the nature of this computational method there are, of course, limitations considering simulation time scale and system size. On the one hand, while bond vibrations occur on the time scale of femto- and picoseconds, protein folding for example often takes milli- to seconds. Nevertheless, the common simulation length in this thesis remains in the regime of nanoseconds. On the other hand, the maximum system size which can be simulated is limited by the computational power at hand as well. Although the biggest simulated system is in the range of a few millions of atoms, usual system sizes are of a few hundred-thousands of atoms.
M.2. Techniques for Simulation & Analysis

In this section, the applied MD simulation techniques are presented and described in more detail. Depending on the type of protein, e.g., porin or transporter, the applied methods have been chosen carefully to consider the structural and dynamical properties. As a long channel, TolC does not change conformation at the same magnitude as AcrB. Hence, these two demand different choices from the list of available MD methods. For TolC, the opening of the closed periplasmic tip was of particular interest, which should occur during docking to AcrB. The latter one required a more general induction of conformational changes, i.e., forcing the transitions between already known conformational states.

M.2.1. Steering with an Electric Field

To speed up the MD simulations of TolC with mutated periplasmic tip, an electric field was applied to the system. The resulting electrostatic forces assisted in the removal of ions, especially potassium, from unfavorable locations such as the closed tip of the alpha-helical channel. This approach is similar to so-called grid-steered MD, which was used to steer molecules through membrane pores [86]. As described in Sec. 2.1, the MD simulations were performed in correspondence to equivalent mutant structures from X-ray crystallography.

In NAMD [64], the electric field is applied homogeneously along the user-defined field direction. Unfortunately, this method does not allow for differing environmental adaptions of the dielectric constant, e.g., in the lipid bilayer or within a membrane pore. Nevertheless, the effect of this electric field on the ions is not major, just adding a minor push to the movement of the ions (see Sec. 2.1 for more details). The direction of the applied positive voltage can be seen in Fig. 2.1C, which steers the cations downward from the extracellular side toward the periplasm and thereby adjacent to the described cation pockets. The gradient of the resulting electrostatic potential is different to the homogeneous electric field, therefore, an electrostatic potential map is shown in Fig. 2.6. There, a continuous gradient through the open channel strongly differs to the sudden decline in the closed wildtype. Fig. M2.1 presents slices through such potential right through the middle of the channel. As it was already shown in Sec. 2.1, the slope of the potential does not only depend on the structural properties, but also on the direction of the applied voltage. This is also due to the accumulation of ions at the periplasmic leaflet of the membrane where the most extreme potential values can be measured (compare Fig. 2.6).

Although this method allowed for quicker removal of cations from the described ion...
pockets, the major parameter which influenced the results of the performed simulations was the concentration of KCl in the solution, which resulted in a partial opening during the initial unbiased simulation step ($\text{MGlu0}$ in Sec. 2.1). Interestingly, the direction of the electric field had an additional effect, because the positive voltage induced a closing of this already opened structure, due to the cations being lead through the channel and close to the ion pockets.

M.2.2. Selective Boundary Forces

The second approach used to investigate the opening transition of TolC was adapted from a method which is offered in NAMD [64]: ‘Tcl Boundary Forces’. This user-definable method allows for the placement of virtual barriers in a simulation box. This approach could be used to mimic the effect of a surface or membrane by instituting a potential wall with a user-defined mathematical function, e.g., exponential, to describe the repulsion of selected atoms if they come close to this barrier. Another idea, which was used for this project, is to define a spherical boundary which could either confine a collection of atoms within or keep them out of a defined region. The latter approach was used dynamically to push the tips of the TolC outward. This can be imagined as placing a balloon in the center of the aperture and inflate it (see left panel of Fig. M2.2). This way the individual direction of movement of each helix is not predefined as long as it slightly points outward from the center. This is in contrast to standard “steered MD”, where the direction of pulling has to be defined beforehand, which only allows for one chosen direction. As the magnitude of pushing can be varied by altering the speed of “inflating” and stiffness of the barrier, the speed of helical tilting can be adapted to allow
for the proposed iris-like, rotational movement away from the center. Furthermore, this can help to not break the secondary structure of the helices.

Figure M2.2: Left: Description of the “inflation of the balloon” from the center of the TolC aperture; Center: Trial result by pushing three out of six helices; Right: Trial result by pushing all six helices.

While a sequence of transition states was obtained, no quantitative measures have been able to be obtained. A next step would be to implement a force measurement procedure to this method or to combine this idea with tools that calculate the change of free energy during this transition. Therefore, the radius of the virtual balloon would be the simplest collective variable.

M.2.3. Targeted Molecular Dynamics Simulations

If information about possible conformational states is available acquired, e.g., by X-ray crystallography, another steering approach becomes feasible, called targeted MD (TMD). This method allows to interpolate between this kind of states. Upon application, the root mean square deviation (RMSD, see Eq. M.5) between initial and target structure is attempted to be reduced linearly over the predefined TMD simulation time (see Fig. M2.3). Thereby, each selected atom is pulled in the direction of the target state, with the force scaled by the RMSD difference between the current state and the ideal RMSD value of the linear decline. In Ch. M3 TMD was applied to mimic the conformational changes between monomeric states from structures obtained by Sennhauser et al. [28].

\[
\text{RMSD}(r_{\text{sel}}, r_{\text{ref}}) = \sqrt{\mathbb{E}((r_{\text{sel}} - r_{\text{ref}})^2)} = \sqrt{\frac{\sum_{i=1}^{N_{\text{atoms}}} (r_{\text{sel},i} - r_{\text{ref},i})^2}{N_{\text{atoms}}}} \quad (\text{M.5})
\]
As it can be seen in the presented studies (Ch. 3), there are several crucial parameters which can significantly alter the outcome of TMD simulations. Firstly, the time scale is usually much less than the corresponding in vivo process. While more general quantities have been found to be comparable for simulations in the range of one to ten nanoseconds, the details might differ notably. One reason might be that the steered parts do not always follow the path of minimum transition energy. During the investigation of AcrB, this aspect has been examined and no tremendous change in interaction energy could be measured (Sec. 3.1). Secondly, the defined force constant renders the pulling forces too stiff. This issue was part of the same study and did only show minor differences in the scope there. The last and probably most interesting TMD parameter is the scheme of selecting the atoms to be steered. In Sec. 3.1, two more local schemes have been compared to verify the results described there (see also Fig. M2.4). While pulling only the Cα atoms would keep the side chains untouched, the steering of all heavy atoms of the protein lead to less variance in the results (see Fig. S1.8). Interestingly, some of the Cα-steered simulations also resulted in similar positions of the drug as for the other scheme. While the consistency is higher in the “all heavy atoms” simulations, the variation of the Cα simulations might help to clarify the crucial details of this extrusion.
process. To understand the underlying interactions of the substrate and its changing environment during the extrusion from the binding pocket, only the $C_\alpha$ atoms have been steered in the second subproject (Sec. 3.2).

![Figure M2.4: Examples of the applied atom selections for TMD simulations: two sample residues in stick representation, highlighted are “all heavy atoms” (all spheres) and only $C_\alpha$ atoms (only red spheres); the backbone is drawn in blue.](image)

The initial idea of steering the entire protein from one state to another one was evaluated critically during the progress of this thesis. To study the effects of conformational changes of domains onto each other, the scheme of TMD selection was altered in another investigation. On the one hand, the proposed pathway of mechanical energy - from the transmembrane domain to the pore domain - has been investigated by steering the lower part. This allowed to observe the crucial elements for the energy transmission and the effect on the unbiased regions in the periplasmic region. On the other hand, one monomer was chosen as the observed, unsteered segment, which is the occupied one, because the same system was used as in the previous studies. Then, both neighboring monomers have been steered using TMD and their influence on the unbiased monomer was examined. Furthermore, both selection schemes have been combined to watch their cooperative effects.

**M.2.4. Statistics of the Residence of Water Molecules**

The following section compiles information from different sources as a manual for the procedure to obtain statistics of the residence of waters close to a particular atom or
group of atoms. This has been assembled with the help of Amit Kumar. The basic method was introduced by Impey et al. in 1983 [148]. In 2001, this method was used to estimate the average residence times of waters close to a lysozyme [149]. Furthermore, Kumar et al. [63] used this approach to examine the waters in porins such as OmpF. Despite the effort, no conclusive results could be obtained during the study of AcrB with doxorubicin inside.

**Definitions**

At first, the following table introduces some definitions which are required later on:

<table>
<thead>
<tr>
<th>variable</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V(t)$</td>
<td>region of interest</td>
</tr>
<tr>
<td>$N_W$</td>
<td>total number of distinct waters in $V(t)$ during simulation</td>
</tr>
<tr>
<td>$N_T$</td>
<td>number of discrete time steps (frames)</td>
</tr>
<tr>
<td>$O_j(t_n, \Delta t; t^*)$</td>
<td>occupancy by water $\in {0; 1}$</td>
</tr>
<tr>
<td>$j$</td>
<td>index of water oxygen $\in [1; N_W]$</td>
</tr>
<tr>
<td>$t_n$</td>
<td>discrete time of trajectory with index $n$</td>
</tr>
<tr>
<td>$\Delta t$</td>
<td>time span</td>
</tr>
<tr>
<td>$t^*$</td>
<td>time span describing temporary leave of waters from $V(t)$</td>
</tr>
<tr>
<td>$\bar{N}_W(\Delta t)$</td>
<td>average number of waters with a residence time of $\Delta t$</td>
</tr>
<tr>
<td>$P_S$</td>
<td>survival probability of any water within $V(t)$</td>
</tr>
<tr>
<td>$\bar{N}_f, \tau_f$</td>
<td>average number and lifetime of fast waters (short lifetime)</td>
</tr>
<tr>
<td>$\bar{N}_m, \tau_m$</td>
<td>average number and lifetime of medium waters (intermediate lifetime)</td>
</tr>
<tr>
<td>$\bar{N}_s, \tau_s$</td>
<td>average number and lifetime of slow waters (long lifetime)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>stretching parameter of fast waters (diffusion parameter?)</td>
</tr>
</tbody>
</table>

**Theory**

As described in Ref. [148], $\bar{N}_W(\Delta t)$ can be calculate by:

$$\bar{N}_W(\Delta t) = \frac{1}{N_T} \sum_{j=1}^{N_W} \sum_{n=1}^{N_T} O_j(t_n, \Delta t; t^*) \in \{0; 1\}$$ \hspace{1cm} (M.6)

$$P_S = \frac{\bar{N}_W(\Delta t)}{N_W}$$ \hspace{1cm} (M.7)

The resulting data from Eq. M.6 can be used to describe the mobility of waters in $V(t)$ (see Sec. M.2.4). But before going into details of any further analyses, the following section will describe how to obtain $O_j(t_n, \Delta t; t^*)$. 

92
Determining the Occupation of $V(t)$ by Waters  To calculate $O_j(t_n, \Delta t; t^*)$, each water in $V(t)$ has to be taken into account. Therefore, the presence of a water $j$ in $V(t)$ will be counted. Thereby, each event, i.e., the presence of water $j$ for a certain time span $\Delta t$, is recorded including its length. Here, one particular detail is of importance. If water $j$ remains in $V(t)$ for $\Delta t$, events of shorter than $\Delta t$ have to be counted as well:

<table>
<thead>
<tr>
<th>Data: simulation trajectory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input: $V(t), t^*$</td>
</tr>
<tr>
<td>foreach water $j$ and time step $t_n$ do</td>
</tr>
<tr>
<td>if water $j$ is in $V(t)$ from $t_n$ until $(t_n + \Delta t)$ without leaving longer than $t^*$ then</td>
</tr>
<tr>
<td>$O_j(t_n, \Delta t; t^*) = 1$</td>
</tr>
<tr>
<td>else</td>
</tr>
<tr>
<td>$O_j(t_n, \Delta t; t^*) = 0$</td>
</tr>
<tr>
<td>end</td>
</tr>
<tr>
<td>end</td>
</tr>
<tr>
<td>Output: $O_j(t_n, \Delta t; t^*)$</td>
</tr>
</tbody>
</table>

Algorithm 1: Occupation of $V(t)$

Using Alg. 1 all probable occupation events are taken into account without ignoring sub-events of longer occupations. After processing this algorithm and continuing with Eq. M.6, Sterpone et al. suggested to divide the resulting dataset into three groups with different average residence times [149].

Groups of Differing Residence Times of Waters  The division into three groups of average residence times is performed by fitting three exponentials to the dataset of $\bar{N}_W(\Delta t)$. This fitting function was described in Ref. [149]:

$$\bar{N}_W(\Delta t) \simeq f(t) = \bar{N}_f \cdot \exp(-t/\tau_f^*) + \bar{N}_m \cdot \exp(-t/\tau_m^*) + \bar{N}_s \cdot \exp(-t/\tau_s^*)$$ (M.8)

During the fitting process, a few issues should be kept in mind:

- Set reasonable initial values for all parameters of eq. M.8:
  - $\bar{N}_W(0) = N_f + N_m + N_s$
  - $\gamma \in (0; 1]$ (supposedly subdiffusion)

- If any $N_i$ is fitted lower than zero, fix it to zero or set bounds; consider neglecting the corresponding $\tau_i$ for the fit.

- Ref. [149] suggests to subtract the number of permanently residing waters.
• Ref. [63] uses this approach to describe water behavior within segments of a channel.

**Additional Remarks** In Ref. [148], $t^*$ was suggested to be in the range of $[0; 2]$. Therein, they set $t^* = 2$ ps for simulation trajectories with a time step of 1 ps. In case of a time step of 5 ps, $t^* = 0$ ps is favorable.
S.1. Functional Rotation of the Transporter AcrB: Insights into Drug Extrusion from Simulations (Sec. 3.1)
Figure S1.1: Drug displacement vs. protein conformational change: Plot of the distance between the COMs of doxorubicin and that of the binding pocket, $d_{\text{DOX-BP}}$, (red full line) and of the RMSD of the whole protein with respect to the target structure (black dashed line), as a function of the TMD simulation time during the T→O transition. The time evolution of the RMSD of the binding pocket from the target structure is also shown (blue dot-dashed line).

Figure S1.2: Profiles of $d_{\text{DOX-BP}}$ for two different conformation and orientation of the drug along the T→O transition: A) Profiles of $d_{\text{DOX-BP}}$ for the TMD simulation discussed in the main text (red full line) and for one simulation in which doxorubicin is in a different conformation and orientation within the binding pocket (green dashed line). In the inset is reported the behavior of the interaction energy between the drug and the residues of the binding pocket. Also shown are the two different initial positions - B) top view; D) side view - as well as (D) the final positions of the drug.
Figure S1.3: Drug displacement vs. simulation parameters along the T→O transition. A) Distance $d_{\text{DOX-BP}}$ as a function of the percentage of TMD simulation time for the set of simulations where all heavy atoms have been targeted. To better identify different behaviors, we report running averages of length 10 of the raw data. Varying the initial velocities within the set of simulations of same length and $k$ value does not remarkably alter the profile of $d_{\text{DOX-BP}}$. Extending the simulation time does not sensitively affect the final position of the drug, although the profile of $d_{\text{DOX-BP}}$ show some differences with respect to the former set; B) Final positions of doxorubicin in the same set of simulations. The COMs of the drug are shown as filled spheres to highlight the similar displacements of the drug despite the difference which can be seen in the orientation.
Figure S1.4: Profile of $d_{\text{DOX-BP}}$ after the $T\rightarrow O$ step. The panel reports the behavior of $d_{\text{DOX-BP}}$ in eight post-TMD simulations. Two sets of such simulations (each 2 ns long) have been performed starting from the final configurations found in each of the TMD simulations with $k = 3$ and TMD time 1 ns. In the first set we have removed all the restraints from the system, in the second we have restrained C$_{\alpha}$ atoms. It can be seen that in half of the simulations without restraints the drug moves further towards the gate, while in the remaining ones it oscillates around the final position. Importantly, doxorubicin never goes back towards the binding pocket.

Figure S1.5: Drug displacement vs. subdomains movements in the periplasmic region. A) Time evolution of $d_{\text{sub}}$, the distance between COMs of the subdomains (shown in panel B) of AcrB mostly involved in the conformational changes during the $T\rightarrow O$ transition (see also Video S1.10). Larger changes in the PC1-PC2 (red line), PN2-PC2 (yellow line) and PC1-PN2 (magenta line) distances occur within the first half of the simulation, while the displacement of the substrate, $d_{\text{DOX-BP}}$ (black dotted line, arbitrary units), essentially increases in the second half. B) Top view of the aforementioned subdomains (doxorubicin is shown as red-sticks).
“Peristaltic” closing of the binding pocket induces squeezing of the drug. Profile of the minimum distance $d_{\text{min}}$ between selected pairs of residues within the binding pocket, as a function of the TMD simulation time during the T→O transition. Results are shown from eight simulations in which all heavy atoms are steered. In each panel we reported the corresponding profile of $d_{\text{DOX-BP}}$ (red dotted lines, rescaled to fit in the graph). To better identify different behaviors, we report running averages of length 10 of the raw data. In 4 out of 5 the 1-ns-long TMD simulations (panels A, C, D, E) a three-step zipper-like closure of the residues lining the binding pocket can be roughly appreciated. Panel B shows a slightly different behavior, which could be viewed as a limit process in which the last step is very short. It is worthwhile to point out that also in this case the closure of the binding pocket occurs in a sequential manner, with outermost pairs F176-F615 and F136-F615 closing after the innermost ones. In panel F (5-ns-long

Figure S1.6
TMD) two steps can be roughly identified; again, the innermost pair F136-F628 closes before the others and the outermost 176-615 as last. The second 5-ns long simulation, panel G, shows a clearer three-steps behavior. The 10-ns simulation (panel H) also shows a zipper-like closure of residues in the binding pocket, although this appear less evident than in the previous cases. Indeed, distances between the innermost pairs, F136-F628 and F136-F617, reduce from ~7.5 to ~4 Å in about 3 ns, while during this time interval the outermost pairs F136-F615 and F176-F615 close only partially, going the corresponding minimum distances from ~9 to ~7 Å and from ~7 to ~5 Å, respectively. Careful inspection of the graphs reveals that a further step, occurring about 1 ns later, is necessary for their complete closure. Additionally, it can be seen that the distances between outermost pairs drops at almost equivalent time. Concerning the 5-ns and 10-ns simulations, we would like to stress that, due to the large computational time needed to perform them, obtaining a relevant statistics is out of reach.
Figure S1.7: Effect of targeting the BP-Gate path on the movement of the drug. A) View of the BP-Gate path. Amino acids in monomers T and O are shown in blue and red, respectively (in silver the binding pocket, in transparent yellow the gate); B) Final configurations obtained respectively from TMD simulations performed with (doxorubicin colored in red) and without (cyan) bias applied to the BP-Gate path for the (T→O transition; the initial position of the drug is shown in transparent yellow). The differences in the structure of the BP-Gate path between the “standard” TMD simulation and the one without the bias on the BP-Gate path are represented using a color scale tuned on the value of the RMSD with respect to the final structure from “standard” TMD run; C) Evolution of $d_{DOX,BP}$ (main graph) and of the RMSD (inset) as a function of TMD simulation time for the two simulations (color code as in B).
Figure S1.8: Effect of targeting side chains on the displacement of doxorubicin. Plot of $d_{\text{DOX-BP}}$ as a function of the simulation time during the $T \rightarrow O$ transition for two sets of TMD simulations (averages are shown): Steering all heavy atoms (red line); Steering only $C_{\alpha}$ (green line).

Figure S1.9: Drug displacement vs. value of the force constant used for TMD. A) Plot of $d_{\text{DOX-BP}}$ (main graph) and of the RMSD of the binding pocket (inset) as a function of the TMD simulation time during the $T \rightarrow O$ transition for two values of the force constant applied on all heavy atoms, $k = 3 \text{ kcal-mol}^{-1} \cdot \text{Å}^{-2}$ (red full curve represents an average from all the runs having different initial velocities), $k = 2 \text{ kcal-mol}^{-1} \cdot \text{Å}^{-2}$ (magenta dashed curve). B) Representation of the initial (yellow) and final (color code as in panel A) configurations of the drug. Note how small differences in the structure of the binding pocket affect the value of the drug displacement.
<table>
<thead>
<tr>
<th>Cycle step</th>
<th>TMD [ns]</th>
<th>$k_{\text{atom}}$ [kcal·mol$^{-1}$·Å$^{-2}$]</th>
<th>Targeted atoms</th>
<th>$N_{\text{sim}}$</th>
<th>post TMD [ns]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Free</td>
<td>Restr. C$_{\alpha}$</td>
</tr>
<tr>
<td>TO</td>
<td>1</td>
<td>heavy</td>
<td>1</td>
<td>0.25</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>heavy</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>C$_{\alpha}$</td>
<td>10</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>not exit path</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>heavy</td>
<td>2</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>heavy</td>
<td>2</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>TL</td>
<td>1</td>
<td>heavy</td>
<td>1</td>
<td>0.25</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>heavy</td>
<td>1</td>
<td>0.25</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>heavy</td>
<td>3</td>
<td>0.25</td>
<td>-</td>
</tr>
<tr>
<td>LO</td>
<td>1</td>
<td>heavy</td>
<td>2</td>
<td>0.25</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>heavy</td>
<td>2</td>
<td>0.25</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table S1.1:** Details of the different simulations. Shown are the cycle direction, the simulation time, the force constant, the selected atoms for the TMD, the number of simulations with the same setup, and the length of post-equilibration standard MD simulations after the targeting was finished, with or without restraints on the C$_{\alpha}$ atoms.
Figure S1.10: Movement of doxorubicin and of the four subdomains PC1, PC2, PN1, PN2 extracted from one TMD simulation along the T→O transition. Firstly the whole trimer is shown, then zoom is performed on the T monomer, in the region around the drug. Doxorubicin is shown initially in blue, and becomes red following the TMD time. Some residues of the binding pocket are also highlighted. Found at: doi:10.1371/journal.pcbi.1000806.s011 (8.52 MB MP4)
S.2. Role of Water during the Extrusion of Substrates by the Efflux Transporter AcrB (Sec. 3.2)

S.2.1. Analysis Methods

Throughout the entire article, several nonstandard methods are employed to describe and analyze the extrusion process in molecular detail. All of these methods have been implemented in VMD [77].

S.2.2. Grid Path for Direction of Water Movement

The grid path (see Figs. 3.13, 3.14, and 3.16) is modeled from the entrance between the subdomains PC1 and PC2, here simply calculated as the COM of these domains, via the COM of the waters surrounding the BP to the COM of the exit gate. As there are only few waters in the BP, the point of measurement at the BP was shifted to the bordering waters (the selected waters are in a region of 3.5 Å from the COM of the BP) because we are more interested in the movement of internal waters in general. The grid path is divided into five equidistant segments along both parts, i.e., five segments from the entrance to the BP and another five from the BP to the exit resulting in eleven grid points. At each of these grid points, waters within a sphere of 9 Å radius are selected. Then the distance each water molecule moves within 50 ps is measured. At the end of the procedure, those displacements are averaged per grid point leading to average directed displacements per 5 ps per water molecule.
S.2.3. Additional Graphs

Figure S2.11: Number of water molecules surrounding the substrate as a function of time for representative TO trajectories.

Figure S2.12: Ratio of hydrophilic over hydrophobic amino acids in the vicinity (3.5 Å) of the substrate as a function of the drug-BP distance of all TO simulations including a running average.
**Figure S2.13:** Number of hydrogen bonds as a function of the drug-BP distance of all TO simulations including a running average.

**Figure S2.14:** Number of water molecules in the vicinity of the drug, separated by the head and the tail, as a function of the drug-BP distance of all TO simulations including a running average.
**Figure S2.15:** Sketch of the spherical coordinate system and the model path from the entrance through the interior of the pore domain toward the exit. Please note that this is the same path as already indicated in Figure 2. Furthermore, the orientation angles of the drug are shown ($\phi$ in the $x$-$y$ plane parallel to the membrane and $\theta$ down from the $z$ axis to the major axis of the substrate).

**Figure S2.16:** Plot of the drug-BP distance as a function of the relative simulation time. Results from all the simulations are shown as dots, and a running average is represented by the solid line.
**Figure S2.17:** Orientation of the drug in the AcrB pore domain along TL simulations. The vertical line represent the initial state again.
Figure S2.18: Ratio of hydrophilic over hydrophobic amino acids in the vicinity (3.5 Å) of the substrate as a function of the drug-BP distance of all TL simulations including a running average.


