University of Cagliari

PhD in Physics

Computational Study on MFP MexA in MexAB-OprM and RND transporter AcrB in AcrAB-TolC Efflux Pumps

FIS/03

Tutor:

Prof. Paolo Ruggerone

Co-tutor:

Dr. Attilio V. Vargiu

Coordinator:

Prof. Paolo Ruggerone

PhD Candidate:

Pierpaolo Cacciotto

XXVII Cycle - Academic year 2013/2014

Contents

Abbreviations

ABC   ATP-Binding Cassette

AcrA   Acriflavin resistance protein A

AcrB   Acriflavin resistance protein B

DD   Docking Domain

GPU   Graphics Processing Unit

MATE   Multi-drug And Toxic-compound Extrusion

MD   Molecular Dynamics

MFP   Membrane fusion protein

MFS   Major Facilitator Superfamily

OMF   Outer membrane factor

OMP   Outer membrane protein

PBC   Periodic Boundary Condition

PD   Porter Domain

PME   Particle mesh Ewald

RDF   Radial distribution function

RMSD   Root mean square deviation

RMSF   Root mean square fluctuation

RND   Resistance Nodulation Division

RoG   Radius of Gyration

SMR   Small Multidrug Resistant

SS   Secondary Structure

TolC   Tolerance to Collicins

TMD   Transmembrane Domain

Chapter 1

Introduction

Bacteria are microscopic single-cell organisms that populate every conceivable environment all over the world. With a million-years old evolution bacteria developed an high propensity to adapt themselves to external conditions. In fact, they have been reported to live in soil, deep within the earth's crust or the sea [1] and dormant in the poles [2], they survive in radioactive waste [3] and in the extreme conditions of the space flights [4]. Some of them became symbiont with other organisms from which they benefit nutrients and relatively safe environments. An example of such cooperation is the human body, which contains 10^{14} microbes [5–7] in skin, mouth or the intestine. They are part of the so-called *bacterial flora* and their presence is generally of benefit for human since they contribute, as example, to the food digestion process. However, they can also act as opportunistic pathogens if they access other parts of human body that are typically sterile or in presence of a compromised immune system [5].

Since the discovery of the first antibiotic molecule in 1940s, the treatment of bacterial infections was based on the idea that antibiotic therapies could cure any
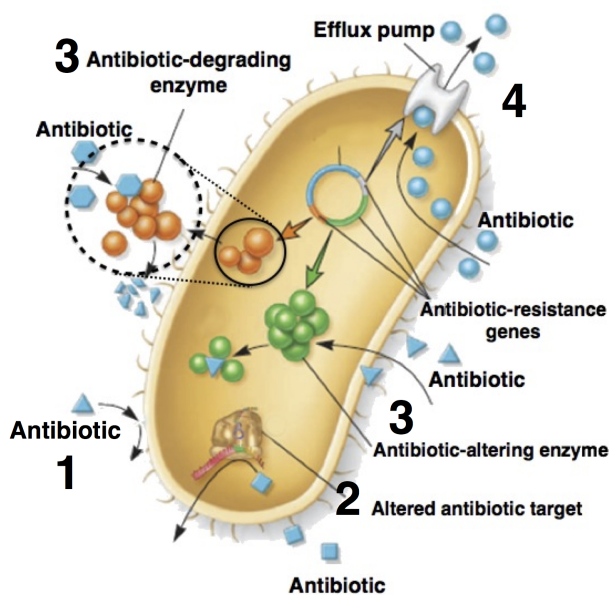
Figure 1.0.1: Scheme of the major defense mechanisms developed by bacteria to survive to external threats: (1) low membrane permeability that regulate the entrance of molecules, (2) mutation in the antibiotic target site, (3) enzymes can inactive or disrupt the antibiotic molecule, (4) efflux pumps can extrude toxic compounds from the inside to the outside.

infectious disease. However, the adaptation ability of bacteria made them resistant to most of the common antimicrobial agents [8–11]. In fact, while the number of antibiotics molecules developed each year is continuously decreasing we observe an increase of the bacterial resistance to the common drugs with the risk to reach the point where bacteria will be resistant to all the known antimicrobial agents [12–14]. Several mechanisms have been developed by bacteria to resist the action of antibiotics. These include the possibility to prevent the antibiotic molecules from entering the bacterial cell by lowering the membrane permeability, to induce mutation in antibiotics target, to produce enzymes that could inactivate the antibiotic molecules, or to actively export them through the so-called efflux pumps [8,9,15–17], as described

in Figure 1.0.1. In the following a more accurate description of the extrusion process and the efflux pumps machineries will be provided, with particular attention to the *Escherichia coli*, *Salmonella typhimurium* and *Psaeudomonas aeruginosa* case.

1.1 Multi-drug resistance

Efflux pumps are continuously expressed in bacteria to fulfill physiological functions, like bile extrusion [18] or calcium transport [19]. However, their over-expression in case of external attack has been proved to be one of the main defense system that confers bacteria the so-called multi-drug resistance (MDR) through an active extrusion of toxic molecules [20–25]. These molecular machineries that act as pumps can be specific for one substrate or can transport several unrelated compounds. This makes them particularly interesting as an understanding of their functionality and the mechanisms involved in the extrusion process might improve the therapeutic strategies. An example of such improvements is the design of a new class of inhibitors to reduce the effect of the efflux pumps in the extrusion of antibiotics [25–31].

A large number of proteins are involved in the efflux pumps assembly and they can divided into five families on the basis of the number of components, the energy source and the type of substrates that the pumps export [9,15,32]: the ATP-binding cassette (ABC) superfamily, the major facilitator superfamily (MFS), the multi-drug and toxic-compound extrusion (MATE) family, the small multidrug resistant (SMR) family and the resistance nodulation division (RND) family. As described in Figure 1.1.1 the members of these families are located in the inner membrane of Gram-negative (or in the single membrane of Gram-positive) bacteria or can span the whole periplasm, connecting the two membranes. To accomplish their role, efflux pumps that belong to different families can cooperate in synergistic ways to extrude

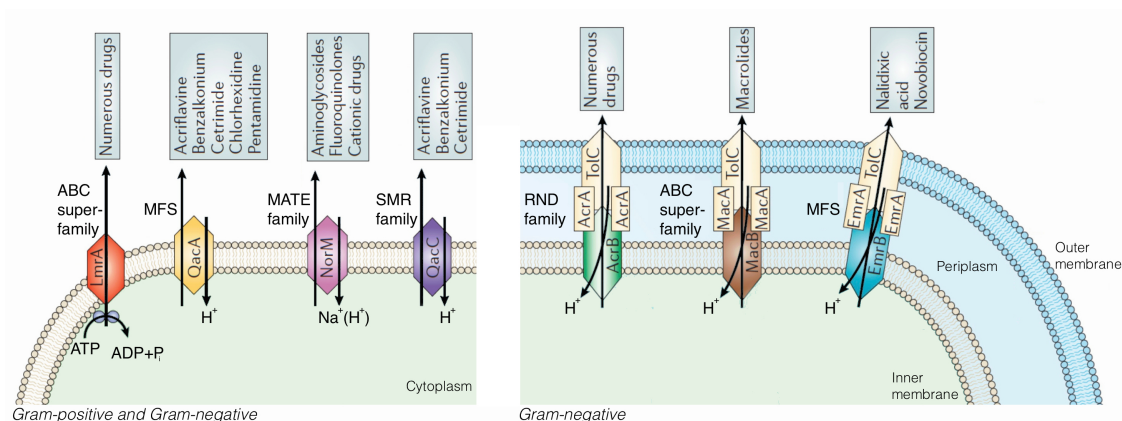the compounds out of the cell, from the cytoplasm or the periplasm [33], as described in Figure 1.1.2.

Figure 1.1.1: Efflux pumps that confer multi-drug resistance, divided into five families.the ATP-binding cassette (ABC) superfamily, the major facilitator superfamily (MFS), the multi-drug and toxic-compound extrusion (MATE) family, the small multidrug resistant (SMR) family and the resistance nodulation division (RND) family. In the picture on the left is reported an example of efflux pumps expresses both in Gram-positive and Gram-negative bacteria; on the right an example of efflux pumps expressed only in Gram-negative. In this last case, the machineries can span the whole periplasm connecting the two membranes. Image adapted from [15].

The proteins studied in this thesis belong to the RND family, whose members play the major role in the intrinsic and acquired MDR in Gram-negative bacteria [15, 34, 35]. The efflux pumps related to the protein studied in this thesis will be described in the following paragraphs.
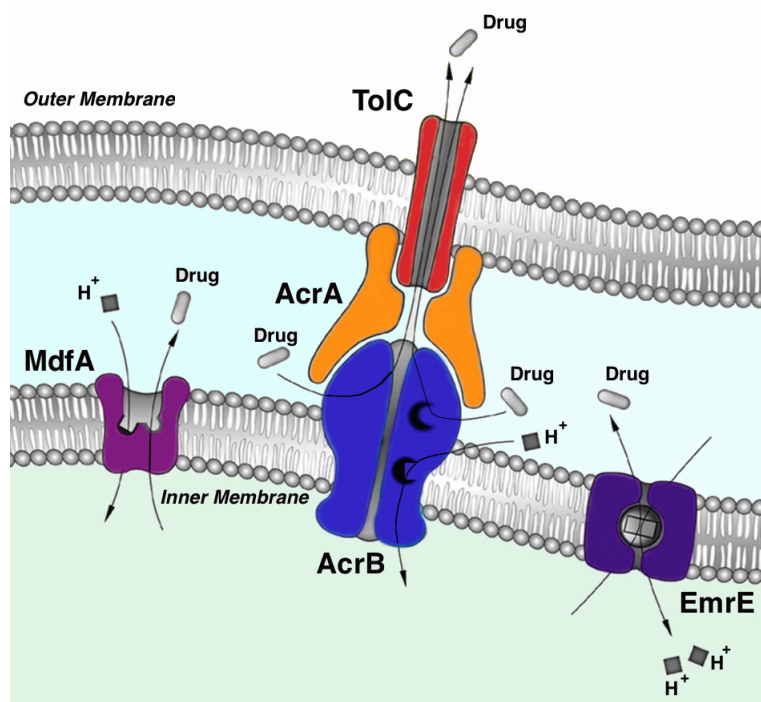
Figure 1.1.2: Example of functional interaction among EmrE (SMR), MdfA (MFS) and AcrAB-TolC (RND) in *E. coli*: EmrE and MdfA remove drug from the cytoplasm into the periplasm, where AcrAB-TolC captures it and extrudes to the outside. Image adapted from [33].

1.2 AcrAB-TolC efflux pump

The RND complex AcrAB-TolC is one of the most studied and most important efflux pumps expressed in *Escherichia coli* and *Salmonella enterica* [33, 36–38].

E. coli is a Gram-negative bacterium that lives in symbiosis with several animals, included humans. It is part of the intestinal flora and it is generally of benefit for the hosts since it assists the digestion process and can prevent other pathogens to establish [39]. However, it can be also a pathogen for the human body causing several infections in the urinary tract, neonatal meningitis, acute diarrhea and gastroenteri-

tis [40–43]. *S. enterica* is another Gram-negative bacterium and pathogen of humans and animals. Some serovars are more prevalent in certain hosts, like *Typhi* in humans, while others are promiscuous, like *S. typhimurium* [38] which is the serovar studied in this thesis. Depending of the host and the serovar, it can mediate either a gastrointestinal disease or systemic disease, can infect macrophages and promote mucosal inflammation. Both *E. coli* and *S. enterica* are considered as extremely dangerous for humans, and classified respectively as urgent and serious threat by CDC [44].

As members of Gram-negative bacteria, *E. coli* and *S. enterica* have two membranes with a peptidoglycan layer in the periplasmic space. Their intrinsic resistance has long been attributed to the outer membrane, for which the limited permeability could be a barrier against drugs [45]. However, it was soon shown that *E. coli* mutants with no *acr*AB genes loose part of their multi-drug resistance [46], suggesting a more important role for the efflux pumps than other mechanisms [22], at least during the first steps of antibiotic therapies.

Further studies showed that the AcrAB-TolC efflux pump is one of the major responsible of the multi-drug resistance in *E. coli* and *S. enterica* [31, 36, 37, 47–52]. The pump spans from the inner to the outer membrane allowing a direct efflux of substrates [47, 53–56], as described in Figure 1.1.2. In the following, a more detailed description of the pump component is provided.

1.2.1   AcrB

The acriflavin resistance protein B (AcrB) is the RND transporter of the AcrAB-TolC efflux pump partially embedded in the inner membrane of *E. coli* and *S. enterica* [57]. The protein has been resolved as an asymmetric homotrimer. Each monomer has
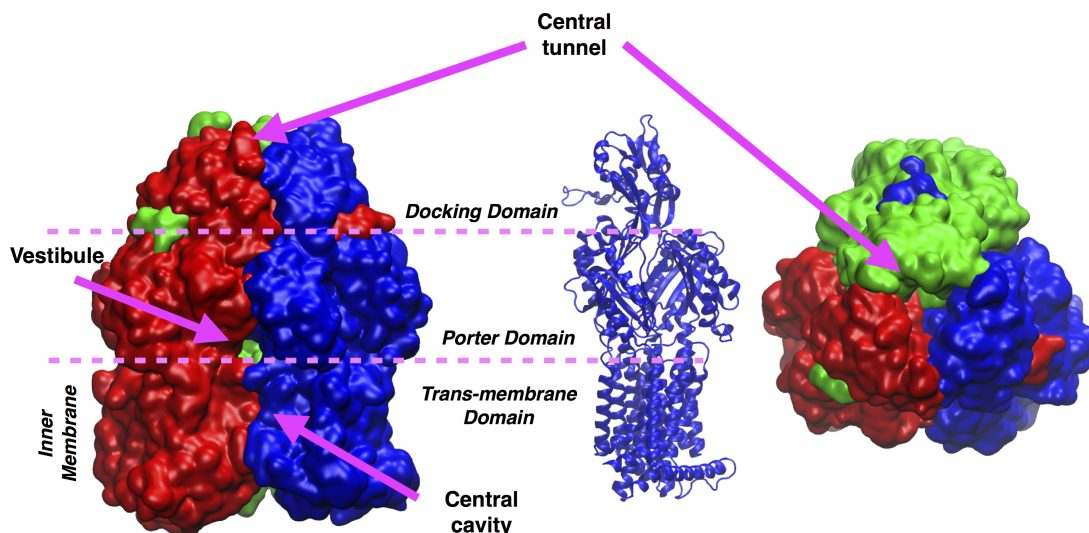
Figure 1.2.1: Crystal structure of AcrB structure (PDB code: 4DX5) divided into its three subunits: monomer 1 in the L state [Blue], monomer 2 in the T [Green] and monomer 3 in the O state [Red]. The three AcrB domains are located in the upper part (docking domain), middle part (porter domain) and lower part (trans-membrane domain). In the TM domain the three monomers form a lipid-filled central cavity connected to the periplasm through three vestibules. The central tunnel in the docking domain is the exit from which drugs are supposed to pass from AcrB to TolC.

a conformation representative of the different extrusion process steps [53, 58–60], as described below in more detail.

Further divisions of each monomer identify three main domains: *a*) the transmembrane domain (TMD) *b*) the porter domain (PD) and *c*) the TolC-docking domain (DD) , as shown in Figure 1.2.1, which are respectively involved in the energy transfer, drug capture and drug extrusion. In addition, AcrB monomers form a central transmembrane cavity which has been suggested to be filled by lipids. Further divisions in each monomer identify other subdomains, named PN1, PN2 and DN, which correspond to the amino-terminal part in the PD and DD domain. Other subdo-

mains are PC1, PC2 and DC, located in the carboxyl-terminal part among the same domains of the amino-terminal subdomains, as shown in Figure 1.2.2.

AcrB and its homologues are the major multi-drug efflux transporters in Gram-negative bacteria and confer intrinsic drug tolerance and multi-drug resistance when overproduced [8, 11, 20, 35, 53, 61–63]. The overproduction process can be induced through mutations of the local repressor or global regulatory gene, mutation of the promoter region of the efflux pump gene or elements insertions upstream of the transporter gene [48, 51, 64–68]. As well as its homologues MexB, AcrB is known to export a wide variety of antibiotics, β-lactams, antiseptics and toxic compounds including anionic, cationic, zwitterionic and neutral compounds [50, 52, 69–76] directly out of the cells. As part of the RND transporters, AcrB acts as the engine of the efflux pump and the activating energy is provided by a proton transfer from the periplasm to the cytoplasm [53, 59, 63, 70]. The proton flux has been predicted to induce conformational changes in the monomers structure, which have been proposed to follow a functional rotation in three states, designed as *loose* (access), *tight* (bind) and *open* (extrusion) [63]. In this scheme, the drug molecules enters the protein in the L (loose) monomer between the PC1 and PC2 subdomains. Subsequently, the L state change into T (tight) state as consequence of the binding process between the protein and the molecule. In the final step of the functional rotation, the T monomer change its state in O (open), in which the molecule is pushed to the central channel in the TolC-docking domain. During this steps the other monomers, change their states in the same way so that the extrusion of toxic compounds is a continuos process, as described in Figure 1.2.3. Several computational studies a been produced concerning the different steps, from the binding of substrates to their exit [73, 75, 77–81] supporting the experimental results.

The study of AcrB is described in Chapter 3, in which AcrB mutant expressed
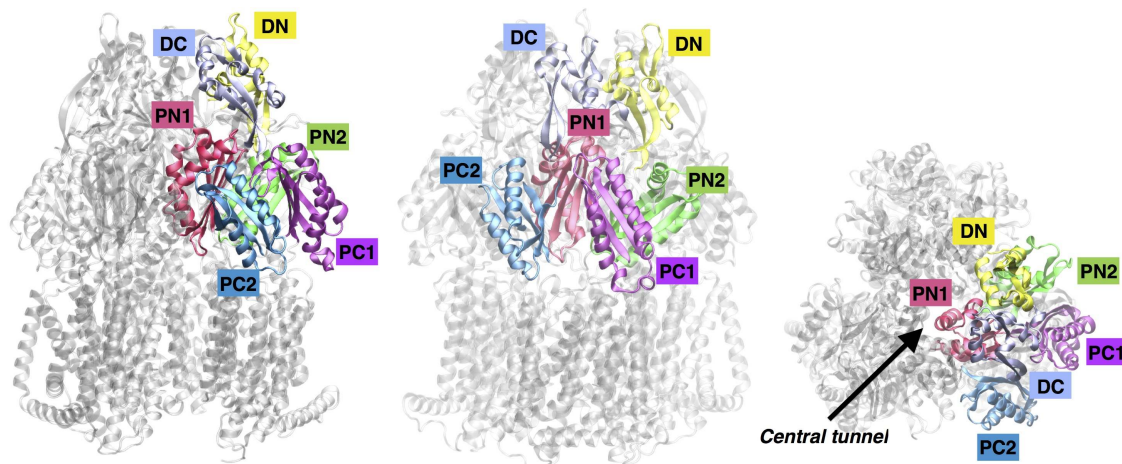
Figure 1.2.2: AcrB division in subdomains: PN1 [light Red] and PN2 [Green] in the amino-terminal part of the pore domain; PC1 [Violet] and PC2 [light Blue] of the carboxyl-terminal part of the pore domain; DN [Yellow] and DC [Ice Blue] in the amino- and carboxyl-terminal part of the docking domain.

in *Salmonella Typhimurium* was shown to increase the resistance to ciprofloxacin by changing the drug binding affinity in the distal binding pocket.

1.2.2 AcrA

The AcrA protein belongs to the membrane fusion protein family (MFP) [82,83] and it is co-expressed with the inner membrane RND transporter AcrB [84]. It is the periplasmic part of the AcrAB-TolC efflux pump and has been proposed to be the adaptor between the RND transporter AcrB and the outer membrane (OM) protein TolC [47,85]. However, its role in the assembly is nowadays reconsidered as more active in the assembly building and in the functionality of the pump [86–89].

AcrA, as most of the MFPs, is an elongated protein that extends from the outer leaflet of the inner membrane deep into the periplasm to meet the OMF TolC [90]. The structure could be divided into four domains (see Figure 1.2.4): the membrane-
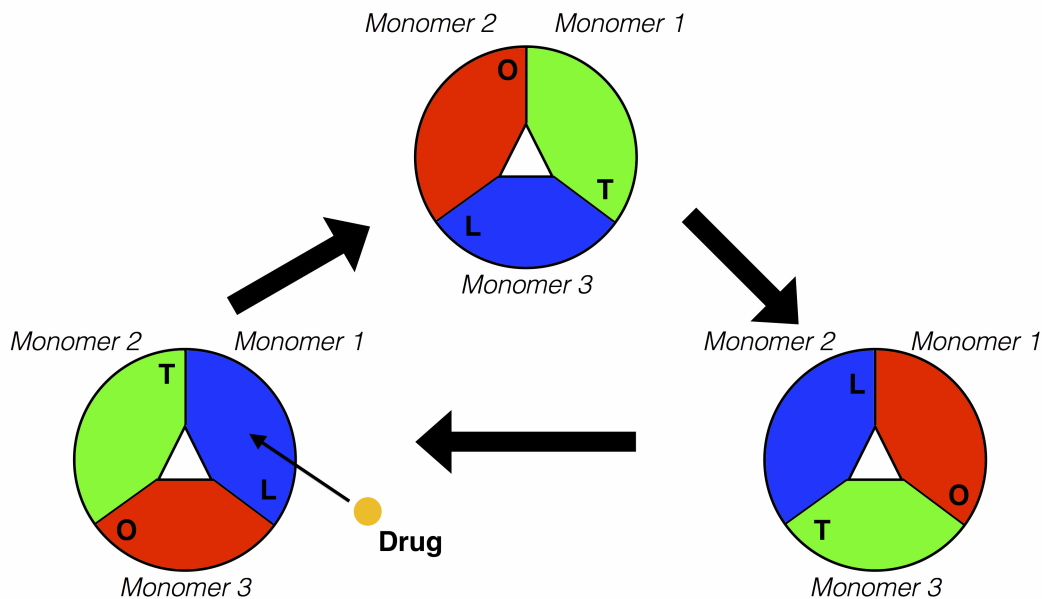
Figure 1.2.3: Scheme of the functional rotation of AcrB monomer. First, the drug molecule enters Monomer 1, which is in the *loose* state [Blue]. Subsequently, Monomer 1 changes its conformation into *tight* state [Green], in which the drug molecules binds the interior of the protein. Finally, Monomer 1 changes into the *open* [Red] conformation, in which the drug is pushed out through the central tunnel. The other monomers (Monomer 2 and Monomer 3) follow the same functional rotation so that the drug extrusion is a continuos process.

proximal (MP) domain, also known as the carboxyl-terminal domain and lacking in the crystal structure [91]; the β-*barrel* domain, which consists of six antiparallel β strands and small α-helix and is predicted to be involved in the AcrA-AcrB binding; the *lipoyl* domain, composed by four β strands and which has been recently proposed to be involved in AcrA dimerization [86, 88, 92] and the α-*haripin* domain which interacts with TolC [93–95]. These domains are connected by mostly unstructured regions which give the protein the flexibility that is believed to be important for functional interactions with the transporter and the OMF.
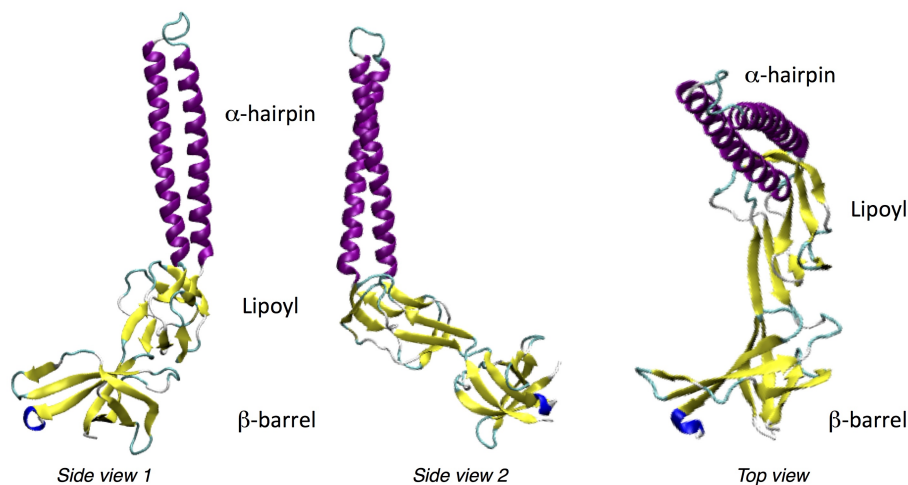
Figure 1.2.4: AcrA structure based on the crystal structure that lacks the MP domain (PDB code: 2F1M). The *α-hairpin* domain is involved in AcrA-TolC interaction while the *β-barrel* domain is supposed to be involved in AcrA-AcrB interaction. The *lipoyl* domain has been recently proposed to be involved in the possible AcrA dimerization.

1.2.3 TolC

The outer membrane channel TolC is the exit duct for the AcrAB-TolC efflux pump although it is able to form functional complexes with inner membrane components that belong to other families like, as example, embark of the MFS superfamily [96] and HlyBD of the ABC family [97]. The first three-dimensional structure of TolC was resolved in 2000 [98]. The structure showed a global cylindrical shape composed by three identical chains of ≈14 nm in length. The upper part of the protein is composed by 12 β-strands, arranged in a 3 nm wide *β-barrel,* which is fully inserted in the outer membrane. The protein extends into the periplasm with an *α-barrel* structure of 14 nm in length. The periplasmic side of TolC is composed by 12 *α-helices* that share similar sequence and inter-helical interactions. In the middle of the *α-barrel* a belt of α- and *β-barrel* forms the equatorial domain as shown in Figure
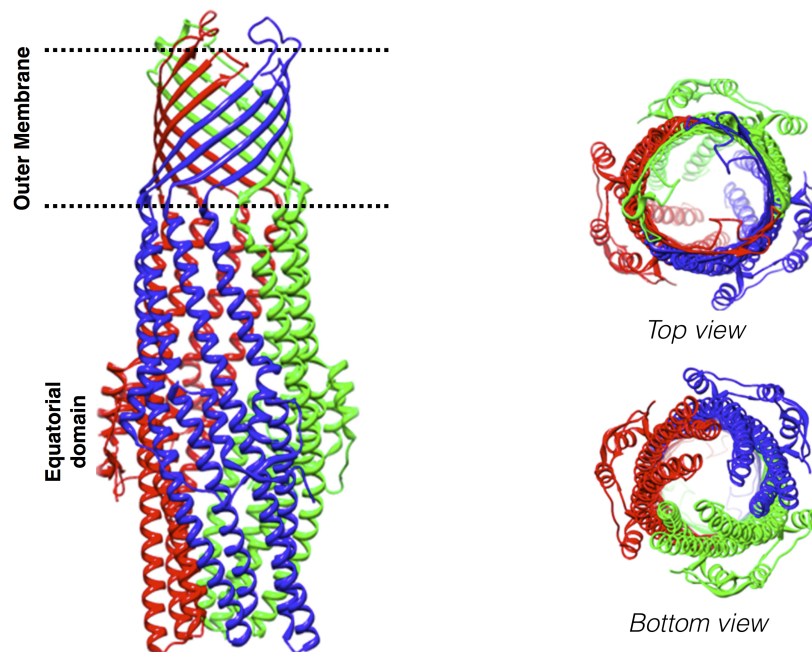
Figure 1.2.5: TolC structure based on the crystal structure (PDB code: 1TQQ). The colors [Blue, Red and Green] identify each monomer. The β-*barrel* is inserted in the outer membrane while the α-*barrel* extends into the periplasm. In the middle of the periplasmic side, α- and β-*barrel* forms the equatorial domain equatorial. Image adapted from [99].

1.2.5.

From the crystal structure analysis the TolC periplasmic side appear to be normally closed by the presence of conserved salt bridges and hydrogen bonds [98, 100–106] but must open in the presence of the other components of the efflux pump. Therefore, it has been suggested that these ionic bridges should disrupt as consequence of the interaction with AcrB [104, 107–109].

1.2.4 Assembly

Although the three-dimensional structure for each component of the pump is available [60, 85, 98, 110], a detailed description of the AcrAB-TolC assembly is still to be reached. In vivo mutageneis studies on the proteins involved in the assembly provided several hints about the possible final structure of the pump. The MFP component AcrA has been proved to interact with the α-*barrel* of TolC in cross-linking and mutagenesis studies [93, 94, 111] and with AcrB [53, 55, 112, 113]. At the same a direct interaction between AcrB and TolC has been proposed based on the crystal structure [110], and then confirmed by several other experiments [103, 109, 114]. Because of these experimental results and because of the symmetry of the RND transporter AcrB and the OM protein TolC, a model with a 3:3:3 stoichiometry was first proposed [60, 103, 115], in which three AcrA proteins were recruited by AcrB and TolC. However, other experiments suggests a possible 3:6:3 stoichiometry, where 6 membrane fusion proteins participate in the assembly building with a more active role in the extrusion process [86–89, 114], as described in Figure 1.2.6. In particular, for the 3:6:3 stoichiometry model, an AcrA hexameric-funnel like structure has been proposed to be inserted between the RND transporter and the OM channel, supported by a comparison with the structure of the MFP CusB, co-crystalized as hexamer with the inner membrane transporter CusA [116].
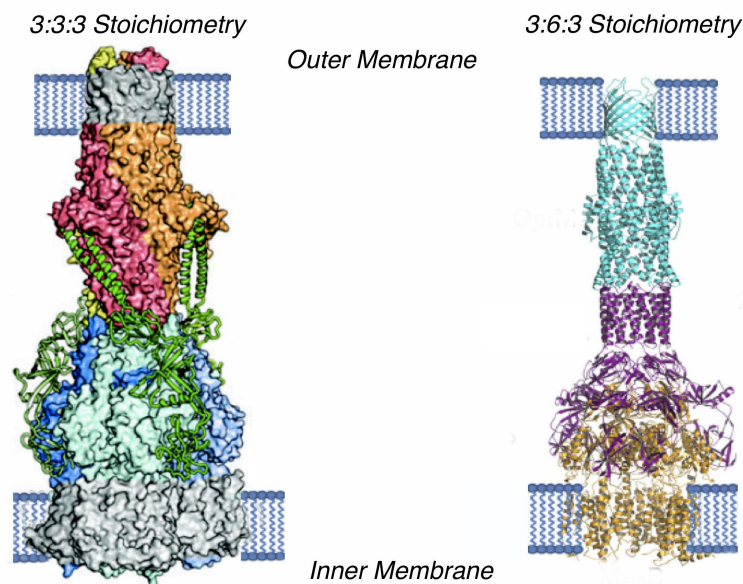
Figure 1.2.6: Most probable stoichiometries for the assembly as suggested by several experiments. In the model with 3:3:3 stoichiometry three MFP components dock the OM protein in the equatorial domain and the OM protein and the RND transporter are directly connected. In the model with 3:6:3 stoichiometry a funnel-like structure is inserted between the RND transporter and the OM protein. The figures are extracted from [115] and [86]

1.3   MexAB-OprM

Psaeudomonas aeruginosa is an opportunistic human pathogen that causes infections with a high mortality rate [35,117–121]. Treatments of *P.aeruginosa* infections of are particularly difficult [122–125] because of its ability to resist a wide variety of unrelated compounds [50,52,69–71,76,126,127]. As part of Gram-negative bacteria, the defense mechanisms of *P.aeruginosa* are similar to *E.coli* with few but crucial differences. As examples, *P.aeruginosa* has an outer membrane with much lower permeability than *E. coli* [128,129]. In addition, to resist to carbapenems, the most com-

mon mechanism is an alteration of the outer membrane porin protein OprD [130–132]. Furthermore, it has been proved to be particularly resistant to fluoroquinolones, i.e. ciprofloxacin, that are regularly used in therapies against *P.aeruginosa* infections. In this case, *P.aeruginosa* can mutate the target sites of fluoroquinolones, by alterations in the enzymes DNA gyrase and topoisomerase IV [133–136]. A number of amino-glycosides are also used in the treatment of *P.aeruginosa* infections but their use is usually associated to resistance development [137]. However, the active extrusion of toxic compounds remains one of the most important process that confers *P. aerugi-nosa* multi-drug resistance [8,138]. The major efflux pump expressed in *P.aeruginosa* is MexAB-OprM, which can extrude the largest range of β-lactams and thus is most frequently linked to β-lactam resistance in clinical isolates [60,98,110,139–141]. Although MexAB-OprM complex shares large similarities with the AcrAB-TolC efflux pump, whose components have been described in the previous paragraphs, a more detailed description of its components is needed to point out the differences and the peculiarity of such pump.

1.3.1 MexB

The RND transporter MexB, similarly to its homologue AcrB in *E.coli*, has been crystallized as an asymmetric homotrimer where each subunit adopts a different conformation representing three states of the transport cycle [141]. It shares high structural similarities (See Figure 1.3.1a) with AcrB, with an overall Cα-RMSD of only 1.4 Å, and a very high closeness in sequence (69.8% identity and 83.2% similarity in a sequence of 1046 amino acid residues). Despite the high similarity, the crystalized L state of MexB appeared to have a more constricted entrance in the PCs subdomains if compared with AcrB (See Figure 1.3.1b and 1.3.1c), presumably
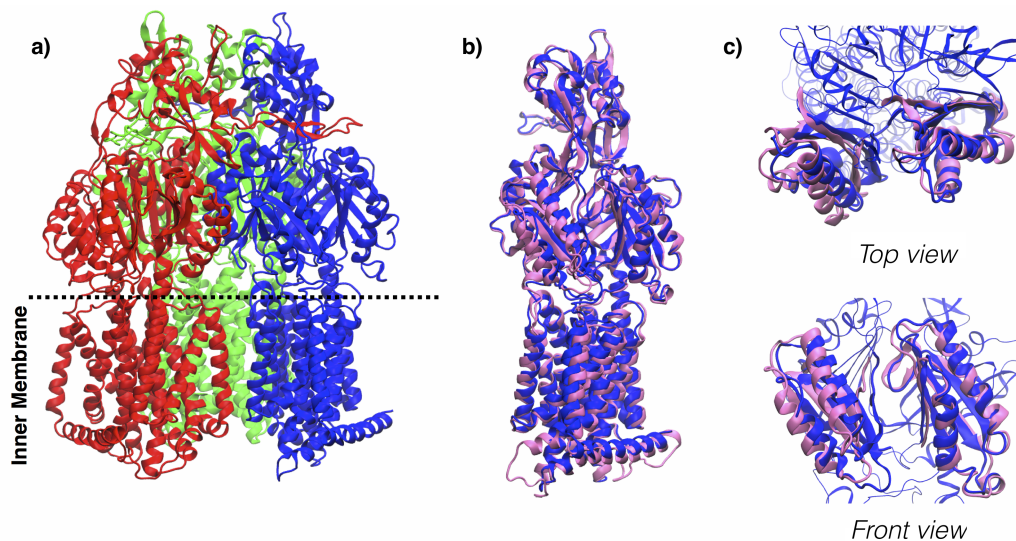
Figure 1.3.1: **a)** Structure of MexB (PDB code: 2V50) in which the different colors [Blue, Green and Red] identify the different conformational state of each monomer. **b)** Comparison of MexB [Blue] and AcrB [Mauve] (PDB code: 2J8S) structures. **c)** Zoom of the structural differences between MexB and AcrB in PC1/PC2 domains

due to a different re-arrangement of the helices in the transmembrane domain [141].

In addition, it was shown that when expressed in *E.coli* the contribution of MexAB-OprM efflux pump differs remarkably from AcrAB-TolC [142] and it is not possible to interchange the two proteins to form hybrid complexes, AcrA-MexB-TolC as example [142, 143].

1.3.2   MexA

The MFP component of the MexAB-OprM efflux pump MexA, like its homologue AcrA has been crystalized as an elongated asymmetric protein [144] in which, it was shown a well-organized membrane proximal domain [55], with two possible orientations with respect to the *β-barrel* domain. In the so-called unrotated state, which was the most frequent among the different chains in the crystal structure, the *MP*

domain is connected to the β-*barrel* domain through an ordered β-ribbon linker.

The state that was designated as rotated showed a conformation where the *MP* domain is globally unaltered but twisted through ≈ 85 clockwise with a different conformation in the β linker [55] as shown in Figure 1.3.2a and Figure 1.3.2b. This state was then proposed to be a possible experimental artifact, as the conformational changes in the linker might have been induced by close crystal contacts [55]. The unrotated state was then proposed to be representative of the functional state of the protein [15, 55]. However, the rotated state was not ignored in literature and it was used to model the possible rotated state of the MP domain in AcrA [145] and proposed to be involved in the MexA dimerization [86, 87] to form the three subunits of the hexameric structure. The building of the homology model for the AcrA MP domain with MexA as template in previous work [145], was justified by the high similarities between MexA and AcrA proteins both in structure and in sequence, as they share 62% and 73% of sequence identity and similarities respectively. However, MexA contains seven residues less than AcrA resulting in length of 89 Å versus 105 Å of AcrA as shown in Figure 1.3.2c. As its homologue AcrA, MexA protein has been proved to interact with the RND transporter MexB through the β-*barrel* domain [146, 147] and with the OM protein OprM through the α-*hairpin* domain [92, 147] in which computational studies revealed an high flexibility of the protein [148]. Furthermore, it was shown through mutagenesis studies that the MexA *lipoyl* domain might be involved in the possible MexA dimerization [92] as shown in Figure 4.0.1.
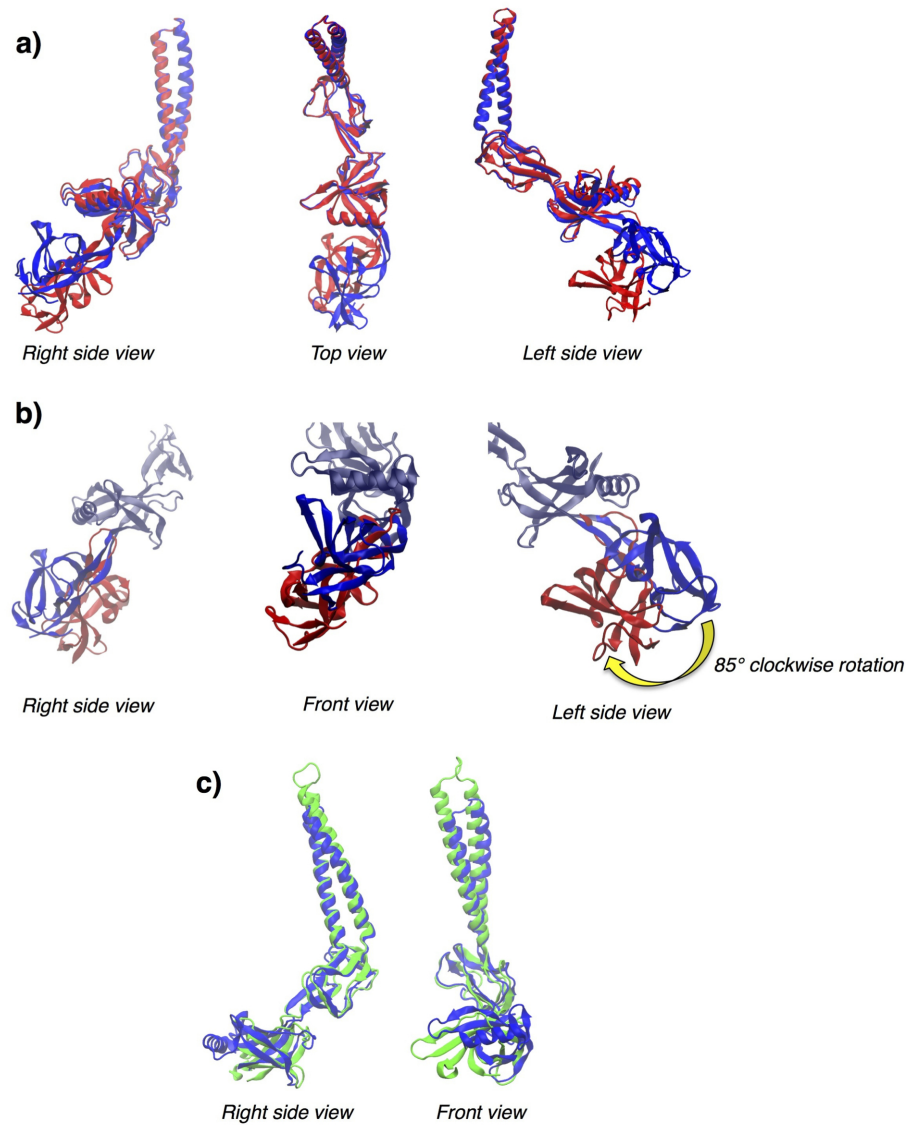
Figure 1.3.2: **a)** MexA re-refined structure (PDB code: 2V4D) with the presence of the MP domain in the unrotated [Blue] and rotated [Red] state. **b)** Focus on the MP domain in the unrotated [Blue] and rotated [Red] state. **c)** Comparison between MexA re-refined structure (PDB code: 2V4D) without the MP domain [Blue] and the AcrA crystal structure (PDB code: 2F1M) in [Red] .

In particular, this result is the starting point for the study of this thesis described

in Chapter 4. Further mutagenesis studies showed also that mutations in the MexB porter domain inhibit the functionality of the pumps but their effects can be partially neutralized by other mutations in the β-*barrel* and MP domains of MexA [146, 147]. These results confirm the interactions between MexA and MexB that might guide the assembly building.

1.3.3   OprM

The OprM protein is the outer membrane component of the MexAB-OprM efflux pump, in which it is assumed to serve as a drug discharge duct [139]. It shares with its homologue TolC in *E.coli*, a relatively low values of similarity (40%) and identity (21%) in sequence, but a very high structural similarity, with a Cα-RMSD of 1.6 Å. The OprM overall trimeric structure exhibits three main domains: the membrane-anchoring β-*barrel*, the periplasmic cavity-forming α-*barrel* and the equatorial domain, as described in Figure 1.3.3. The β-*barrel* structure is inserted in the outer membrane and has a diameter, in the crystal structure, of 6-8 Å, which is not enough for the antibiotic exit [139]. .
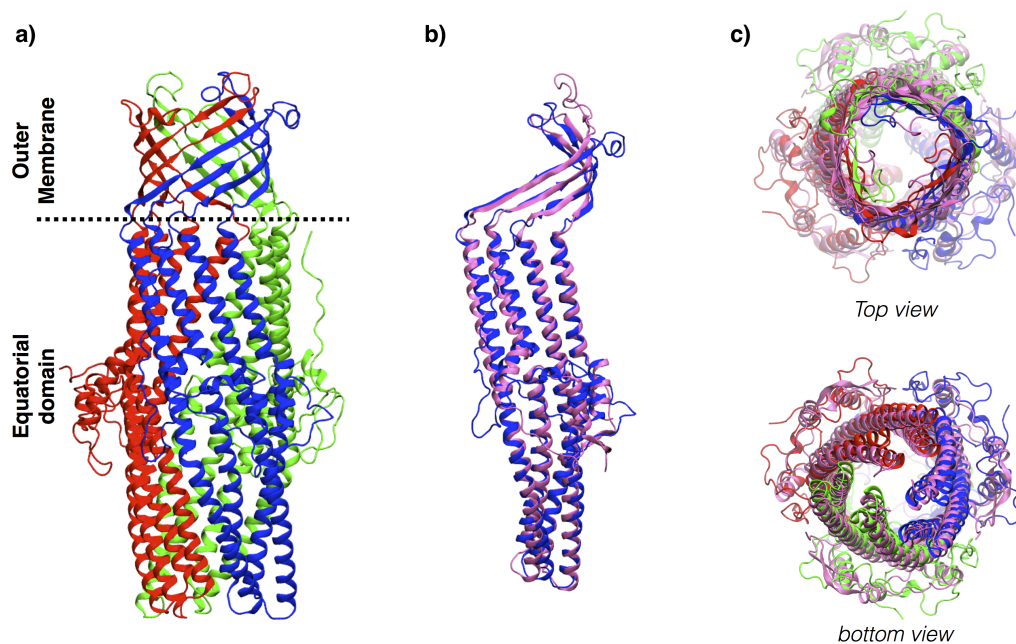
Figure 1.3.3: **a)** OprM crystal structure (PDB code 1WP1), in which different colors refer to different monomers. As it homologue TolC in *E.coli*, the *β-barrel* is inserted in the outer membrane while an *α-barrel* extends into the periplasm. In the middle of the periplasmic side, *α*- and *β-barrel* forms the equatorial domain equatorial. **b)** Comparison of OprM [Blue] and TolC [Mauve] (PDB code 1TQQ) monomer. Despite the low similarity in sequence, the two structures show high structural similarities. **c)** Comparison of upper and bottom parts of OprM and TolC.

1.3.4 Assembly

Because of many similarities with AcrAB-TolC both the assemblies have been predicted to have the same stoichiometry. However, which stoichiometry may represent the real structure of these efflux pumps is still on debate. Although different experiments suggest a 3:3:3 stoichiometry [60, 115, 143], many others suggest a stoichiometry in which six MexA proteins participate in the assembly [86–88, 92, 114]. In the latter case, a prior dimerization of the membrane fusion protein seems a pre-requisite to

form the MFP hexameric structure that might be inserted between the RND transporter MexB (and AcrB) and the OM protein OprM (and TolC) [86,88,99] as shown in Figure 1.2.6. The effectiveness of the dimerization process is described in Chapter 4 of this thesis.

1.4   Thesis outline

This thesis aims to investigate the structural properties of the membrane fusion protein MexA in the MexAB-OprM efflux pump and to gain insight into the RND transporter AcrB inhibition process.

The study described in the first part of the thesis concerns the first case in which a mutation on RND efflux pump is related to an increased multi-drug resistance in clinical settings. The work is a result of experimental and computational study, in which the effects of the mutation in the binding and recognition of different substrates are investigated.

In the second part of the thesis, we investigated the role and the possible stoichiometry of MexA proteins in the MexAB-OprM assembly. In particular, we studied the effect of a mutation that is known to impair the functionality of the pump, presumably by affecting MexA multimerization. Following different strategies, we showed the propensity for MexA (wt only) to form stable dimeric units. This result, could help to gain insights on MFP stoichiometry and, thus, role in the whole assembly.

Chapter 2

Computational Methods

The constant development of computer technology in the last decades increased considerably the accuracy of the computational description of biological world [149–160]. In fact, it is nowadays common to refer to computer simulations as *in silico* experiment as they allow scientists to monitor chemical reactions, conformational changes and other thermodynamics quantities of the systems under study. Statistical mechanics provides the theoretical framework that connects real-world experiments and computer simulations through a relation between ensemble averages, over the microscopic configurations explored by the system, and macroscopic observables.

The work of this thesis is based on computer simulations and has the aim to describe, at atomistic level, the mechanisms involved in efflux pumps inhibitions and in efflux pumps building. Therefore, in this chapter will be presented an overview of the computational methods used in these studies such as Molecular Dynamics (MD) simulations, homology modeling and protein-protein docking, and the theoretical background that support them.

2.1 Molecular Dynamics

MD simulation is one of the principal tools to study biological molecules. In fact, the MD technique aims to mimic real experiments, by following the same approach to study the samples which, in this case, consist on N interacting particles. As well as in real experiments the samples are equilibrated, by solving the equations that describe the motion of each particle of the systems. After the equilibration, the actual measurements are performed.

From the very first algorithm, developed in the late 1950s to investigate collisions of hard spheres [161,162], MD simulations progress led to the first description of liquid water [163,164] and proteins [165]. Nowadays, several MD simulations programs [154, 166–169] are routinely used with constantly improved hardwares [150, 152, 153, 155–157, 159, 170] to study solvated [145, 148, 171] and membrane proteins [?, 73, 77, 172], or to validate homology models for proteins whose three-dimensional structure is still missing.

2.1.1 Theoretical framework

The theoretical basis of MD simulations to describe biological world lies in statistical mechanics, which provides mathematical expressions that relate microscopic properties such as atomic positions and velocities, computed in the simulations, to macroscopic observables such as pressure, energy, etc. In order to describe how statistical mechanics connects the macroscopic world to its microscopic components, few definitions should be first recalled.

The *thermodynamic state* of a system is defined by the set of its macroscopic parameters, like temperature T, pressure P and number of particles N, while the *microscopic state* is defined by the positions, r, and momenta, p of all the particles

of the system. The collection of these phase space points that belong to a particular thermodynamic state is called *ensemble* and corresponds to the possible conformations of the system. The macroscopical observables are then defined as ensemble averages as described in equation 2.1.1.

$$\langle A \rangle_{ensemble} = \int \int dp^N dr^N A\left(p^N, r^N\right) \rho\left(p^N, r^N\right), \qquad (2.1.1)$$

where $A\left(p^N, r^N\right)$ is a function of the position r and momenta p and represents the observable of interest, re-weighted by the probability density

$$\rho\left(p^N, r^N\right) = \frac{1}{Q} \exp^{\left[-H\left(p^N, r^N\right)/K_B T\right]}, \qquad (2.1.2)$$

where H is the hamiltonian of the system, T the temperature, k_B the Boltzmann constant and Q is the partition function

$$Q = \int \int dp^N dr^N \exp^{\left[-H\left(p^N, r^N\right)/K_B T\right]}. \qquad (2.1.3)$$

As the integral over the phase space is extremely difficult to solve, MD simulations calculate each point of the phase space step by step. Therefore, to calculate the ensemble average properly the simulations must explore all the possible states accessible by the system. This means that, in principle, the trajectories need to be long enough to explore the whole phase space. However, such condition is not always guaranteed, even with multiple MD simulation [173].

The way MD simulations determine the observable values is, however, trough a time average of A as described in the following equation

$$\langle A \rangle_{time} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_{\tau=0}^{\tau} A\left(p^N(t), r^N(t)\right) dt \approx \frac{1}{M} \sum_{t=1}^{M} A\left(p^N, r^N\right), \qquad (2.1.4)$$

where M is the number of time step in the simulation, t is the simulation time and $A\left(p^N, r^N\right)$ is the value of A at each step. To relate the time averages provided by MD simulations and the ensemble averages that define the experimental observables a strong hypothesis is to be made, the *ergodic hypothesis* which can be resumed by the equation 2.1.5.

$$\langle A \rangle_{ensemble} = \langle A \rangle_{time}. \qquad (2.1.5)$$

The idea that support the ergodic hypothesis is that, if the system is free to evolve indefinitely in time it will pass through all the accessible states. To preserve the validity of such hypothesis is thus necessary that MD simulations generate enough conformations that might be representatives of all the states of the system.

As the macroscopic experiments might be performed in different conditions so do MD simulations, and the choice of the ensemble does make a difference when computing the observables averages [174]. However, techniques exist to relate fluctuations in different ensembles [175] allowing a comparison of the results of MD simulations performed in different conditions. Several algorithms have been developed to mimic the different ensembles properties, such as the Andersen, Nosé-Hoover, Berendsen and Langevin dynamics algorithms [176–180]. With this methods is thus possible to perform MD simulations in: **NVE** *(micro-canonical)* ensemble, in which the number of particles N, the volume V and the total energy E are constant; **NVT**, one of the main ensembles used in MD simulations, with conserved number of particles N,

volume V and temperature T; $\boldsymbol{NPT}$ with constant number of N particles, pressure P and temperature T.

2.1.2 Classical Mechanics

In the previous paragraph it was introduced the theoretical basis of the MD simulations method. The aim of this paragraph is to describe how the instantaneous value of $A\left(p^N, r^N\right)$ are computed through the evaluation of the forces that act on each particle of the system. To efficiently compute the observables values, it is necessary a first approximation to neglect the subnuclear and nuclear interactions among the particles, which can be separated from the electronic interactions in the so-called Born-Oppenheimer approximation. Since the energy scale for sub- and nuclear interactions is way beyond the biological interactions energy scale, in most of the circumstances the quantum nature of the nuclei can be safely neglected and they can be considered as classical particles that move in an effective potential. In this approximation, the effective potential is due only to the quantum nature of the electrons in the so-called Ehrenfest dynamics scheme. Therefore the major task of quantum mechanics is, in the Born-Oppenheimer approximation, to find the solution of the Scrödinger equation for the electrons. However, the solution of such equation is computationally demanding and, as additional limitation, only relatively small system can be simulated. This led to a more heuristic way to study macromolecules systems, and the developments of semi-empirical effective potentials, which are much faster to integrate. These potentials will be described later in this Chapter.

In this classical approximation the particles follow the Newton's equation of motion 2.1.6, which relates the sum of the forces that act on a particle i with mass m and its acceleration $\partial^2 \mathbf{r_i(t)}/\partial t^2$.

$$\sum_j F_{ij}(t) = m_i \frac{\partial^2 \mathbf{r}_i(t)}{\partial t^2} \qquad (2.1.6)$$

However, if the classical approximation considerably simplifies the physical description of the system, it makes theoretically impossible to find an analytical solution for equation 2.1.6. In fact, such solution for a system of N interacting particles is possible only for $N = 2$. The first step to move around this limit is to assume that the force $F_i(t)$ that acts on particle i at the time t is constant for a Δt time interval, so that the new particle position $r_i(t + \Delta t)$ and velocity $v_i(t + \Delta t)$ can be computed. Several algorithm based on Taylor series have been developed to integrate the equation of motion [181–185]. An example is reported in equation 2.1.8 that describes the *verlet leap frog* algorithm, that evaluates the positions at time $t + \frac{1}{2}\Delta t$ by using the velocities at time $t - \frac{1}{2}\Delta t$.

$$\mathbf{F}_i = \text{-}\nabla U_i, \ \ i = 1 \ldots N \qquad (2.1.7)$$

$$v_i\left(t + \frac{1}{2}\Delta t\right) = v_i\left(t - \frac{1}{2}\Delta t\right) + a(t)\Delta t \qquad (2.1.8)$$

$$r_i(t + \Delta t) = r_i(t) + v_i\Delta t\left(t + \frac{1}{2}\Delta\right). \qquad (2.1.9)$$

The second step to move around the limitations of the N-particles system is to parametrize the resultant force F_i in equation 2.1.6 with the so-called *force field*, where the potential energy of the system (derived from equation 2.1.7) is described as sum of semi-empirical terms that describe all the interactions between the particle i and the the other particles j of the system, as shown in equation 2.1.10.

$$U_i \quad = \quad \sum_j U_{i,j}^{bonds} + \sum_j U_{i,j}^{angels} + \sum_j U_{i,j}^{dihedrals} + \qquad (2.1.10)$$
$$+ \quad \sum_j U_{i,j}^{electrostatic} + \sum_j U_{i,j}^{VdW}$$

The terms in equation 2.1.10 that contribute to the total energy of the system can be further divided into *bonded* and *non − bonded* interactions, described in the next paragraphs with more details.

Bonded Interactions This class includes the contributions of the *bond, angles* and *dihedral* interactions and each term has the following functional form [186]:

$$U^{bonded} \quad = \quad U^{bonds} + U^{angels} + U^{dihedrals} \qquad (2.1.11)$$
$$U^{bonds} \quad = \quad \sum_{bonds} k_r \left(r - r_{eq} \right)^2 \qquad (2.1.12)$$
$$U^{angles} \quad = \quad \sum_{angles} k_\theta \left(\theta - \theta_{eq} \right)^2 \qquad (2.1.13)$$
$$U^{dihedrals} \quad = \quad \sum_{dihedrals} \frac{V}{2} \left(1 + \cos \left(n\phi - \gamma \right) \right) \qquad (2.1.14)$$

While a more realistic description for the bonds interaction is provided by the Morse-potential, for small deviations from the equilibrium distance r_{eq}, it can be approximated to the harmonic interaction in equation 2.1.12, as described in Figure 2.1.1 and Figure 2.1.2a. In this approximation the interacting particles are described as rigid bodies, connected by a spring with an empirical k_r constant. Analogue to bonds, angle θ between two particles also fluctuates around a reference value as
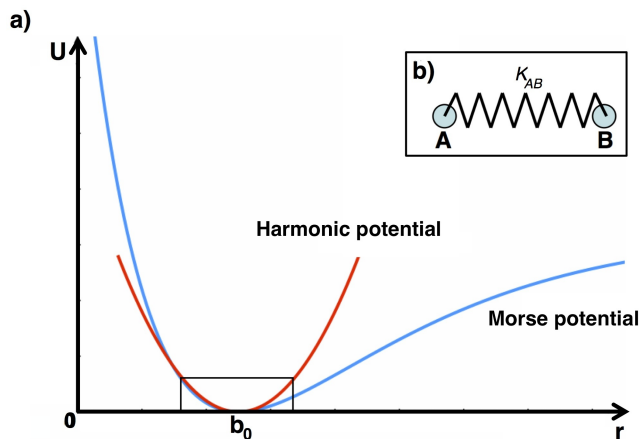
shown in Figure 2.1.2b.

Figure 2.1.1: **a)** The approximation of the Morse potential [Blue] with the harmonic potential [Red] for small deviations from the average distance b_0; **b)** The harmonic interaction between particle A and particle B with a spring with a k_{AB} constant.

The potential energy associated with this motion is an harmonic function, with k_θ as spring constant. The torsional motion along a covalent bond is taken into account in the dihedral potential function, where ϕ (see Figure 2.1.2c) is the angle between the ijk and jkl planes, n is the number of the minima and γ is a phase angle. In addition to the last term of the bonded interactions, eq 2.1.15, describes the improper dihedral contribution, where the angle ξ is chosen to keep planarity in a molecular structure.

$$U^{Improper\,dihedrals} \quad = \quad \sum_{dihedrals} \frac{1}{2} k_\xi \left(\xi - \xi_{eq}\right)^2 \qquad (2.1.15)$$

Non-bonded Interactions Non-bonded interactions are described by the electrostatics and Van der Waals (VdW) interactions. The Coulomb potential describes
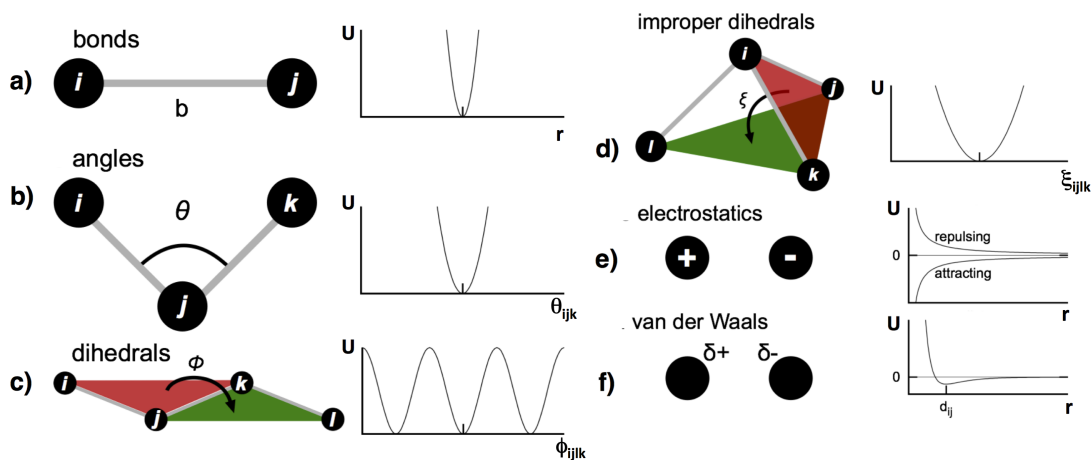
Figure 2.1.2: **a-c)** Bonded interactions for the bonds, angle and proper dihedral terms. **d)** Bonded interaction for improper dihedral contributions. **e-f)** Non-bonded interactions terms described by the Coulomb and the Lennard-Jones potentials

the attractive and repulsive interactions between charged atoms at distance r_{ij} (see Figure 2.1.2e) as described in equation 2.1.16, where the ϵ_0 is the vacuum dielectric constant. Electrostatic interactions are, however, possible also between uncharged atoms, whose asymmetric charge distribution may form transient dipoles as shown in Figure 2.1.2f. The dipoles interacts thought the so-called Van der Waals potential, parametrized in the Lennard-Jones potential in the equation 2.1.17, where the $\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}$ is the repulsive long-range term and $\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}$ is the attractive short-range term, ϵ is the depth of the potential well and σ_{ij} is zero-potential distance between particle i and j.

$$U^{electrostatic} = \sum_{i<j} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \qquad (2.1.16)$$

$$U^{VdW} \quad = \quad \sum_{i<j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \qquad (2.1.17)$$

Force Fields Several force fields have been designed to efficiently compute the potential energy terms described above such as the force fields in the AMBER package [187, 188], used for the simulations in this thesis [186, 189, 190], the force fields in the CHARMM package [169, 191, 192], GROMOS [193–195] and MARTINI, developed as coarse-grained force field [196]

2.1.3 Periodic Boundary Conditions (PBC)

Even in the best laboratory conditions, real experiments are always influenced by the surrounding boundaries. In most cases such influences can be neglected or taken into account as systematic errors. In the computational world the simulated experiments are limited by the simulation box that contains the system. Since the box is not infinite, the edge effect will influence the evolution of the simulation. To avoid this limitation, the so-called Period Boundary Conditions (PBS) have been developed. The idea is to simulate an infinite system, composed by the copies of the finite initial box. The result is a three-dimensional system where identical images are arranged in a periodic array. In this extended box the atoms can leave the original simulation box on one side and enter the simulation box as periodic image on the opposite side as described in Figure 2.1.3. As consequence, the total number of particles in the system is conserved. One drawback of the PBC method is the possible interaction, e.g. electrostatic, between the system under study and its periodic images [197]. In order to avoid such interactions a proper chose of the simulation box is needed.
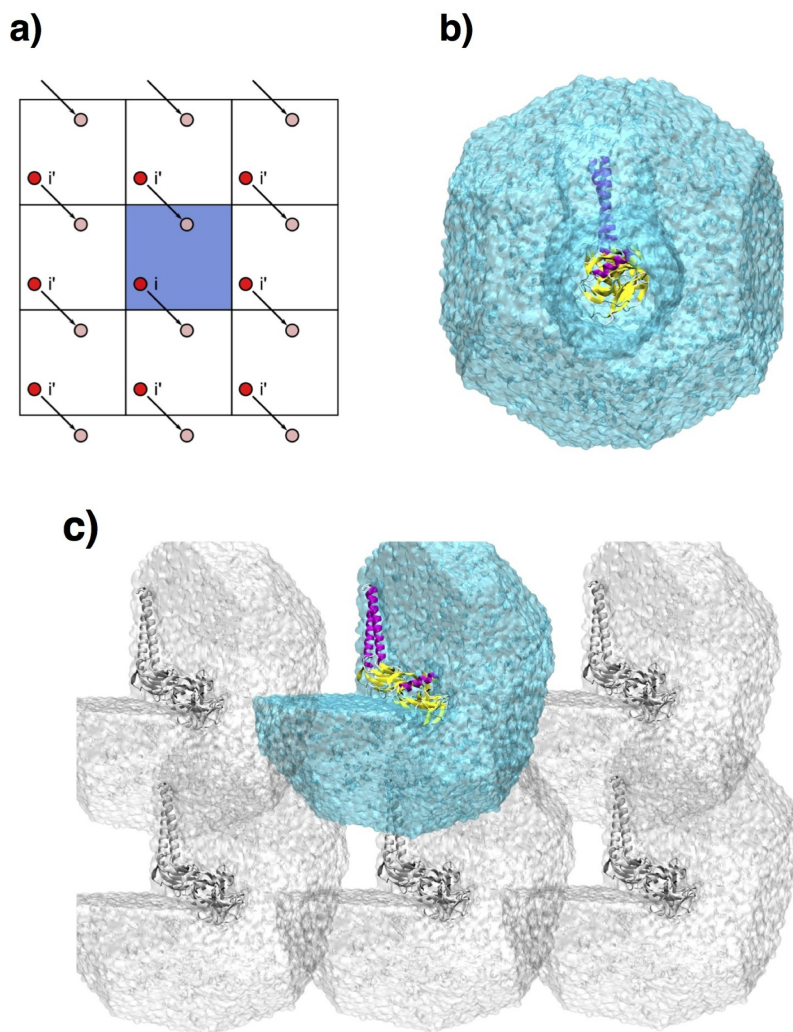
Figure 2.1.3: **a)** Two dimensional representation of periodic boundary conditions (PBC). Atoms that leave a simulation system a side enter the original system on the opposite side. **b)** Truncated octahedral cell box for MexA protein. The water molecules close to the protein have been removed to highlight the protein. **c)** Section of periodic images (in gray) of the initial octahedral cell box. The truncated octahedron cell has been chosen to reduce the number of water molecules to be simulated, maximizing at the same time the distance between the protein images.

However, if a very large box would be potentially the best choice it is necessary to

consider that the higher the number of the atoms the slower the simulation is. Another point that should be taken into account in the PBC context is how long-range interactions can be computed during the simulation of an infinite three-dimensional periodic system.

As example, the Coulomb potential in equation 2.1.16 might include interactions between images that should be avoided. Since a simple truncation of the potential is not allowed [174, 198] it became essential to find a way to efficiently compute long-range interactions in the PBC context. A way to take into account the long-range interaction is through Ewald summation and the Particle Mesh Ewald (PME) methods [199–202]. In a system of N charged particles located in a cube of volume L^3, under PBC, the Coulomb potential in equation 2.1.18 can be written as

$$U^{electrostatic} \quad = \quad \frac{1}{2} \sum_{\mathbf{n}} \sum_{i,j}^{N} \frac{q_i q_j}{4\pi\epsilon_0 |\mathbf{r}_{ij} + \mathbf{n}L|} \tag{2.1.18}$$

where the sum is over all periodic images $\boldsymbol{n}$ and over all the particles (it is assumed that particle i interacts with all its periodic images). The equation 2.1.18 however, has been proved to converge very slowly and thus cannot be used to efficiently compute electrostatic interactions. The idea behind the PME method is to split the Coulomb potential into two parts by using the identity

$$\frac{1}{r} \quad = \quad \frac{f(r)}{r} + \frac{1 - f(r)}{r} \tag{2.1.19}$$

where $f(r)$ is a generic function of the distance r between the particles. The choice of the function $f(r)$ has been proved to be crucial for the efficiency of the

method, although different choices yield to comparable results [203]. The identity 2.1.19 is introduced to separately consider the two main complications of the Coulomb potential, i.e. the large variation at small r and the slow decay for large r. Therefore, the first term of the identity 2.1.19 should be negligible beyond a cutoff r_{max} in order to avoid the divergec when $r \to 0$. At the same time, the second term should vary slowly for all r so that its Fourier transform could be represented by limited number of reciprocal vectors. An example of one traditional function [204] that fits these requirements is the function $2\pi^{-1/2} \int_r^\infty dt \exp\left(-t^2\right)$.

2.2 Homology Modeling

To efficiently use MD simulations in the study of bimolecular systems it necessary a proper initial description of the system under study. As example, since proteins functionality is strongly related to their three-dimensional structure [205–207], a bad initial model description may lead to largely different results [208]. Therefore several experimental techniques have been developed to determine the structure of proteins, such as X-ray crystallography [209, 210], Nuclear Magnetic Resonance (NMR) spectroscopy [211–213] or Electron Microscopy (EM) [214–217]. However an experimental description of the three-dimensional structure for all the proteins in nature is very far to be reached, since only a small fraction of them has been structurally determined to date [218–220]. To fill this gap in the proteins structure knowledge computational methods have been developed in recent years.

One of the main computations approaches is the so-called *homology* (or comparative) *modeling* that predicts three-dimensional structure of proteins. To accomplish this role the modeling procedure requires at least one reference structure that have been previously obtained with experiments [220–224]. To predict a realistic structure

the target and the template proteins should have a significant amino acid sequence similarity [225, 226], where the term *significant* refers to the definitions [226] of the (*i*) easily detected relationship (>30% sequence identity), (*ii*) the twilight zone [227] with statistically significant sequence similarity (10%-30% sequence similarity) and (*iii*) the "midnight zone" [227] of statistically insignificant sequence similarity. Independently of the method used to align sequences, searching in the midnight zones of the sequence-structure relationship often results in false negatives, false positives, or alignments that contain an increasingly large number of gaps and alignment errors [226]. Among different programs that have been developed, the MODELLER package [221, 223, 228] was used to build the proteins structures studied in this thesis.

After the sequence alignment, the MODELLER first step of model building is to compute distance and dihedral angle restraints for the target sequence, derived from the template three-dimensional structure using a database of alignments [229]. Relationships such as the correlations between two equivalent Cα-Cα distances from two related proteins are evaluated and expressed through probability density functions (pdf) and are used as spatial restraints. Then, stereochemical restraints, like bond length and bond angle preferences are extracted from the Charmm22 force field [230] or from the so-called homology-derived restraints analysis, in which restraints are calculated from template protein structures (see Figure 2.2.1 for a schematic description of the modeling process).
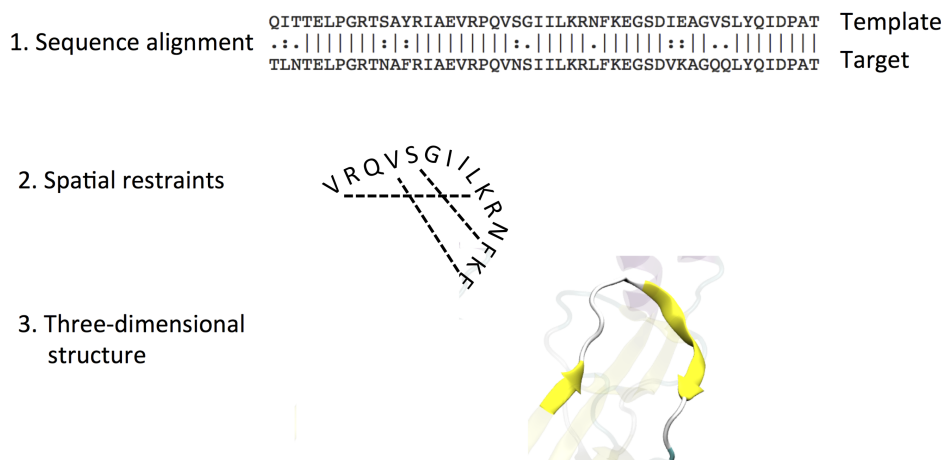
Figure 2.2.1: Main steps in the modeling process with MODELLER. First, the target and the template sequences are aligned. Second, spatial restraints such as Cα–Cα distances, hydrogen bonds, dihedral angles, etc. are transferred from the templates to the target. Finally, the three-dimensional model is obtained by satisfying all the restraints as well as possible

The result of this procedure is a three-dimensional structure for the target protein. If multiple structures are obtained from the homology modeling, the best model can be selected with the lowest value of discrete optimized protein energy (DOPE) scoring function [231]. However, the DOPE score index is not an absolute measure that validates the model but can be used only to rank the models obtained from the same alignment.

2.3 Protein-protein docking

Many biological processes are mediated by protein–protein interactions that lead to a macromolecular assemblies [232–234]. Thus, an understanding of such interactions could deepen the knowledge of the assemblies building process and their

functionality. However, experimental determination of protein-protein complexes is generally more difficult than the determination of each component. In the last decades, the developments in the computational biology led to several methods to predict possible assemblies structures, coded in computational programs like, as example, AUTODOCK, DOCK, GOLD, SwissDock and ATTRACT [235–241]. With the large number programs available it is possible to simulate protein-protein docking with high accuracy [242] using a wide range of options like, rigid-body docking, flexible docking with normal mode computations or ensemble of conformations, etc. For this thesis, the program used to study protein-protein docking was ATTRACT and its description is in the following paragraph.

ATTRACT The protein-protein docking program, ATTRACT has been developed in 2003 by *M. Zacharias* [240] and continuously updated [241, 243–246]. The ATTRACT approach is based on energy minimization in rotational and translational degrees of freedom of two or more protein partners. In case of two proteins-docking a starting point grid for the ligand is generated around the receptor as described in Figure 2.3.1. For each starting point, various orientations of the ligand are then generated. The docking from each starting position is based on energy minimizations of the ligand protein with respect to the receptor during their approach. During the docking process, ATTRACT includes the flexibility of the proteins by including the normal modes analysis or an ensemble of different conformations. Recently, new approaching methods have been introduced (like "fixed receptor and free ligand, with gravitational attraction to the receptor" or "free receptor and ligand with a common gravitational center") and the starting point grid can be replaced with a grid representation of both the ligand and the receptor. Together with the parallelization of the code and the possibility to drive the docking with Cryo-EM maps, these innovations

remarkably increased the velocity of the docking analysis.

Front view   Left side view   Right side view
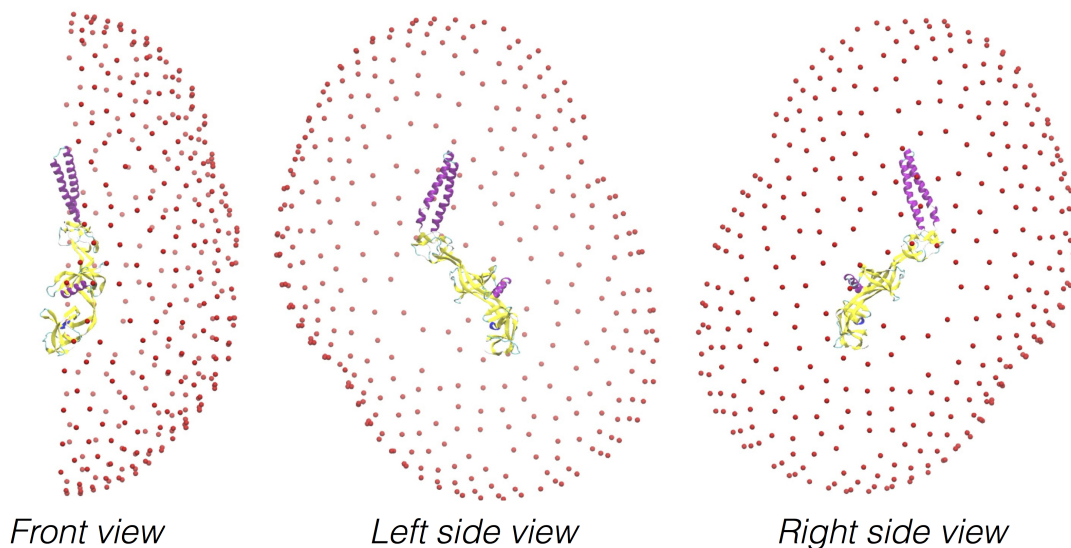
Figure 2.3.1: Section of the starting point grid [Red] for the protein-protein docking with ATTRACT for the Membrane Fusion Protein MexA. The distance between the points is larger than the maximum distance from the receptor atoms and the ligand center.

Reduced protein model   The ATTRACT package employs a reduced model for the proteins, between the full atomic and the residue-based representation. The largest residues are represented by four pseudo atoms, two for the backbone and two for the side-chain, while the side-chains of small residues are identified by on pseudo atom, in the side-chain geometric center. The potential energy is described by a soft Lennard-Jones function and an electrostatic interaction as shown in Figure 2.3.2

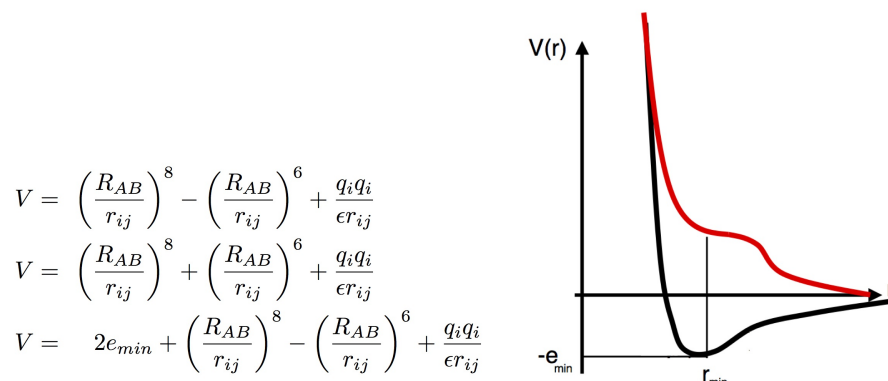$$V = \left(\frac{R_{AB}}{r_{ij}}\right)^8 - \left(\frac{R_{AB}}{r_{ij}}\right)^6 + \frac{q_i q_i}{\epsilon r_{ij}}$$

$$V = \left(\frac{R_{AB}}{r_{ij}}\right)^8 + \left(\frac{R_{AB}}{r_{ij}}\right)^6 + \frac{q_i q_i}{\epsilon r_{ij}}$$

$$V = 2e_{min} + \left(\frac{R_{AB}}{r_{ij}}\right)^8 - \left(\frac{R_{AB}}{r_{ij}}\right)^6 + \frac{q_i q_i}{\epsilon r_{ij}}$$

Figure 2.3.2: Soft Lennard-Jones potential in ATTRACT. In case of an attractive pair [Black] an r^{-8}/r^{-6} Lennard-Jones-type potential is used. For a repulsive pair [Red] the energy minimum is replaced by a saddle point.

2.4 Analysis methods

MD simulations produce as result trajectories of atomic positions (and optionally velocities and energies) as a function of time. Several methods are available to extract meaningful information from the MD trajectories. In this paragraph will be described the main analysis performed for the work of this thesis.

RMSD analysis

The evaluation of the Root Mean Square Deviation (RMSD) is in most cases the first analysis that gives an overview of the global behavior of the system during the MD simulation. Although it cannot be used to estimate the convergence it is useful to check wether the simulation has *not* reached the convergence [247]. To evaluate the RMSD between two structures, the first structure is superimposed to the reference and the RMSD value is computed as described in equation 2.4.1. Generally $C\alpha$ atoms, backbone atoms or heavy atoms are applied as reference atoms in this

analysis.

$$RMSD\left(t_0, t_n\right) \;=\; \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left\|r\left(t_0\right)-r_i\left(t_n\right)\right\|^2} \qquad (2.4.1)$$

A more accurate description of the global behavior of the system is provided by the RMSD-matrix, where each element (i,j) of the matrix consists of the RMSD value between the structure of the protein at time i and j: the lower the value the highest the similarity of the structures (special case for diagonal elements (i,i) with zero value of the RMSD).

Protein compactness

This analysis is mainly based on the Radius of Gyration (RoG) analysis of the structures during the MD simulations. It is defined as the RMS distance between the center of the protein and its ends as described in equation 2.4.2, where r_{cm} is the position of the center of mass of the system and r_i the coordinates of the i-th atom. It is a useful index of the compactness of the protein structure [248]. Other analysis were performed to support the RoG results, such as the inter-domain distance of the protein structure and the distribution of the Center of Mass (CoM) of each domain. These analysis were performed with both CPPTRAJ and home made scripts and were performed in order to gain more information about the relative motion of each domain with respect to the others.

$$R_g \;=\; \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|r_i-r_{cm}\right\|^2} \qquad (2.4.2)$$

Cluster analysis

An useful way to study the MD simulation trajectory is to group molecular configurations into subsets based on the similarity of their conformations [249,250]. Clustering algorithms collect data elements (points) in sets called clusters using a function that measures the distance between pairs of points. The points among in each cluster are more similar to each other than to points of other clusters and thus, the size of each cluster are likely to be different as higher energy substates will be less populated than lower energy substates. Several algorithms have been developed to perform cluster analysis such as hierarchical, centripetal, complete-linkage, centroid-linkage, average-linkage, means, Bayesian and COBWEB. Among this variety of methods in this thesis was used the average-linkage algorithm in the CPPTRAJ package [251], in which the distance between two clusters is defined as the average distance between each point in one cluster and every point in the other clusters. The average-linkage algorithm was selected among the others as it was proved to be one of the most useful algorithms for such analysis [250].

Secondary structure analysis

The protein structures studied in this thesis are the products of homology modeling using crystal structures as templates and the analysis of the secondary structure during the MD simulation allows the identification of possible sites of instabilities. Furthermore, as the MD simulations for the systems aimed to point out the effects of specific mutations, the secondary structure analysis provide a powerful tool to identify changes in the protein structures. The analysis of the secondary structure in this thesis were performed with the CPPTRAJ analysis tool.

Bonds analysis

During the MD simulation the residues of the proteins can interacts to form or break new bonds that can confer stability or flexibility to the system. The analysis of the bonds evolution focused on the H-bonds and Salt bridges. The first family describes the bond between two residues that share an hydrogen atom. The analysis was performed with CCPTRAJ [251] with a distance cut-off of 3.0 Å and an angle cut-off of 135 degrees. The second family describes the bonds between any oxygen atom of acidic residues and the nitrogen atoms of basic residues and, in this case, VMD [252] was used to analyze possible inter-domain Salt bridges with a distance cut-off of 3.2 Å.

Radial Distribution Function

Water plays a crucial role in the structure, dynamics and function of a protein. Depending of the interactions between water molecules and protein, water can be broadly classified into three categories: bulk water, around the protein molecules at greater distance than the van der Waals range; bound water, which forms hydrogen bonds with charged or polar protein atoms inside the protein; hydration water, which interacts with the protein surface. The hydration water contributes to many properties of the protein such as solubility, drug docking and protein folding. Thus, the study of possible changes in the hydration of protein interior is important to understand experimental differences, which have been observed in the drugs efflux as described in Chapter 3. The radial distribution function (RDF) provides a way to study the hydration of a cavity, since it describes variations of matter density around a reference point. First, the reference particle is surrounded by a sphere with radius r. Then, the number of particles within a distance r and $r + \delta r$ is evaluated

as described in Figure 2.4.1. In Chapter 3, the RDF is computed to study the effect of a mutation in the hydration of a cavity, i.e. the distal binding pocket of AcrB.
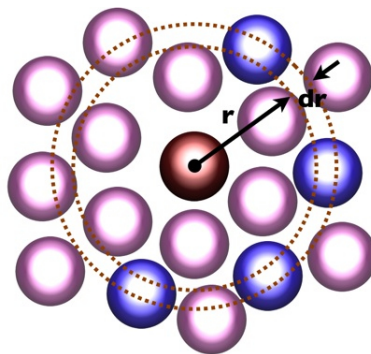

Figure 2.4.1: Scheme of the RDF computation procedure: the reference particle [Red] is first surrounded by a sphere of radius r. Then, the RDF computes the probability of finding a particle in a sphere shell within r and $r + \delta r$ [Blue spheres].

Chapter 3

AcrB

Increased resistance in clinical settings

Adapted from *AcrB drug-binding pocket substitution confers clinically relevant resistance and altered substrate specificity.* Proc Natl Acad Sci (2015): 201419939.

In this Chapter is described the first case of increased drug resistance due to a mutation within an efflux pump, in clinical settings. The study focused on elucidating the mechanism of resistance in a unique set of clinical isolates collected over the course of a complex Salmonella infection [253,254]. The isolates were obtained from a 52-years-old male patient admitted for repair of a leaking abdominal aortic aneurysm graft. The *S.typhimurium* serovar was isolated (strain L3) before the treatment with ciprofloxacin, and was susceptible to ampicillin, sulphonamide, trimethoprim, cefuroxime, chloramphenicol, gentamicin, and ciprofloxacin. During the therapy the patient received i.v. ciprofloxacin, oral ciprofloxacin, i.v. ceftazidime, and i.v. aztreonam. Isolates were taken over the course of the treatments and MDR strains were isolated. The last strain, L18, was isolated after a post antibiotic-free period of two

weeks. A comparison between the pre- and post-therapy isolates showed a less susceptibility to numerous agents, including ciprofloxacin and β-lactams [254], as shown in Table 3.0.1. The genome sequencing revealed a new mechanisms of clinically significant MDR selected during therapy: substitution in a transporter protein, which altered the specificity of the efflux pump for antibacterial drugs. The structural impact of the substitution was investigated by molecular dynamics (MD) simulations and the results will be described in the following paragraphs.

| | Weeks | MIC, μg/mL | | | | | |
Strain	post-therapy	Cip	Nal	Chl	Tet	Atm	AcrB residue 288
L3	0	0.015	2	1	0.16	0.12	G
L10	1	**0.06**	**8**	2	0.12	0.5	G
L11	3	**0.03**	**8**	2	**0.25**	0.25	G
L12	3	**0.5**	**32**	**8**	**0.5**	1	D
L13	3	**0.5**	**16**	**8**	**0.5**	0.5	D
L6	5	**0.5**	**32**	**8**	**0.5**	0.5	D
L16	17	**0.5**	**32**	**8**	**0.5**	1	D
L18	19	**0.5**	**32**	**8**	**0.5**	2	D

Table 3.0.1: Antimicrobial susceptibility of isolates and substitution in AcrB. MIC data presented areare the mode values. MIC values shown in bold are two or more dilutions higher than for the pretherapy isolate, L3. Isolates in bold are those with G288D substitution. Atm, aztreonam; Caz, ceftazidime; Chl, chloramphenicol; Cip, ciprofloxacin; Nal, nalidixic acid; Tet, tetracycline.

3.1 G288D Substitution

Antibiotic resistance, particularly to ciprofloxacin occurred and increased in the isolates along the time course of infection and was first attributed to increased expression of the acrB transcript [255]. Indeed, measurement of acrB expression using an *acrAB* promoter–GFP fusion confirmed increased acrB expression in all strains isolated after the onset of therapy but the expression level did not correlate with the MDR

phenotype. For example, acrB expression in L10 was not significantly different from that of L18, but L18 was more resistant. However, the level of MDR did correlate with which residue was present at position 288 of AcrB; the isolates with the highest level of MDR all carried the G288D substitution in AcrB. Isolates with an aspartic acid at residue 288 were two- to fourfold less susceptible to antibiotics compared with those with glycine at the same residue. The changes in susceptibility were sufficient to confer clinically relevant resistance to the drugs used to treat the patient. To confirm the importance of AcrB to the phenotype, *acrB* was inactivated in the MDR isolate, L18, to give strain L1299 and susceptibility to known substrates of AcrB was determined. Decreased MICs of ciprofloxacin, nalidixic acid, chloramphenicol, and tetracycline, to within one dilution of for the pre-therapy isolate L3, were obtained. In Table 3.1.1 are reported the effect of *acrB* inactivation and G288D mutation on antimicrobial susceptibility.

		MIC, μg/mL					
Strain	Genotype	Cip	Nal	Chl	Tet	Dxr	Min
L3	Pre-therapy isolate	0.015	2	1	0.16	0.12	1
L18	Post-therapy isolate	0.5	32	8	4	>256	4
L1299	L18 *acrB::aph*	0.015	1	1	0.5	4	0.5
L1351	L1299 + p*acrB*-WT	0.06	4	1	1	16	1
L1315	L1299 + pG288D*acrB*	0.25	16	4	2	4	0.25

Table 3.1.1: Antimicrobial susceptibility of isolates and substitution in AcrB. Experiments were carried out on at least three separate occasions and the mode value is presented. MIC values indicated in italic are two or more dilutions lower than for L18. MIC values indicated in bold indicate a difference between the MIC for L18 in which acrB has been inactivated and replaced with mutant (G288D) AcrB or wild-type (WT) AcrB. Chl, chloramphenicol; Cip, ciprofloxacin; Dxr, doxorubicin; Min, minocycline; Nal, nalidixic acid; Tet, tetracycline.

To further investigate the effect of the G288D substitution, the wild-type and mutant *acrB* sequences were cloned into pWKS30 and transformed into the multidrug-

susceptible L1299. Similar levels of wild-type (L1351) and mutant (L1315) $acrB$/AcrB were expressed. Apart from ciprofloxacin, where only a modest decrease in susceptibility was seen (due to the mutation in gyrB) [256], L18 $\Delta acrB$ with wild-type AcrB (L1351) had susceptibility to antibiotics similar to the pre-therapy isolate L3. However, L18 $\Delta acrB$ with mutant G288D AcrB (L1315) was MDR and the MIC values were within one dilution of those for L18 for ciprofloxacin, nalidixic acid, chloramphenicol, and tetracycline. Interestingly, L1315 was more susceptible to minocycline and doxorubicin than L1351, which had wild-type AcrB. This was not seen in L18. This may be due to the level of AcrB being greater in L18 or due to one of the SNPs in another gene confounding the effect of the G288D substitution.

As described in Chapter 1, drug binding within AcrB is complex and is proposed to include several possible sites before extrusion including a vestibule and proximal and distal binding pocket [75, 115]. It has been suggested that low-molecular-mass compounds could bind directly to the phenylalanine-rich distal binding pocket [?, 257, 258], bypassing the proximal binding pocket and vestibule, which is implicated in the recognition of larger compounds [259] such as erythromycin [257, 260]. Structural alignment of the *S.typhimurium* AcrB with the corresponding *E. coli* sequence shows high conservation of the distal pocket residues, suggesting a conserved mechanism of drug binding, and justifying the use of the *E. coli* structure as a guide for rationalization of the effect of the G288D substitution in L18 AcrB. Residue 288 is also highly conserved in AcrB of other *Enterobacteriaceae*, suggesting an important structural role. In the available crystal structures of AcrB in its apo- (PDB ID code 2J8S) [140] and in doxorubicin- (PDB ID code 4DX7) and minocycline-bound forms (PDB ID code 4DX5) [258], G288 is located in the middle of a β-strand, with the neighboring side chains pointing away from the pocket, suggesting that the substituted side-chain will protrude into the pocket cavity, as shown in Figure 3.1.1.
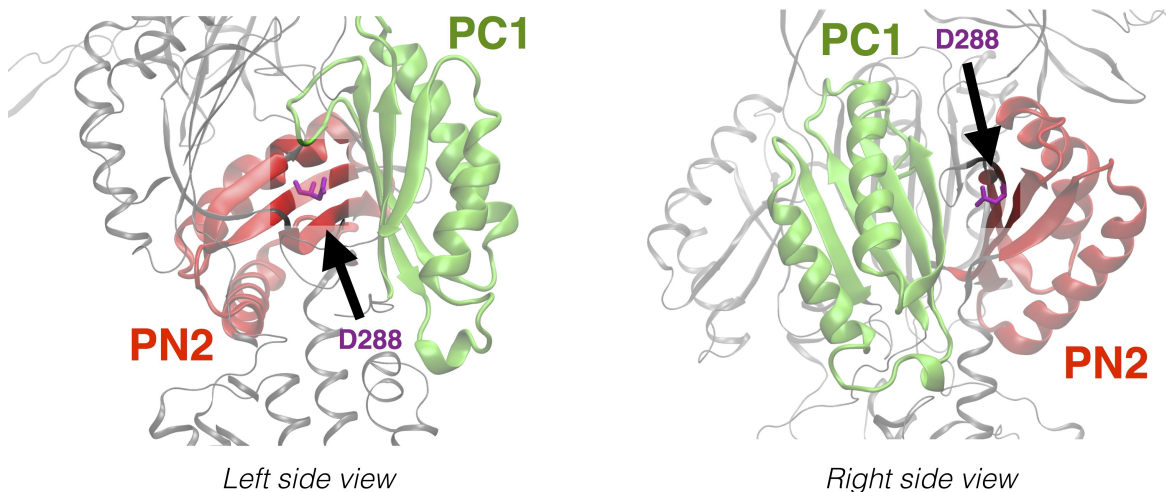
Left side view					*Right side view*

Figure 3.1.1: Mutation G288D in the middle of a β-strand in PN2 domain, suggesting that the substituted side-chain will protrude into the pocket cavity.

The substitution of glycine with this bulky and charged aspartate residue will have a drastic impact on structure and dynamics of this primarily hydrophobic pocket (see Figure (3.1.2) for the example of doxorubicin binding), not only by changing the electrostatic profile and polar environment of the cavity but also by causing steric clashes with several key residues involved in the hydrophobic stacking interactions with the substrate (see Figure (3.1.3)).

Indeed, previous studies have shown that several drug recognition determinants are deeply perturbed by a single substitution in a key region of the transporter [73, 261].
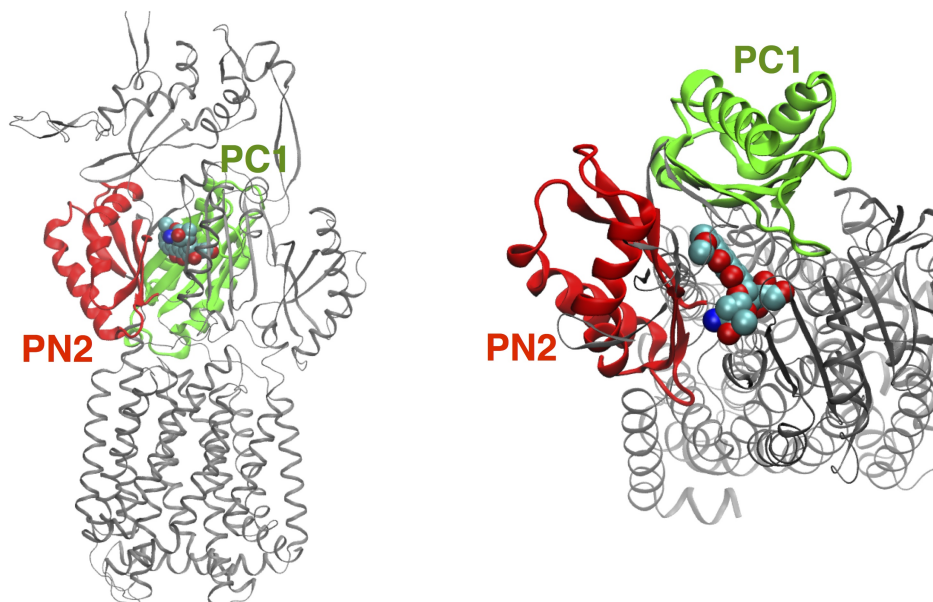
Figure 3.1.2: Binding of doxorubicin to the AcrB monomer. Overall view of the AcrB monomer bound to doxorubicin (shown in space fill) as per PDB ID code 4DX7

3.2 Computational study

In this paragraph are reported the computational methods used to study the effect of G288D mutation in the distal binding pocket and the results of such study.

3.2.1 Homology modeling and MD simulation

To analyze possible changes in the distal binding pocket as a result of the mutation homology models of wild-type and G288D variant of *S.typhimurium* AcrB were built with the Modeler package [221] and using the crystal structure of wild-type *E. coli* AcrB (PDB ID code 2J8S [140]) as a template. For each system 20 different models were generated, which were then sorted by their discrete optimized protein energy score [231]. The best structures according to this criterion were selected for further
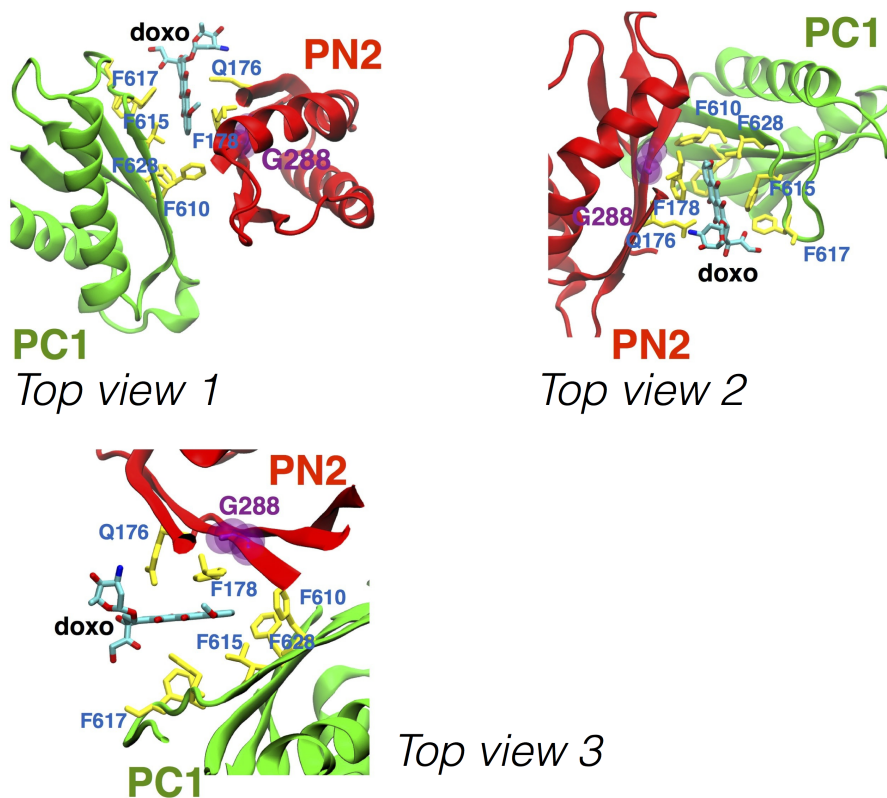
Figure 3.1.3: Different view of doxorubicin binding to AcrB monomer. In purple is reported the mutation site, in yellow the residues involved in drug recognition.

MD simulations. These latter were performed in a 0.1 M KCl water solution, using the NAMD 2.9 simulation package [168] and the ff14SB force field [262]. A reduced model containing only the periplasmic domains (see Figure (3.2.1)) was used as in previous computational studies on *E. coli* AcrB [75, 76, 263]. After structural optimization, the system was equilibrated by linearly increasing the temperature from 0 to 310 K in 1 ns of NPT MD simulation and equilibrated further for additional 5 ns. Then, multi-copy simulations (three for each system: one of 400 ns and two of 200 ns for the wild type and three of 400 ns for the G288D system) in the NPT ensemble were performed both on the wild-type and on the G288D systems.
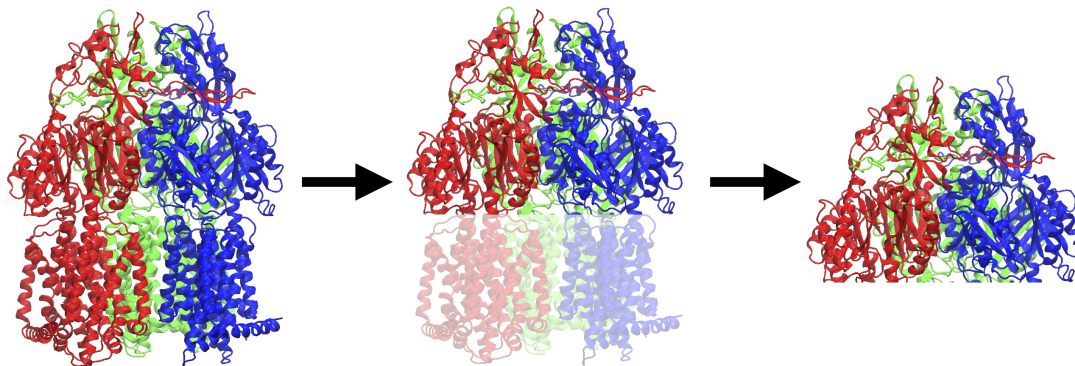
Figure 3.2.1: Reduced AcrB model. The trans-membrane domains were cut and only the periplasmic part of the structure was used in the MD simulations.

3.2.2 Results

Comparison of representative structures (i.e., the conformations with the lowest RMSD from the averages) of the wild-type and mutant proteins revealed that the introduction of the aspartate induces local structural changes in the distal binding pocket (RMSD 3.1 Å). In particular, residues F178 and Q176 assume alternative conformation to avoid steric clashes with the aspartate and F136 shows a large displacement (See Figure (3.2.2)a). As a result, the radius of gyration calculated for the residues lining the distal pocket increased by ∼1Å in the G288D variant (from ∼10.5 Å in wild type to ∼11.5 Å, Figure (3.2.2)b). The radius of gyration is indicative of the level of compaction in the pocket and the measured variation indicates an increased size. However, these changes have only limited effect on the secondary structure of the distal binding pocket as shown in Figure (3.2.3).

In addition to these local structural changes in the distal binding pocket, possible changes in the access regulation were investigated. In particular, since the switch-loop (see Figure 3.2.4) has been already proved to be crucial in the drug access regulation [258, 264], any change in its flexibility might suggest different access
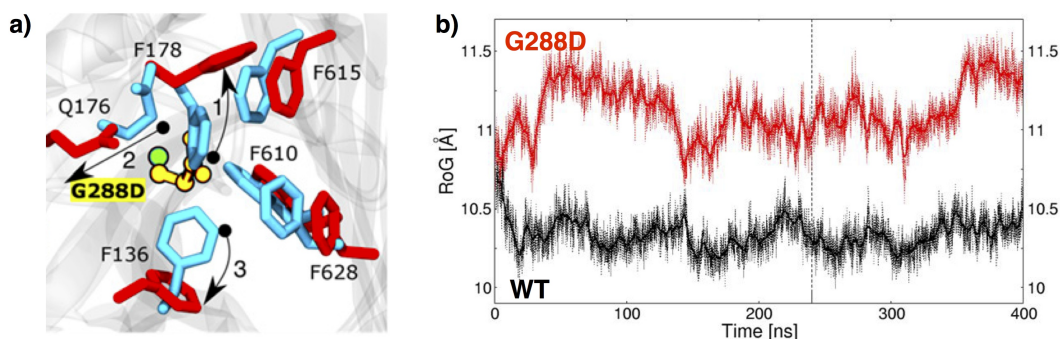
Figure 3.2.2: Effects of the G288D mutation in the distal binding pocket. a) Close view of the binding site where important residues are shown (blue, wild type; red, mutant). Owing to the presence of the side chain of the aspartate residue in the G288D mutant, the side chain of F178 tilts away in the G288D mutant compared with its orientation in the wild type protein (denoted by the arrow and labeled as 1). In turn, the side chain orientation of the residue Q176 is also altered in the mutated protein (denoted by 2) with respect to the wild-type form. Finally, the side chain of F136 also changes orientation (denoted by 3) **b**) Evolution of the distal pocket's radius of gyration along the 400 ns long MD simulations of the wild type (dotted black line) and G288D (dotted red line) variant of AcrB.

mechanisms. Thus, a comparison between the flexibility, in terms of RMSF, of the switch-loop in the wild type and the mutant was performed. In Figure 3.2.5 is reported the difference between the RMSF of the G288D and the wild type in two functional states, i.e. *tight* and *access*. The differences induced by the mutation in both the *tight* and *access* states suggest the possibility of new binding mechanisms because of the higher flexibility in the *tight* state and different access regulation due to the decreased flexibility in the *access* state.
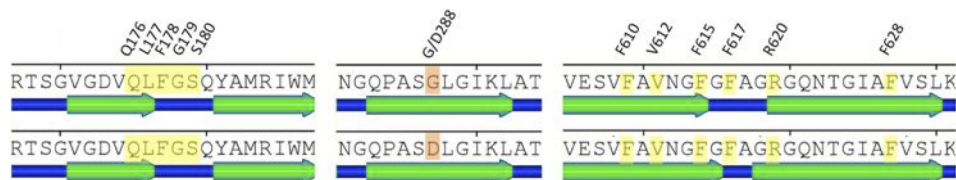
Figure 3.2.3: Schematic diagrams of the secondary structure of the distal drug binding pocket (calculated on the closest-to-average structure extracted from MD simulations) in wild-type (upper diagram) and G288D mutant (lower diagram) of AcrB. Secondary structure elements are numbered according to Murakami et al.
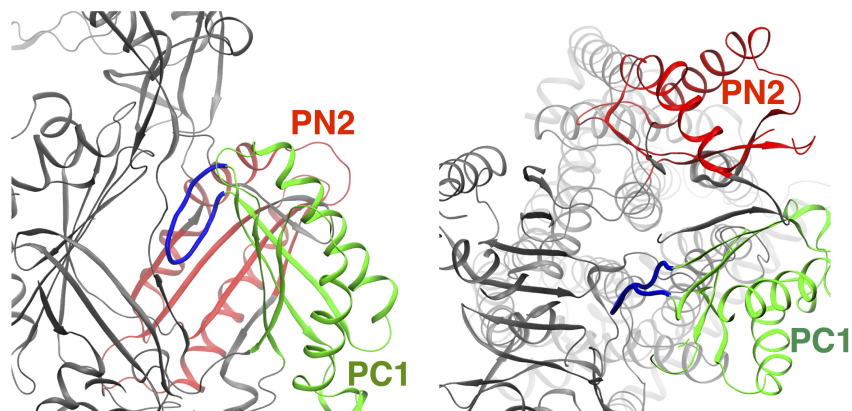
Figure 3.2.4: Switch-loop (Blue) that regulate the access of small drugs in the distal binding pocket (PN2, in Red, and PC1, in Green).
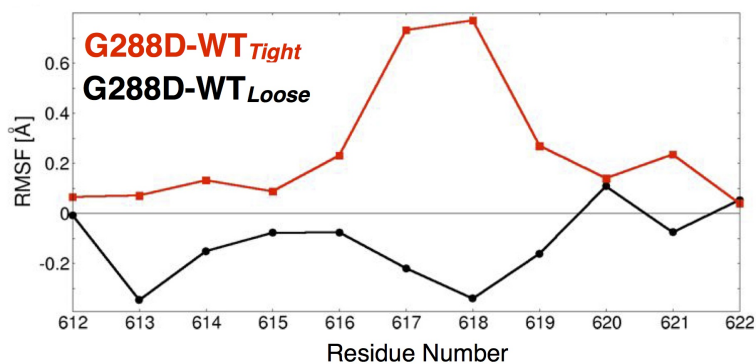

Figure 3.2.5: The difference between the RMSFs of the switch loop in G288D mutant and wild type AcrB for monomers in the access (A, black curve) and binding (B, red curve) configurations. A and B 3 refer to the configurations that the monomers are supposed to assume during the proposed functional rotation.

Furthermore, the introduction of the aspartate increases the polarity of the otherwise highly hydrophobic distal binding pocket (see Figure (3.1.3)) and hence may affect its hydration. To test this, radial distribution function (RDF) profiles, which show the probability of finding water molecules at a certain distance from the G288 / D288 residues, were calculated. Water residence times and average numbers of water molecules were calculated as previously described [265, 266] with the following alterations. First, it was defined fast-, medium-, and long-time waters as those having survival probabilities τ of less than 100 ps, between 100 ps and 1 ns, and higher than 1 ns, respectively. Second, because the trajectory was saved every 10 ps, there was no need to use a stretched exponential to fit survival.

As anticipated in the mutant protein the D288 residue was found to have a higher probability of being surrounded by water molecules during the course of the simulation (Figure (3.2.6), solid lines). The sharp peak at around 2.5 Å from D288 is due to the presence of the first hydration shell. To the contrary, water molecules are on average absent in the close vicinity of G288, consistent with the more hydrophobic environment in the wild-type protein (Fig. 2, dashed lines). In addition, the mutant exhibits a notable change in residence time for water molecules within the distal pocket as described in Table 3.2.1.
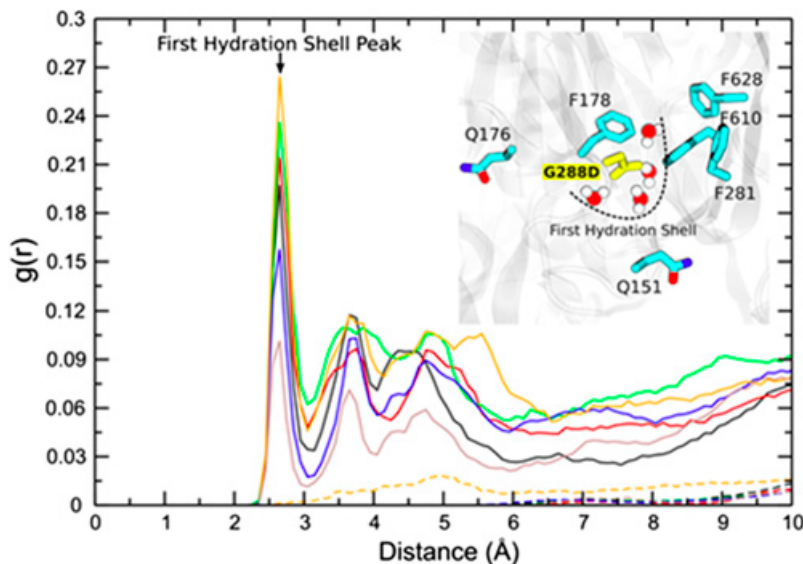
Figure 3.2.6: RDF profiles indicating the distribution of water molecules near the D288 residue in the mutant (solid lines) protein and near the G288 residue in the wild-type protein (dashed lines). Differently colored lines (solid or dashed) represent the RDF profiles calculated from two different simulations and for each of the monomers of the AcrB protein. RDF profiles are calculated based on the distance between the water oxygen to center of mass of G288/D288 residue's side chain. (Inset) A representative snapshot of the presence of water molecules in the binding pocket of the G288D mutant protein. Water molecules at the corresponding positions are absent in the wild-type protein.

Model	n_s	τ_s	n_m	τ_m	n_1	τ_1
WT	26.7	72.8	7.3	392	1.1	4278
G288D	30.9	92.4	14.9	469	3.9	2546

Table 3.2.1: Number and average residence times (in ps) of short (n_s, τ_s), medium (n_m, τ_m) and long (n_1, τ_1) residence time waters within the distal pocket of monomer B. The values are computed by fitting the reduced survival probability $\Delta Nw(t)$ with three exponentials.

Then the effect of the G288D substitution on doxorubicin, minocycline, and ciprofloxacin efflux was analyzed. Doxorubicin and minocycline, although not clinically relevant, have been co-crystallized with E. coli AcrB, making phenotypic and

structural data easier to correlate and interpret. A comparison between these co-crystal structures and the results from the MD runs of the G288D variant was then performed. Consistent with MIC and drug accumulation data below, and based on the assumption that the binding of tetra-ring structures follows that in the co-crystal structures, the G288D substitution would result in pronounced steric clashes with both doxorubicin and minocycline, due also to displacement of amino acid F178 discussed above. Although an experimental structure of the ciprofloxacin–AcrB complex is lacking, a recent computational study [75] suggested that it binds to the bottom of the distal pocket, in a different position than doxorubicin (See Figure 3.2.7).
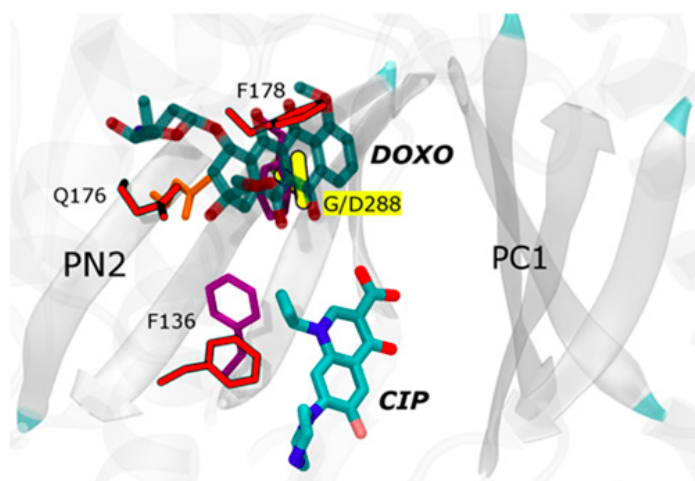
Figure 3.2.7: Predicted effects of the mutation on the binding of the doxorubicin (DOXO) and ciprofloxacin (CIP) in the distal binding pocket. DOXO and CIP are rendered with transparent and solid thick sticks, respectively, colored by atom type. The Cα atom of G288 is rendered in green and the side chain of D288 with yellow sticks. Side chains of other most relevant residues are shown with thinner sticks, colored by residue type and red in wild type and G288D, respectively. The orientation of CIP is derived from the MD simulations (27), and that of DOXO is based on the experimental X-ray structure as from PDB ID code 4DX7.

In contrast to doxorubicin and minocycline, the G288D substitution should not

impair the binding of ciprofloxacin; instead, the additional hydroxyl groups present in this compound may form hydrogen bonds with the introduced aspartate and resident waters, improving its recognition. In agreement with these hypotheses, G288D in L1315 (L18 $\Delta acrB$ pG288D$acrB$) not only led to increased susceptibility to doxorubicin and minocycline, likely due to decreased efflux, but also conferred increased efflux of, and decreased susceptibility to, ciprofloxacin and other fluoroquinolones. Although the differential effect of the mutation on the different classes of antibacterial drugs is not immediately obvious, it is notable that the efflux of more hydrophobic substrates such as doxorubicin and minocycline is correlated with the observed changes in the hydrophobicity of the distal binding pocket. Accordingly, minocycline and tetracycline seem differentially affected, consistent with the former being 10-fold more lipophilic than the latter. The computational results suggest a decrease in ciprofloxacin susceptibility for AcrB G288D strains owing to an increase efficiency of the its efflux. Subsequently, these prediction has been confirmed by experimental tests in which the mutant AcrB protein was expressed in two well-recognized standard antibiotic-susceptible reference strains, S. Typhimurium SL1344 and E. coli MG1655. In this way, unlike in L18, there were no other mutations that could affect antibacterial drug susceptibility and drug efflux. This allowed study of the G288D substitution in the absence of other mutations that could confound the phenotype. To explore whether efflux activity was altered due to the AcrB G288D substitution alone, the accumulation of ciprofloxacin, doxorubicin, and minocycline was measured. In both Salmonella and E. coli the strain expressing mutant AcrB (L1352 and I972, respectively) accumulated less ciprofloxacin than the strain expressing wild-type AcrB (L1365 and I971, respectively), suggesting more efficient efflux of ciprofloxacin by the mutated protein (see Figure 3.2.8).
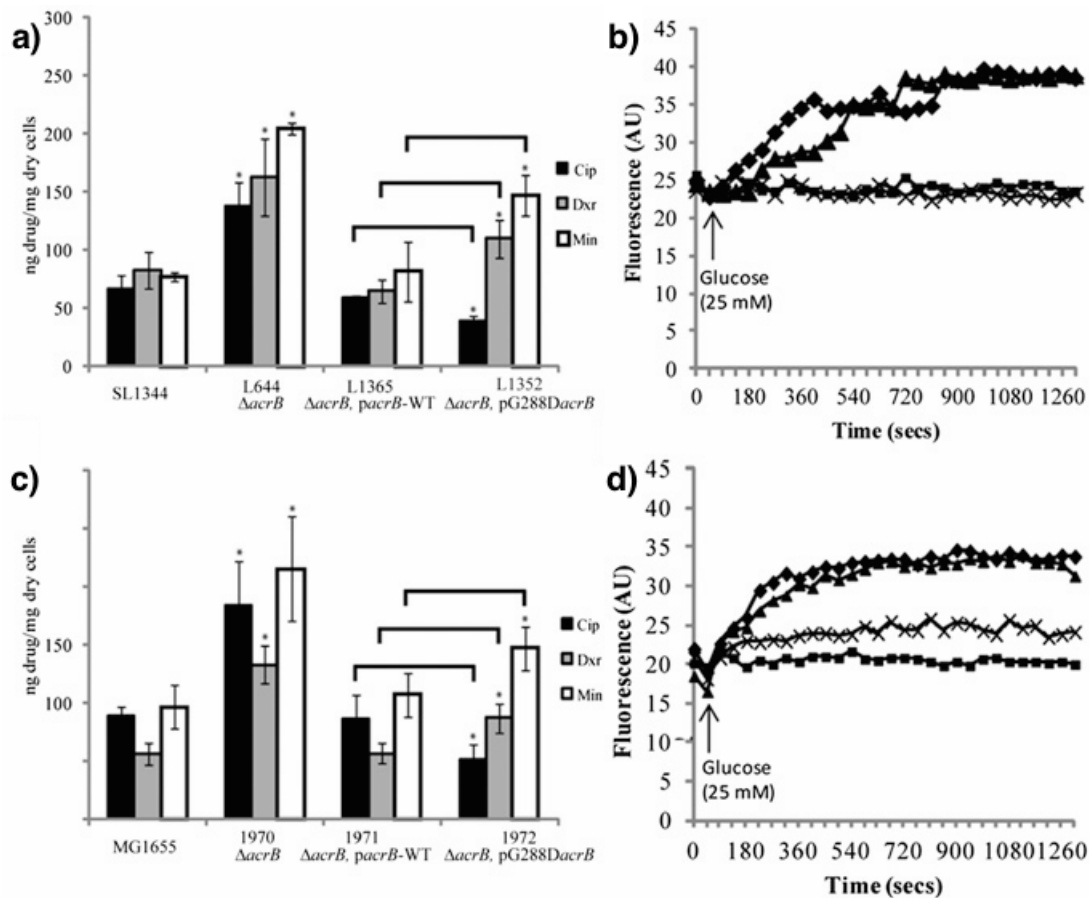
Figure 3.2.8: The effect of the G288D mutation on the accumulation of ciprofloxacin (Cip), doxorubicin (Dxr), and minocycline (Min) and on the efflux of doxorubicin in Salmonella (**a** and **b**) and E. coli (**c** and **d**). **a)** and **c)** show accumulation of Cip, Dxr, and Min in Salmonella and E. coli, respectively. Black bars show Cip accumulation, gray bars show Dxr accumulation, and white bars show Min accumulation. Data presented are the mean of the steady-state values for three biological replicates $\pm$ SD. Asterisk denotes values that are significantly different from the corresponding wild-type parental strain (SL1344 or MG1655). Lines denote values, for strains carrying the G288D mutation, which are significantly different to the strain carrying the wild-type AcrB. **b)** and **c)** and **d)** show the efflux of Dxr by Salmonella SL1344 and MG1655, respectively. In each case diamonds denote the wild type, squares denote ΔacrB, triangles denote the acrB mutant complemented with wild-type acrB, and crosses denote the acrB mutant complemented with G288D acrB.

Conversely, strains of both species expressing mutant AcrB accumulated more

doxorubicin and minocycline, indicating reduced efflux of these compounds (Figure 3.2.8a and c). In addition, *Salmonella* and *E. coli* with G288D AcrB extruded doxorubicin less than strains with wild-type AcrB (Figure 3.2.8b and d). These data support the hypothesis that the G288D mutation in AcrB improves efflux of ciprofloxacin and so confers resistance to ciprofloxacin (and other fluoroquinolones) while perturbing doxorubicin and minocycline binding and reducing efflux, so increasing susceptibility to these agents. Furthermore, experiments examined whether the G288D substitution affected any other *Salmonella* attributes that could have contributed to the disseminated salmonellosis and the failure to effectively treat the patient. As result, the G288D mutation did not alter fitness when grown in competition with a strain with wild-type AcrB and did not affect invasion in the tissue culture model of infection.

3.3   Discussion

In this study an interdisciplinary collaboration investigated the increased resistance to ciprofloxacin in *S.typhimurium*, acquired during clinical treatments. During the course of the therapy isolates were taken and the increased resistance appeared together with a substitution of one glycine reside with a negatively charged aspartic acid, in the distal binding pocket of AcrB. Thus, a single nucleotide change caused MDR by alteration of the substrate specificity of an MDR efflux pump. X-ray crystallography of RND transporter proteins with substrates has proven very challenging, giving information only for three antibiotics, erythromycin, minocycline, and rifampicin [58, 257, 258] and one inhibitor [267]. In the absence of X-ray crystallography data for AcrB co-crystallized with a fluoroquinolone antibiotic, computational structural analysis of AcrB with the G288D substitution was used to predict any

effect upon ciprofloxacin binding.

The simulations were performed only on the apo form of the transporter. Thus, in rationalizing experimental data, it was implicitly assumed that also in AcrB G288D variant the key interactions with substrates occur in the same distal binding pocket identified in the wild-type protein. Clearly, this must be proven; nonetheless, the structural modeling that we carried out demonstrated that G288D would substantially change the environment of the distal binding pocket because a bulky and charged residue would perturb the primarily hydrophobic region and in particular the conserved phenylalanine cluster of the pocket. Because this site has been shown to be extremely adaptable to many substrates [75], it cannot be ruled out that despite the large effect due to G288D substitution the pocket would still be able to accommodate most compounds but with altered affinities. In agreement with such an interpretation, it was observed that the G288D substitution caused an increase in intracellular accumulation of doxorubicin and minocycline, which have been reported to bind to the phenylalanine cluster in the distal binding pocket [58], and concomitant increased susceptibility. In contrast, it was detected decreased intracellular accumulation of, and resistance to, ciprofloxacin and other more hydrophilic drugs including tetracycline. In addition, the study showed that this phenotype could be recreated in well-characterized standard strains of *Salmonella* and *E. coli*, confirming that the G288D mutation, and not any of the other SNPs detected in the clinical isolate, L18, was responsible for MDR. Thus, the modeling provided a plausible mechanistic explanation for the effect of the mutation G288D. A recent study using MD simulations has highlighted an unexpectedly complex pathway of drug expulsion, with several possible routes of export through AcrB [259]. Full investigation of the interaction between drugs and the mutated protein clearly merits a systematic, stand-alone full MD simulation and/or a crystallographic solution.

MDR in clinically relevant infections has previously been associated with over-expression of RND efflux pumps owing to mutations in regulatory genes [11]. The same association was previously made for L18 [255]; indeed, increased expression of acrB has been confirmed in this study. However, based upon our recent data showing that the bacterial cell responds to lack of efflux (irrespective of mechanism) [268], we now postulate that increased transcription of acrB relative to L3 is in response to altered efflux owing to the G288D substitution.

The literature contains many examples of engineered mutations in RND efflux pumps but very few naturally occurring mutations have been described. Before this study there was no indication that substitutions in AcrB would confer antibiotic resistance; the accepted dogma was that AcrB-mediated resistance was caused by overproduction of the protein. Furthermore, in this study the G288D substitution arose during the course of clinically appropriate antimicrobial treatment and the impact was profound; the infecting organism became resistant to agents with which the patient was treated, therapy failed, and the patient died due to bacteraemia. The computational study showed that the mutation subtly affects several possible determinants responsible for the recognition and selectivity of the transporter, including orientation of side chains, polarity, electrostatic potentials, and steric properties.

Chapter 4

Membrane Fusion Protein: MexA

The role of the MFP component in the MexAB-OprM efflux pump, expressed in *Psaeudomonas aeruginosa,* has been largely studied with direct *in vitro* experiments and through the study of its homologues in other bacteria, like *E. coli* [8, 32, 55, 60, 90, 92, 114, 144, 146, 147, 269–275]. As result, different models about the MexA contribution in the assembly stoichiometry, and its role in the efflux process has been proposed [55, 86–88, 114, 273]. As consequence of the different number of MexA proteins that are supposed to participate in the assembly, different interacting mechanisms between the other two components of the pump have been suggested. One of the first models proposed predicts the participation of three MexA monomers [55, 103, 273] with a direct connection between the RND transporter MexB and the outer channel OprM. However, recent structural studies support the 3:6:3 stoichiometry model as described in Figure 1.2.6. As example, the MFP MacA protein was shown to form an hexamer, which is required for the interaction with the ABC-type transporter MacB [276, 277]. Furthermore, structural studies of a CusAB complex have revealed an hexamer of the MFP CusB bounded to the inner membrane protein

CusA [278], providing an indirect support to the 3:6:3 model for the MexAB-OprM efflux pump. More direct support is provided by studies on chimeric proteins AcrA-MacA and MexA-MacA [86, 87]. The results of these studies suggest that the two MFP components could form a funnel-like hexameric structure, in agreement with other experiments [88]. One of the main differences introduced in the 3:6:3 stoichiometry model is the not-direct connection between the RND transporter MexB and the outer channel OprM, indicating a new active role of the MFP in controlling the conformational changes that lead to the channel opening [88, 114]. In this context of uncertainty about the number of membrane fusion proteins involved in the MexAB-OprM efflux pump, the studies described in this chapter aim to provide a better understanding about the effectiveness of the 3:6:3 stoichiometry model and to enrich the relatively few MFP computational studies [145, 148]. In particular, the study focused on the possibility for MexA proteins to form stable dimers, which seems to be the pre-condition required to form the MexA hexameric structure proposed [86, 99]. In order to understand if such dimerization process might be possible, wild type monomers and dimers were studied multiple 2μs MD simulations. In addition, an homology model for the G72S variant was built. The mutation is located in the *lipoyl* domain (see Figure 4.0.1 for more details) and it is known to inhibit the MexAB-OprM efflux pump, presumably by affecting the MFPs dimerization [92]. To study the dimerization process a protein-protein docking program was used [240, 241] and the results assessed with multiple 500ns MD simulations and with a comparison with structures from the MD simulation for the dimeric homology model.
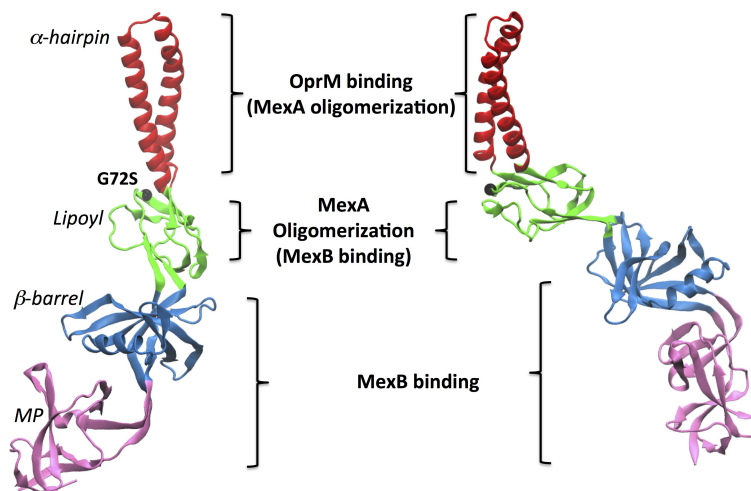
Figure 4.0.1: MexA monomer structure divided into four domains: *a–hairpin* [Red], involved in MexA-OprM binding; *lipoyl* [Green], supposed to be involved in MexA dimerization; *β–barrel* [Blue] and the *Membrane Proximal* [Mauve] domain, both involved in MexA-MexB binding. The black sphere in the *lipoyl* domain identifies the mutation site.

4.1 System and Methods

In this section will be described the details about the study on MexA, from the homology modeling of the monomeric and dimeric units to the MD simulation and docking details.

4.1.1 MexA monomeric unit

MexA monomers in the re-refined crystal structure are in two main state, the so-called rotate and unrotated state [55], with the latter state conventionally considered as the representative of the functional state [31,55]. The rotated state, however, has not been completely ignored as it was used in *in silico* studies to model different states of AcrA [145] and to predict the hexameric funnel-like structure of MexA and
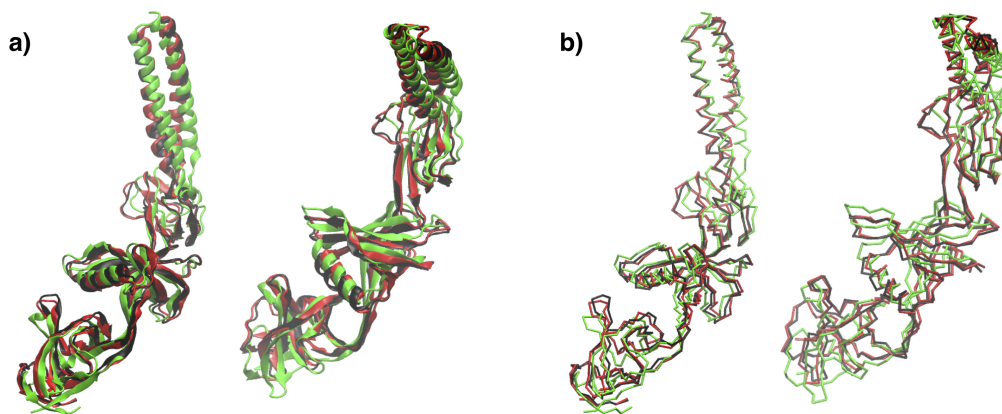
Figure 4.1.1: Comparison between the reference (chain L, PDB code 2V4D) structure [Green], G72S-1 [Red] and G72S-2 [Black] in the new-cartoon **(a)** and the trace **(b)** representation.

AcrA [86].

To model the mutated monomeric unit, the MODELLER package [221] was used with the unrotated state of the crystal structure (PDB code 2V4D [55]), as template. The mutation was introduced in the FASTA sequence by replacing the glycine residue in the *lipoyl* domain, G72, with a serine residue (see Figure 4.0.1 for more details). As reference, chain L was selected following previous works [86, 87] although other chains in the same unrotated state were available. During the modeling process two independent runs were performed. For each run, 50 structures have been generated and the DOPE score criteria [231] was used to select the best models, named G72S-1 and G72S-2 (see Figure 4.1.1 for more details). The Cα-RMSD between models G72S-1/2 and the chains in the unrotated state was comparable with the Cα-RMSD between the chain L and the other chains. Subsequently, a Then, the modeling of the mutant was performed with the same criteria of the wild type case.

4.1.2   MexA dimeric unit

The MexA hexameric funnel-like structure proposed in previous works was either based on three dimeric subunits [86] or six independent monomer [88]. The first model, is also supported by mutagenesis study that suggest MexA-MexA dimerization [92] and in vitro experiments [274]. To assess the possibility of the formation of a stable MexA dimer, an homology dimeric model was built following the same procedures described for the monomeric units. The reference structure for the dimer modeling was suggested by previous works [86] to be a combination of MexA unrotated and rotated state, namely chains L and M.

Thus, the study of the possible MexA dimerization was divided into two main fronts:

1. Building of dimeric homology model based on chain L and chain M (see Figure 4.1.2). This front was chosen to validate the hexameric funnel-like structure and to understand wether the rotated state could be representative of the functional state of the protein.

2. Docking runs for the MexA monomers based on chain L. Starting from the definition of the unrotated state as potentially functional, monomers extracted from the MD simulations were selected to study the dimerization process.

4.1.3   MD simulation

The size of the monomeric units was ~5K atoms for both the wild type and the mutant, while the dimeric unit contained ~10K atoms. The systems were prepared for the simulations by solvation with a truncated octahedron water cells and by adding the counter-ions to neutralize the net charge. The size of the final structures was
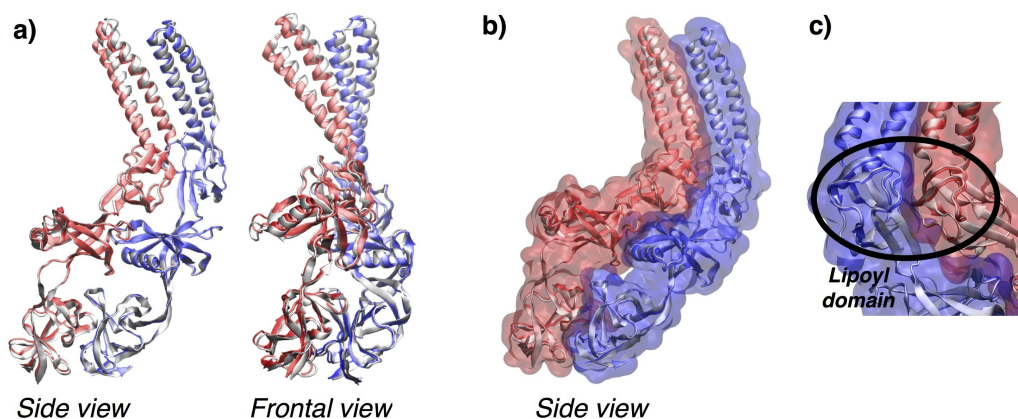
Figure 4.1.2: **a)** Comparison between the crystal structure [White] and the homology model in chain L [Blue] and M [Red]. **b)** Homology dimeric model in surface representation. **b)** Focus on the *lipoyl* domains that are supposed to be involved in the MexA-MexA dimerization.

$\sim$200K atoms. After the solvation, each system was was first structurally relaxed for 1,000 steps (500 with the steepest descent algorithm, followed by 500 steps of relaxation with the conjugate gradient method) in presence of soft restraints (force constant of 10.0 $kcal/mol \cdot \mathring{A}^2$) on all non-hydrogen atoms of the protein. The restraints were then removed and the system was further optimized for 10,000 steps. Then, heating to 310K was performed in 1 ns under NVT conditions, using the Langevin thermostat followed by 2 ns-long equilibration under NPT ensemble (using the Berendsen barostat). Then, the MD simulations were run up to 2 µs with AMBER12 [279].

System	WT1	WT2	G72S-1	G72S-2	Dimer
Number of atoms	200K	200K	200K	200K	240K
Simulated Time (μs)	2	2	2	2	2

Table 4.1.1: Resume of the MD simulations run to study MexA dimerization: two wild type system to investigate monomeric stability and dimerization process; G72S mutant that inhibits the pump functionality by affecting MexA dimerization; homology dimer to investigate the stability of the possible dimeric units.

4.1.4 MexA-MexA dimerization

As described in Chapter 2 the docking runs were performed with a protein-protein docking program ATTRACT to study the possible dimerization of the wild type monomers and the mutants as well. The flexibility of the proteins was taken into account by including the ensembles of different conformations of each system. The ensembles contained the representatives of the most populated cluster for the monomers during the simulations. During the docking a soft harmonic restraint, k $= 2 \, kcal/mol \cdot \mathring{A}^2$ for the maximum distance between specific regions in the MexA monomers, was introduced to prevent the outcome of unphysical results, i.e. proteins oriented in opposite direction as described in Figure 4.1.3b. The atoms selected to be restrained were the closest to the center of mass of the α-hairpin and the MP domain as shown in Figure 4.1.3a.

The selection of the best structure was based on the Cα-RMSD between the dimeric homology model and all the possible configurations from the docking runs. The best structures were then prepared for new MD simulations in order to study the stability of such possible dimers. The systems preparation for the MD simulation followed the procedures described in the previous paragraph.
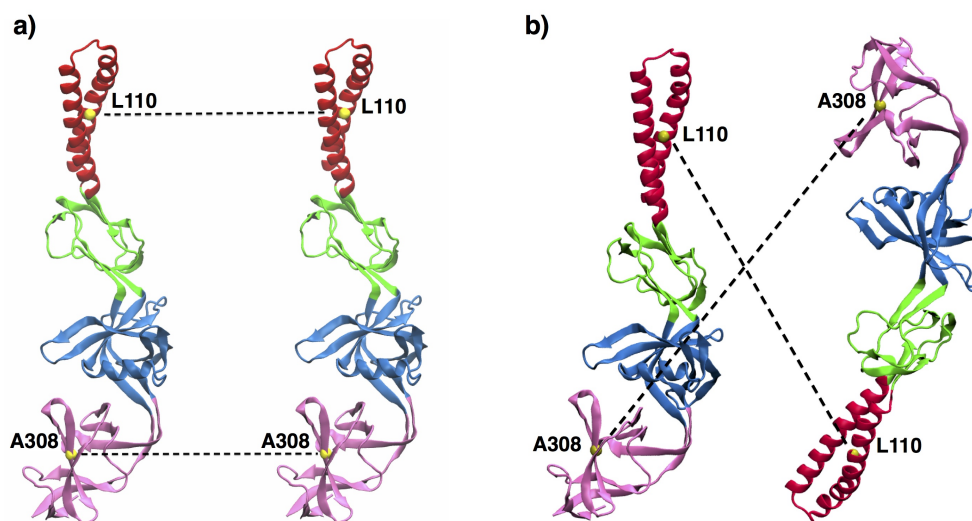
Figure 4.1.3: Scheme of the maximum distance harmonics restrained for the MexA-MexA docking **(a)**. The atoms selected, in yellow, were in the center of mass of the *a–hairpin* and the *MP doma*in. The restraints were applied to prevent unphysical structures, such as two monomers in opposite direction **(b).** The maximum distance has been chosen as 50 Å to guarantee enough flexibility for the possible structures.

4.2 Results

In this section the results of the MD simulations for the MexA monomers, wild type and G72S mutant, and for the MexA homology dimeric model will be presented. The simulations were run up to 2 μs for all the systems with 250 ns extra MD simulations for the two mutant monomers. The extra simulations were needed because, although both the mutants showed large structural changes, they occurred in different time. The analysis of 250 ns after the 2 μs MD simulations, revealed a convergence of the mutants to a common deformed structure, with respect to the wild type case.

4.2.1 MexA monomeric unit

The analysis of the global fluctuations for the MexA MD simulations revealed large RMSD values for both the wild type and the mutant with respect to their initial configurations. In the wild type case, however, the RMSD, despite the high values from the initial configuration, can be considered relatively stable especially if compared with the G72S mutant behavior. In particular, the WT1 system revealed an oscillating behavior while WT2 resulted to be more stable. The differences between the two models will be discussed in more details later in this paragraph. The mutant, on the other hand, in both the simulations showed a large drift suggesting larger global changes in the protein structure, shown in Figure 4.2.1a. However, these changes occur in different time and because of that the two mutant MD simulations were extended for extra 250 ns. The resulting RMSD analysis confirmed a stabilization of the RMSD (Figure 4.2.1a), providing a first index for a common deformed conformation. The analysis of the Cα-RMSD matrix confirmed the different behaviors between the wild type monomers and the mutant. The two wild type structures appear, once again, to be more stable during the simulation although WT1 showed a global oscillating evolution. On the other hand the mutant structures appear to change dramatically during the simulation respect to their initial structure, as shown in Figure 4.2.1b.A visual inspection of the trajectories showed a global rearrangements in the β–barrel and MP domain of the two mutant monomers, followed by a bending of the a-hairpin in the β–barrel domain direction, as shown in Figure 4.2.4.
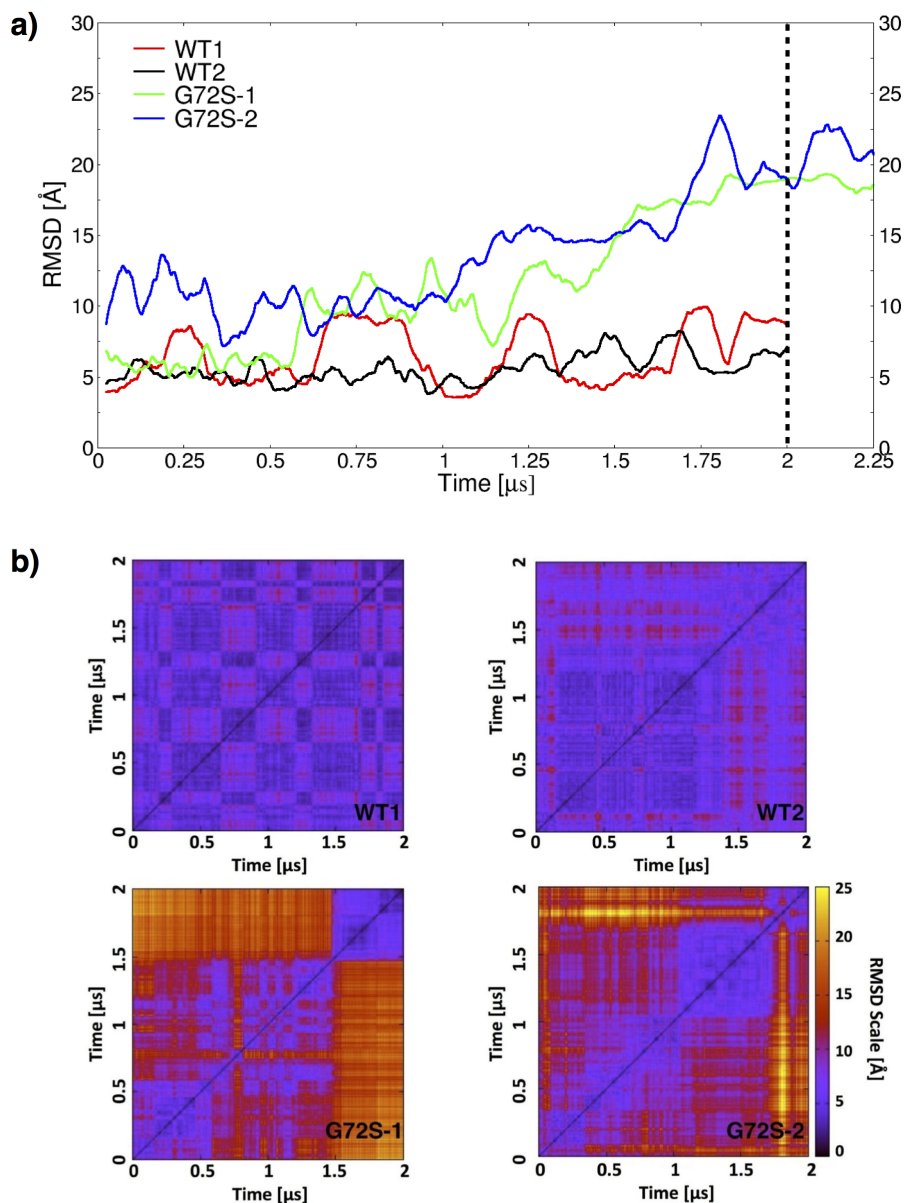
Figure 4.2.1: **a)** Cα-RMSD for WT1 [Red], WT2 [Black], G72S-1 [Green] and G72S-2 [Blue]. The extra simulated time refers to the MD simulation for G72S-1 and G72S-2. **b)** Cα-RMSD matrix for the MexA monomer simulations: WT1 (top left), WT2 (top right), G72S-1 (bottom left) and G72S-2 (bottom right). Each element (i, j) of the matrix consists of the RMSD (Å) value between the structure of the protein at time i and j: the lower the value the higher the similarity of the structures (special case for diagonal elements (i, i) with zero value of the RMSD).

While the Cα-RMSD analysis revealed large global changes between the wild type and the mutant, Cα-RMSF is a way to evaluate local differences as it describes atomic deviations with respect to a reference configuration. The analysis showed that, while the two wild type monomers have almost the same behavior the mutation appears to reduce the internal flexibility of the domains as shown in Figure 4.2.2b. The largest differences are located in the *a–hairpin* domain, which is known to have large flexibility [145, 148], presumably to adapt to OprM surface (see figure 4.2.2a for the interaction scheme). The loss of flexibility in this region suggests a reduced ability for MexA to interact efficiently with OprM. Although the largest differences appear to be located in the *a–hairpin* domain, also the *β–barrel* (MexA-MexB interaction) and *lipoyl* (presumably MexA-MexA interaction) domain are affected by the G72S mutation.
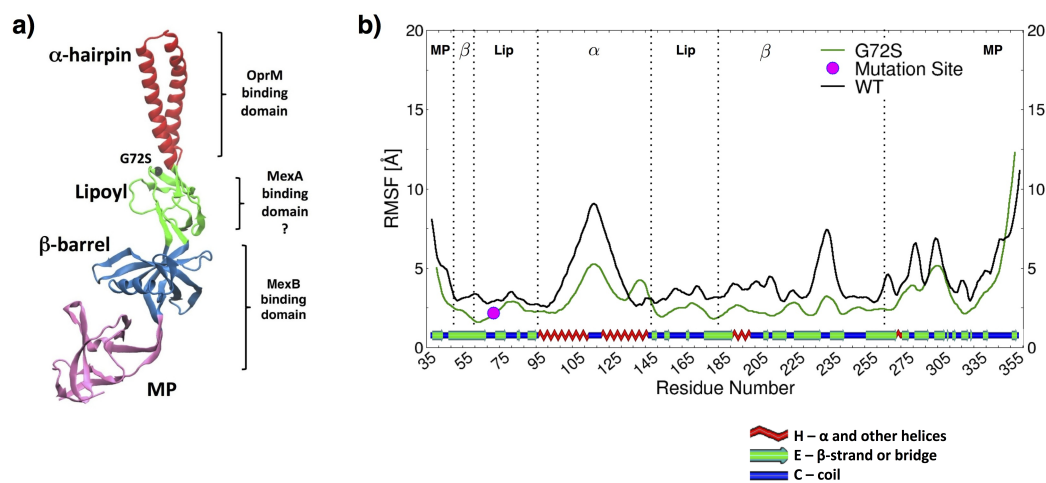

Figure 4.2.2: **a)** MexA interacting and domain schemes. Cα-RMSF for MexA monomers. **b)** average Cα-atoms fluctuation for WT [Black] and G72S [Green]. The mutation site is identified by the Magenta dot. The Secondary Structure was obtained with POLYVIEW-2D [280]

While the Cα-RMSD and Cα-RMSF analysis pointed out global and local differ-

ences between the wild type and the mutant, a visual inspection of the trajectories revealed large inter-domain rearrangements with the bending of the *a–hairpin* and *MP* domains in the mutant. To confirm this behavior, the analysis of the Radius of Gyration (RoG) was performed, showing a trend for the mutant structures to evolve into a more compact conformation while the wild type remains globally stable, with few fluctuations as described in Figure 4.2.3. Despite the same behavior for the two mutant structures to reach a more compact conformation, only with the analysis of the RoG of the extra MD simulations they showed the same final compactness, as shown in the extra simulated time in Figure 4.2.3.
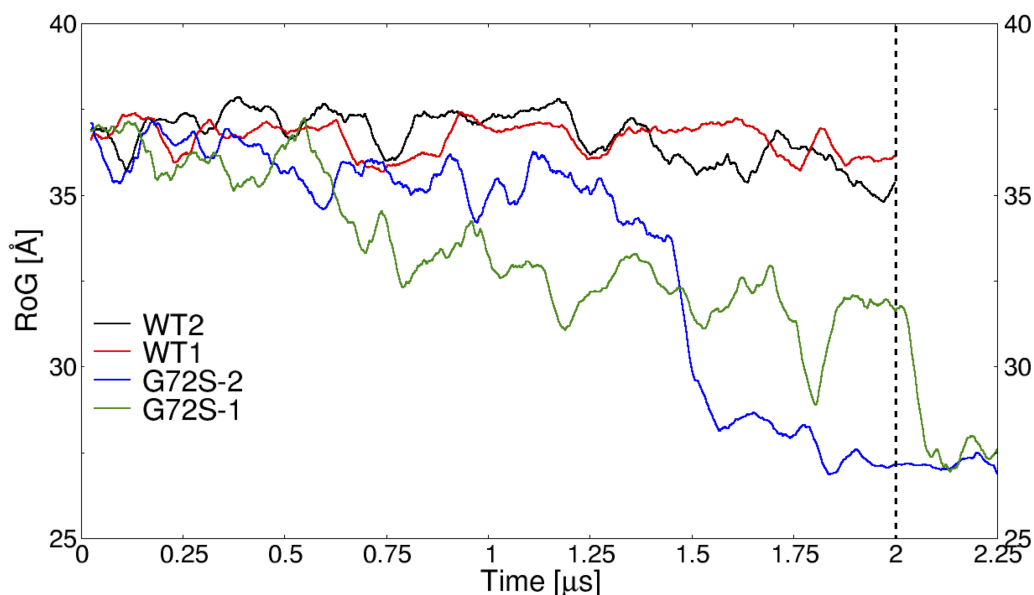
Figure 4.2.3: RoG for WT1 [Red], WT2 [Black], G72S-1 [Green] and G72S-2 [Blue].

The RoG, however, describes only of the global changes in the structure during the MD simulations of the mutant. To identify the domains responsible for such changes, the evolution of the distance between the center of mass of each domain was studied. In the wild type case the domains oscillate around an equilibrium

distance while the mutant showed largely different behavior, as shown in Figure 4.2.5. The analysis of the extra simulations for the MexA mutant, in Figure 4.2.5, revealed a convergence in the domains distance, suggesting a common final structure. An example of conformational changes induced by the mutation are shown in Figure 4.2.4.

Figure 4.2.4: **a)** Initial conformation for the G72S-1 mutant **b)** Final conformation for the G72S-1 mutant with the α-*hairpin* domain [Red] that approaches the β-*barrel* domain [Blue] as well as the *MP domain* [Mauve] **c)** MexA domain scheme and mutation site [Black].
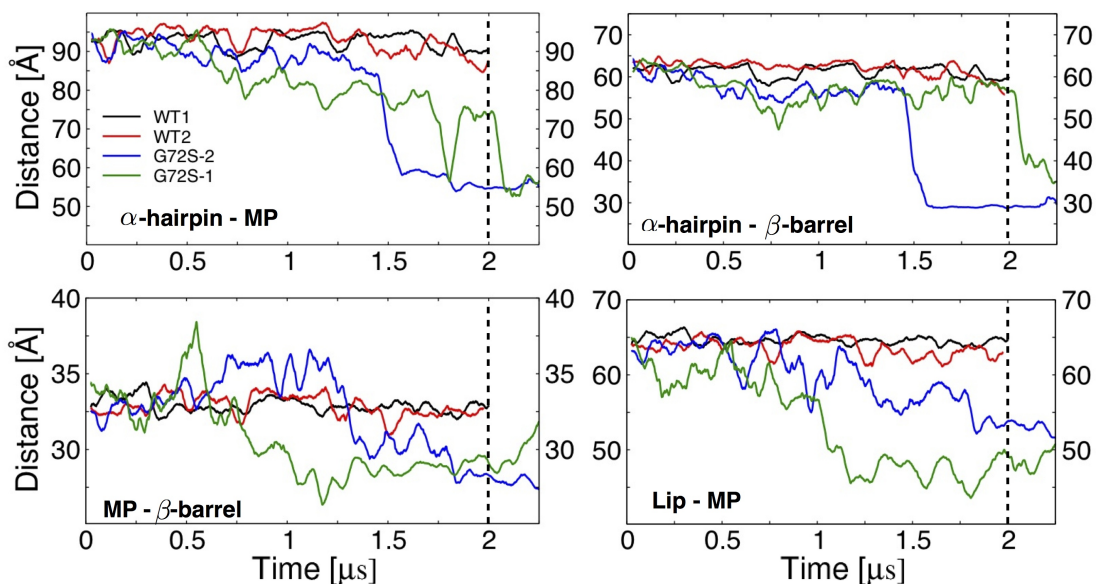
Figure 4.2.5: Domain distance evolution for WT1 [Red], WT2 [Black], G72S-1 [Green] and G72S-2 [Blue] and extended simulation for G72S-1 [Green] and G72S-2 [Blue].
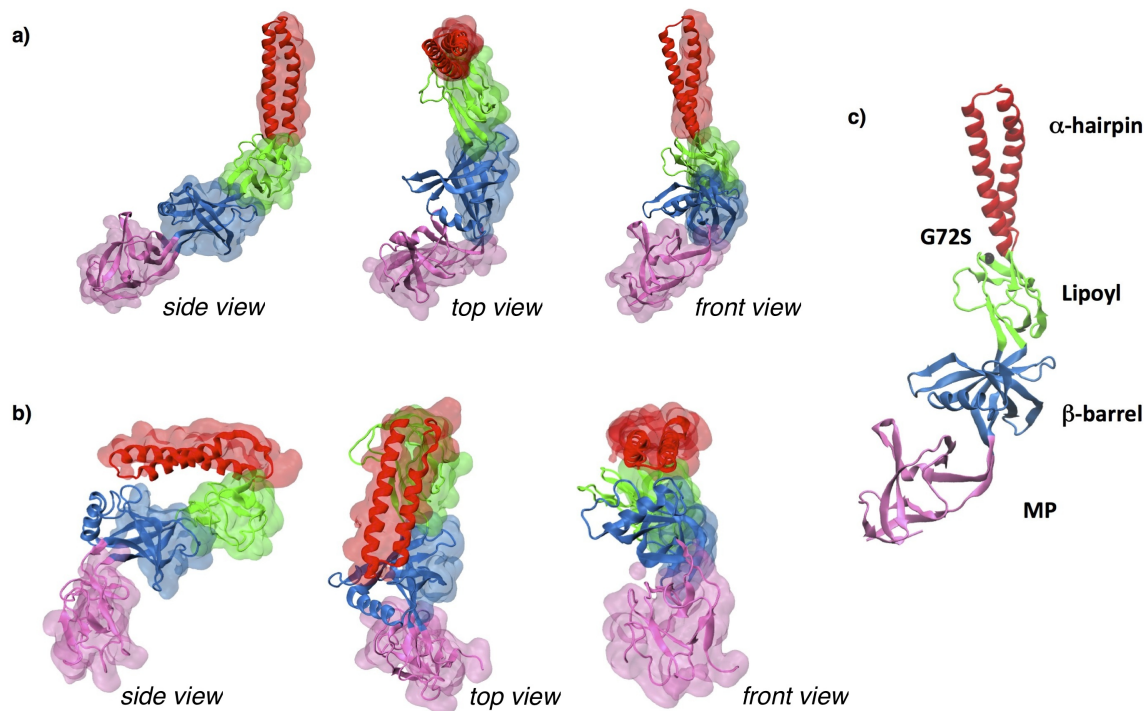
Interestingly, the visual inspection of the trajectory and the previous analysis suggest that the final conformation of the G72S variant is stable. To confirm this hypothesis, further investigations are needed. A quantitative description of the stability of this new conformation is provided by the analysis of bonds that may be established or broken during the dynamics. The analysis focused on the last 1 μs of the simulations, where the new configurations are more persistent and the systems reached an equilibrium conformation. As shown in Figure 4.2.4 the mutant showed a bending of the *MP* and the *α-hairpin* domain in the *β–barrel* direction. The close position between the *MP* and the *β–barrel* domain is stabilized by several new H-bonds, established during the MD simulation, as shown in Figure 4.2.7a, where the thickness of the residues indicates the strength of the bonds (for a quantitative description of the H-bonds persistency see Table 4.2.1).

The H-bonds analysis also revealed that only for the G72S-1 mutant new H-bonds between the α-*hairpin* and the β–*barrel* domain are established, as shown in Figure 4.2.7b and 4.2.7c. In fact, the α-*hairpin* of G72S-2 does not establish any H-bonds with the other domains. Thus, the analysis of salt bridges among the domains was performed in order to investigate other stabilizing mechanisms that might be involved.

Residue	Occupancy (%)
R326-E53	80.8
R326-N51	69.3
K232-G329	14.9
K232-D301	12.8

(a)

Residues	Occupancy (%)
K232-D325	89.5
D235-K303	77.7
R326-E242	48.1
K287-E213	40.74
Q281-D325	39.8
V327-Q238	33.1
D235-Q305	30.5
R326-Q283	29.3

(b)

Table 4.2.1: Occurrence frequency of the new H-bonds established between the *MP* and the β–*barrel* domain during the MD simulations for **a)** G72S-1 and **b)** G72S-2 mutated systems.

The analysis revealed that salt bridges involving different domains were established in the last part of the two simulations. In particular, for G72S-1 the α-*hairpin* and the β–*barrel* domain are bounded by three salt bridges, which increase the stability conferred by the H-bonds. In G72S-2, the analysis showed that the α-*hairpin* established one salt bridge with the *lipoyl* domain as shown in Figure 4.2.6. Interestingly, in both the systems the same residue in the α-*hairpin*, D136, is involved in salt bridges, although with different domains: in G72S-1 the aspartic acid is bounded to two arginine residues in the β–*barrel* domain, while in G72S-2 it is bounded to an arginine residue in the *lipoyl* domain. These results suggest that, during the simu-

lation, G72S-2 have not reached a conformation as stable as G72S-1. The previous analysis showed that the structure of the MexA mutant reached a largely different conformation respect to the wild type. Moreover, such conformation is stabilized by several bonds established among the different domains during the simulations. Further analysis showed that these structural changes are coupled to remarkable losses of secondary structure as shown in Figure 4.2.8a. The analysis was performed using the average-close structure from the *a)* two most populated clusters for the wild type case and *b)* five most populated clusters for the mutant. The residues involved in the mutant secondary structure loss are mainly located in the *a–hairpin* and in the *β–barrel* domains (see Figure 4.2.8b) which are involved in MexA-OprM and MexA-MexB interaction respectively.
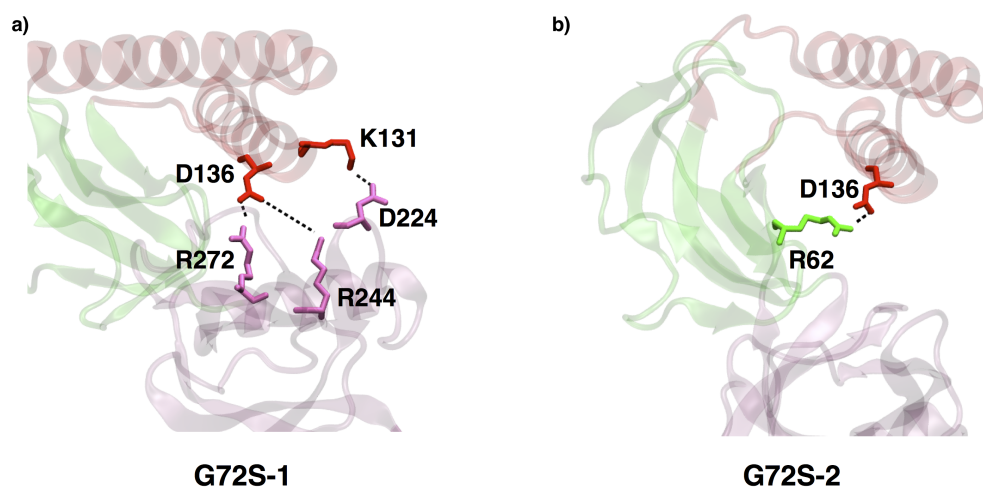
Figure 4.2.6: Salt bridges established during the MD simulations between **a)** *β-barrel* [light Mauve] and *α-hairpin* domains for G72S-1; **b)** *α-hairpin* and *lipoyl* domains for G72S-2.
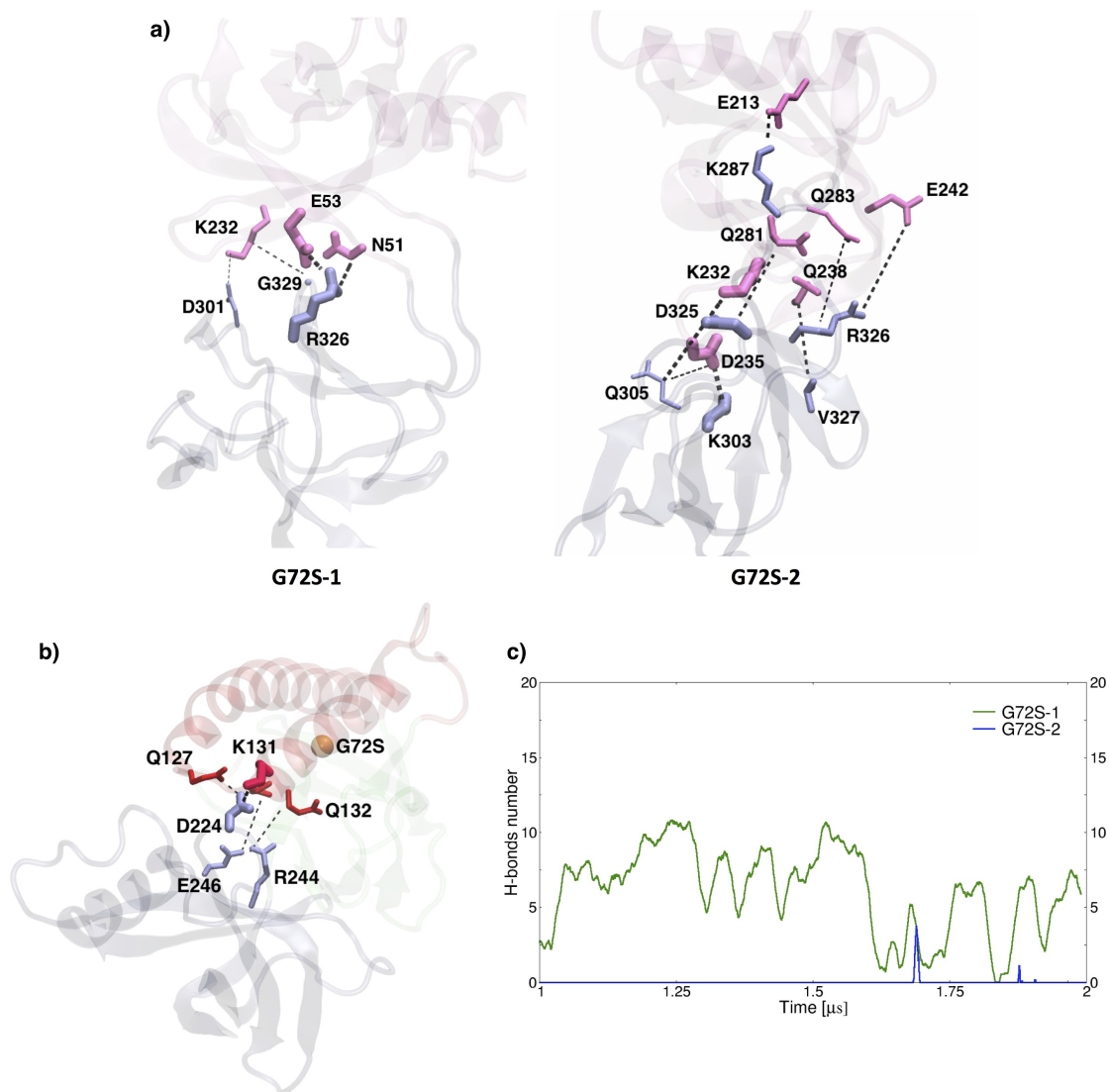
Figure 4.2.7: New H-bonds established during the MD simulations. The size of the residues involved in the H-bonds is proportional to the persistence of the bonds while the color refers to their correspondent domains. **a)** H-bonds between the *MP* [light Blue] and the *β-barrel* [light Mauve] domain. The G72S-1 showed less H-bonds between the two domain respect to G72S-2 but their occurrence frequency is high enough to stabilize the domains in this new conformation. **b)** H-bonds between the *α-hairpin* [light Red] and the *β-barrel* [light Blue] domain for G72S-1. The dimension of the residues involved in the H-bonds is proportional to the persistence of the bonds while the color refers to their correspondent domains. The mutation is identified with an orange sphere. **c)** The number of H-bonds between the *α-hairpin* and the *β-barrel* domain for G72S-1 [Green] and G72S-2 [Blue].
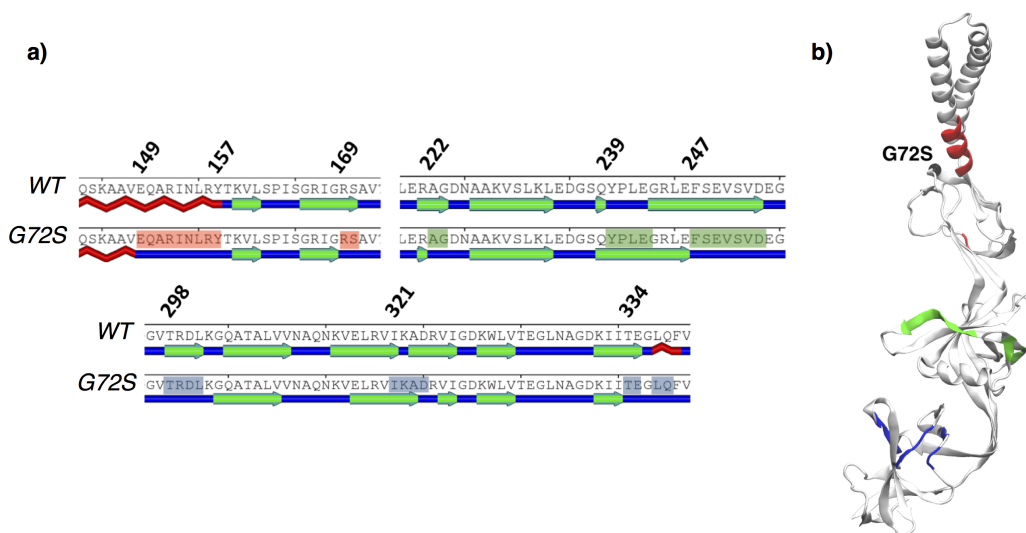
Figure 4.2.8: **a)** Secondary structure differences between the average-close structure of the cluster representatives for WT and G72S. **b)** Residues involved in the secondary loss, which is mainly located in the *a–hairpin* [Red], *β–barrel* [Green] and *MP* domain [Blue]. The mutation site is identified by a black sphere. The Secondary Structure was obtained with POLYVIEW-2D [280]

4.2.2   MexA dimeric unit

The analysis of the global behavior of the homology dimeric unit revealed a very stable structure during the the simulation, with very low values of the Cα-RMSD, as shown in Figure 4.2.9. In fact, a comparison with the relative high Cα-RMSD values for the wild type monomers suggests a flexibility loss in the dimeric unit. This is also confirmed by the Cα-RMSF analysis that suggest a more rigid structure for the dimer, since the atomic fluctuations are much less than in the wild type monomeric case (see Figure 4.2.10 for more details). In order to test the dimerization process, a structure of possible MexA dimer is needed as reference. Thus, the cluster analysis was performed to extract a representative structure of the MD simulation

for the homology dimer. The cluster analysis confirmed the stability of the system by identifying one main cluster for the dimeric unit that contains more than 90% of the whole population. The representative structure of this cluster was used as reference to validate the results of the docking analysis, as described in the next paragraphs.
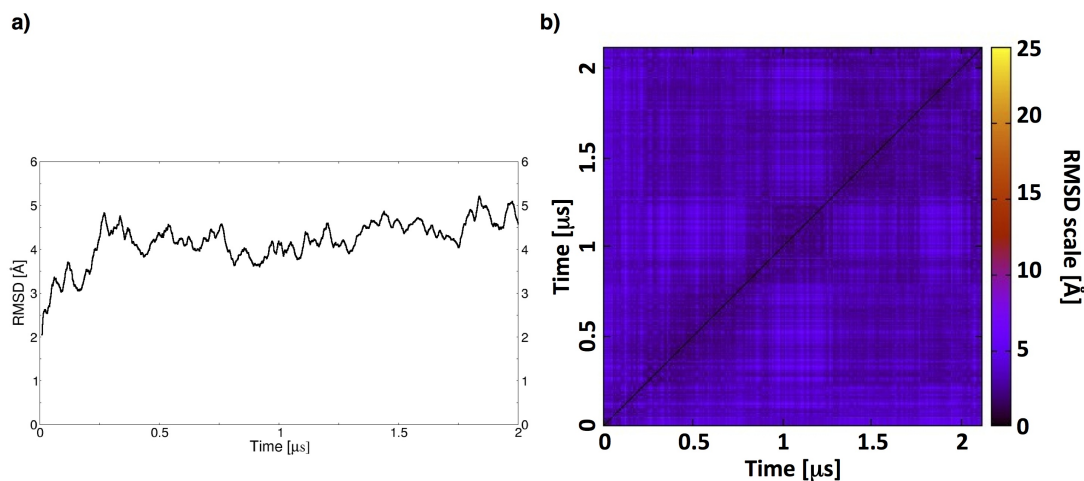
Figure 4.2.9: **a)** Cα-RMSD of the dimeric homology model and **b)** Cα-RMSD matrix.
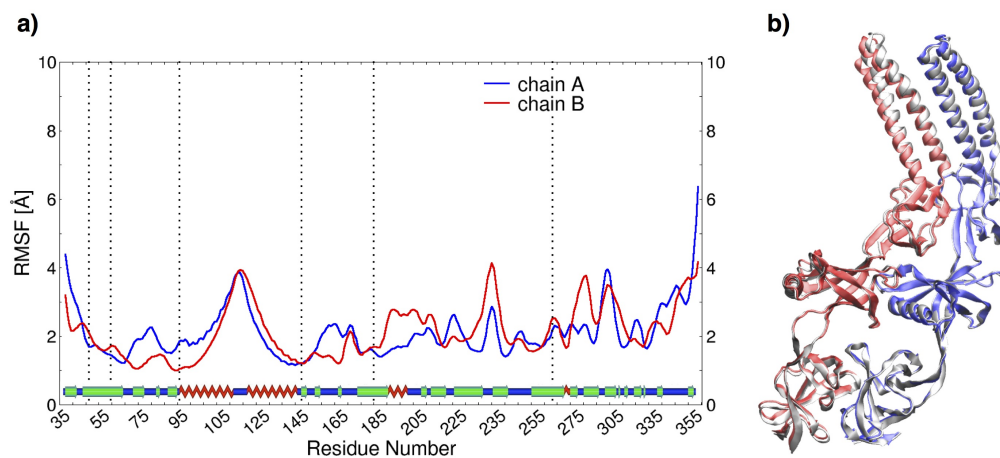
Figure 4.2.10: **a)** Cα-RMSF of the dimeric homology model for chain A [Blue] and chain B [Red]. **b)** Structure of the homology dimer compared with the crystal structure [White].
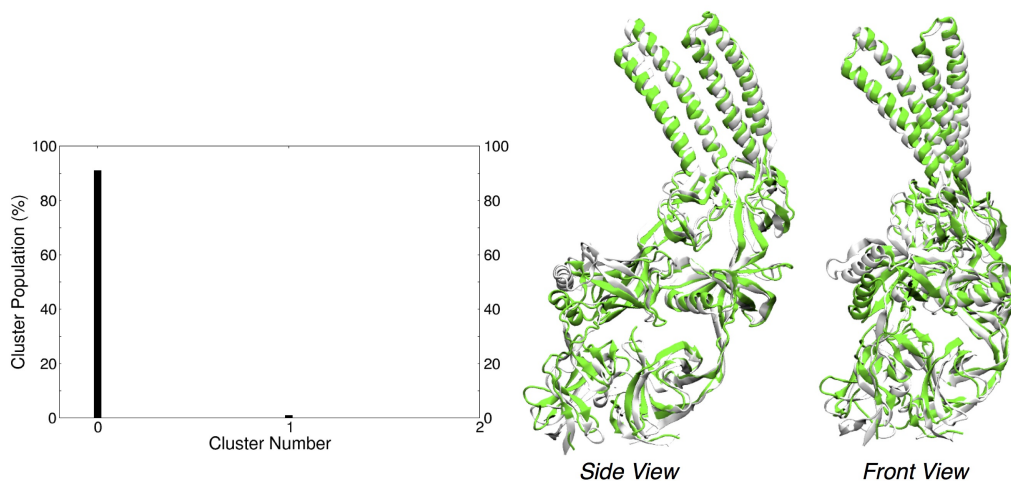
Figure 4.2.11: Cluster population for the dimeric homology model and comparison between the initial structure [White] and the representative of the cluster 0 [Green].

4.2.3 MexA dimerization

The analysis of the MD simulations for MexA monomers confirmed a more compact structure for the mutant with a flexibility loss in the domains involved in MexA-OprM and MexA-MexB interaction. Furthermore, the final configuration of the mutant is stabilized by several H-bonds and Salt bridges. These results suggest large differences also in the multimerization process. Thus, before running the docking simulations it is necessary to select MexA structures that can be representatives of the main conformations of the protein. The cluster analysis provides a way to select representative structures of the dynamics by collecting common structures together. For this analysis, the average-linkage algorithm [281] was used with a 4Å cutoff. The results confirmed, as expected, differences between the wild type and the mutant (see Figure 4.2.12). The two simulations of the wild type have a peaked distribution with the first two clusters that account for more than 50% of the population, while wider distributions were found for the mutant.
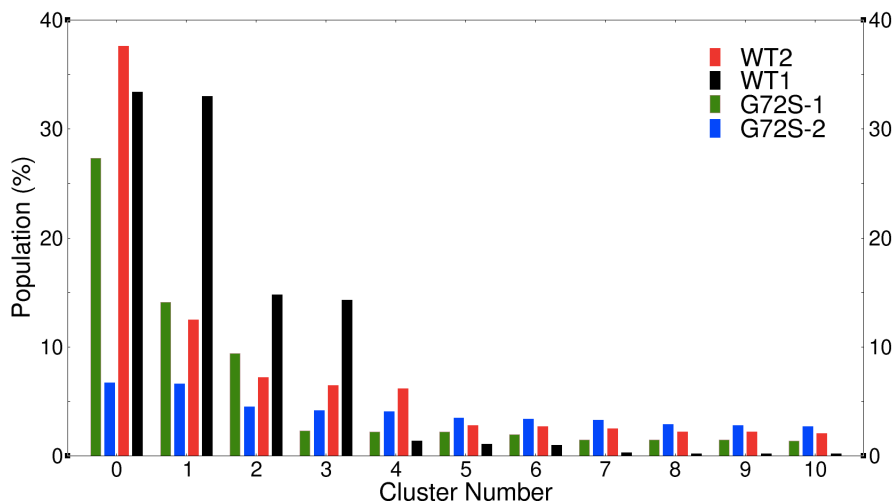
Figure 4.2.12: Distribution of the cluster population for the four MD simulations. The two wild type simulations, WT1 [Black] and WT2 [Red], show a peaked distribution of the cluster population while the simulations of the mutant, G72S-1 [Green] and G72S-2 [Blue], showed wider distributions, in particular for G72S-2 system.

Thus, for the wild type the docking runs were performed with the top five clusters representatives in order to take into account more than 90% of the population. To reach the same population percentage for the mutant, it was necessary to consider the top 50 clusters for the docking runs. An example of the differences among the representative structures of the wt and the G72S variant is shown in Figure 4.2.13 while in Table 4.2.2 is reported a quantitative description, in terms of $C\alpha$-RMSD.

Clus	WT1	WT2	G72S-1	G72S-2	Clus	WT1	WT2	G72S-1	G72S-2
0	9.1	3.6	18.8	15.5	0	-	8.3	18.2	15.5
1	4.5	5.5	18.6	20.1	1	7.8	8.8	18.4	18.4
2	5.5	5.8	16.9	19.1	2	5.2	10.7	16.6	14.5
3	3.3	6.1	12.8	14.5	3	7.2	6.9	10.8	15.7
4	8.7	4.3	16.8	14.1	4	5.3	10.1	16.8	13.3

(a)   (b)

Clus	WT1	WT2	G72S-1	G72S-2	Clus	WT1	WT2	G72S-1	G72S-2
0	8.3	-	19.9	14.7	0	18.2	19.9	-	16.6
1	2.4	2.2	19.7	20.5	1	19.7	20.1	3.9	9.9
2	3.7	4.6	18.1	18.4	2	19.3	19.4	5.2	17.6
3	2.8	3.7	11.7	14.0	3	19.0	19.2	15.7	15.8
4	6.7	4.9	18.0	13.1	4	19.3	20.1	4.9	17.2

(c)   (d)

Table 4.2.2: Cα-RMSD between the representatives of the top five clusters and **a)** chain L in the crystal structure (PDB code 2V4D.pdb; **b)** representative of cluster 0 for the WT1 ; **c)** representative of cluster 0 for the WT2; **d)** representative of cluster 0 for the G72S-1. The Cα-RMSD with the representative of G72S-2 as reference is not shown as the wider distribution of this system prevents the choice of a good reference structure.

After the selection of the structures, the docking was performed with soft harmonic restraints in order prevent unphysical results as described in paragraph 4.1.4. The selection criteria of the resulting structures was based on the Cα-RMSD between the homology model for the dimer and the structures extracted from the docking. The Cα-RMSD cut-off value was set to 10 Å in order to consider the possible distortions induced by the docking process. Among all the possible dimers produced during the docking, for the wild type twelve structures were selected using the dimeric homology model as reference. The docking runs for the mutant produced, in fact, largely different conformations although in few cases the results were comparable

to the wild type case. However, a deeper analysis of these cases showed that they correspond to conformations that MexA maintained for a very short time, before the inter-domain rearrangements. This suggests that these structures might be not suitable to be representatives of the dynamics. In addition, according to the cluster analysis they accounted only for very few percentage of the population.
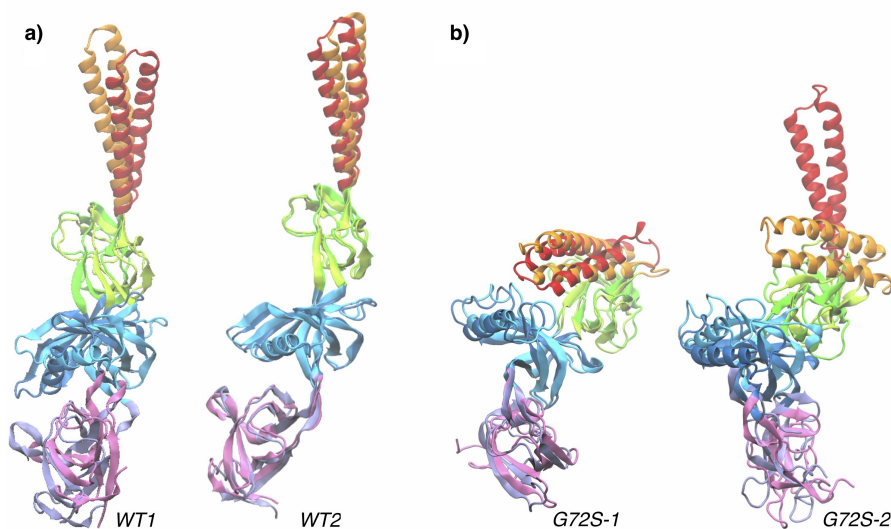
Figure 4.2.13: **a)** Representatives of the two most populated clusters for WT1 and WT2. The two simulations led to similar structures, although the WT1 showed higher flexibility in the a–$hairpin$ and β–$barrel$ domains. **b)** Representatives of the same clusters for G72S-1 and G72S-2. In both cases the color sequence $red - green - blue - mauve$ identifies the representative of cluster 0 while the $orange - green2 - cyan - deep\_blue$ refers to the representative of cluster 1. The simulations for the mutant led to different representatives structures with respect to the wild type case. In G72S-1 case both the representatives showed the bending of the a–$hairpin$ and MP domains in the β–$barrel$ direction. The G72S-2 showed a new position of the MP domains in the β–$barrel$ direction for the cluster 0 while the bending of the a–$hairpin$ is only in cluster 1.

Thus, only the wild type structures were used for further analysis. To study the stability of such dimers, they were first optimized and then new MD simulations were run for the dimeric candidates with the same procedures described in para-

graph 4.1.3. During the first steps of the MD simulations the structures reached a more stable conformation with respect to their initial structure with the exception of the systems identified with names *dock4* and *dock10*, as shown in Figure 4.2.14b. However the Cα-RMSD evolution, using the homology dimer as reference, showed a stable behavior for all the systems, with relatively small Cα-RMSD values (see Figure 4.2.14a). A new cluster analysis, based on the same criteria used for the monomers, led to representative structures relatively close to the homology model as shown in Table 4.2.3.

dock number	1	2	3	4	5	6	7	8	9	10	11	12
RMSD [Å]	6.7	6.5	7.9	6.4	6.7	6.7	6.0	4.6	8.8	6.9	5.9	6.1

(a)

dock number	1	2	3	4	5	6	7	8	9	10	11	12
RMSD [Å]	3.1	6.6	8.2	6.0	6.8	7.1	5.8	5.8	6.3	5.1	5.5	5.7

(b)

dock number	1	2	3	4	5	6	7	8	9	10	11	12
RMSD [Å]	3.4	4.9	4.3	5.0	4.6	3.1	4.2	5.1	4.4	2.9	4.6	4.6

(c)

dock number	1	2	3	4	5	6	7	8	9	10	11	12
RMSD [Å]	4.6	6.0	4.8	5.6	5.4	2.8	5.3	3.5	7.1	5.5	5.7	5.4

(d)

dock number	1	2	3	4	5	6	7	8	9	10	11	12
RMSD [Å]	11.0	7.8	11.9	8.3	8.9	10.5	7.8	4.3	11.9	10.5	5.1	5.5

(e)

Table 4.2.3: Cα-RMSD between the homology model, used as reference, and the representatives of the most populated cluster of the dimer candidates after a global alignment, for **a)** whole system, **b)** *α-hairpin* domain, **c)** *lipoyl* domain, **d)** *β-barrel* domain, **e)** *MP* domain.

The main differences among the systems are located in the *a–hairpin, β–barrel* and the *MP* domains (see Table 4.2.3b,d,e) while the *lipoyl* domain of all the systems is closer to the homology dimeric model, in terms of Cα-RMSD. The *lipoyl* domain has been predicted to be the responsible of the MexA-MexA dimerization [92] and this result might suggest that the two MexA *lipoyl* domains, in case of dimerization, bind always in the same way. On the other hand, the differences in the domains that are known to be involved in MexA-MexB and MexA-OprM interactions [60, 86, 88, 114, 146–148] suggest an high flexibility of the dimer to adapt to the surface of the other components of the assembly (see Figure 4.2.15 and 4.2.16a,b).

The analysis of the surface exposed to the solvent of the dimeric model revealed a common behavior of both the homology dimer and the dimers candidates from the docking analysis. The comparison with the docking results was possible only for the first 500 ns of the MD simulation of the dimeric homology model. However, the SASA analysis for the whole trajectory of the dimer alone revealed not large changes in the surface exposed to the solvent. Combining the Cα-RMSD results and the SASA analysis for the first 500 ns of the simulation it is plausible to deduce a similar behavior for the structures resulting from the docking analysis with a stabilization of the hydrophobic residues of the dimeric candidates.
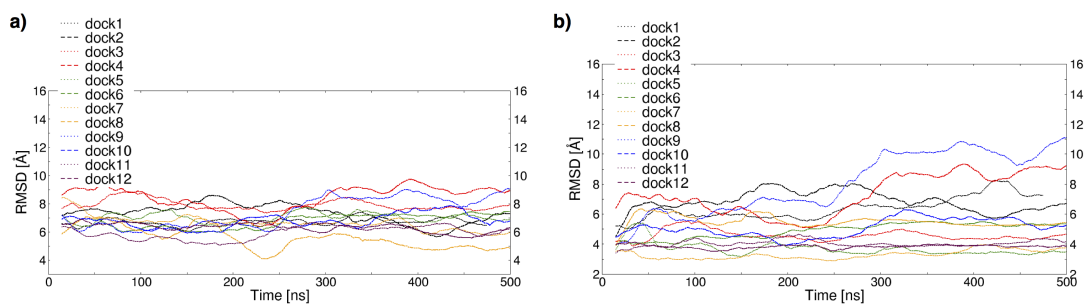
Figure 4.2.14: Cα-RMSD evolution of the twelve dimeric candidates selected after the docking analysis with respect **a)** the homology dimeric model and **b)** their own initial structure.
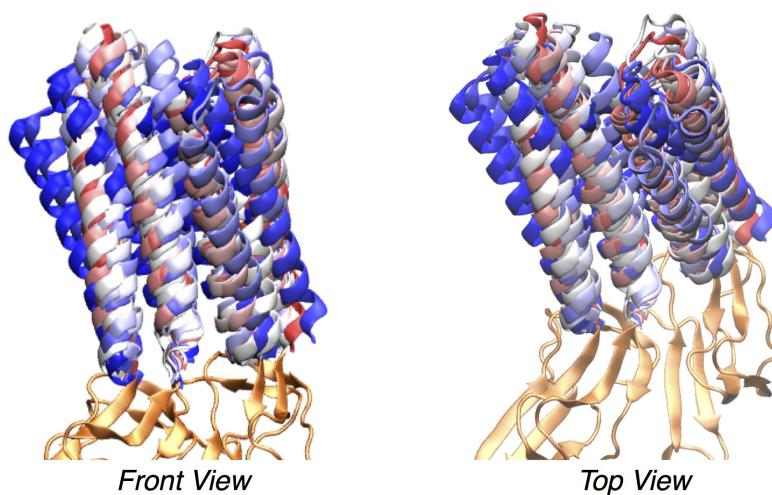
Figure 4.2.15: Differences in the *a–hairpin* domain between the representatives of the most populated cluster of each dimeric candidate. The color range (*TimeStep* color with VMD) identifies the structures from the homology model [dark Blue] to *dock12* system [dark Red].

Front View   *Top View*

(a)

(b)

Figure 4.2.16: Differences between the representatives of the most populated cluster of each dimeric candidate in the **a)**β–*barrel* domain and **a)** *MP* domain. The color range (*TimeStep* color with VMD) identifies the structures from the homology model [dark Blue] to *dock12* system [dark Red].
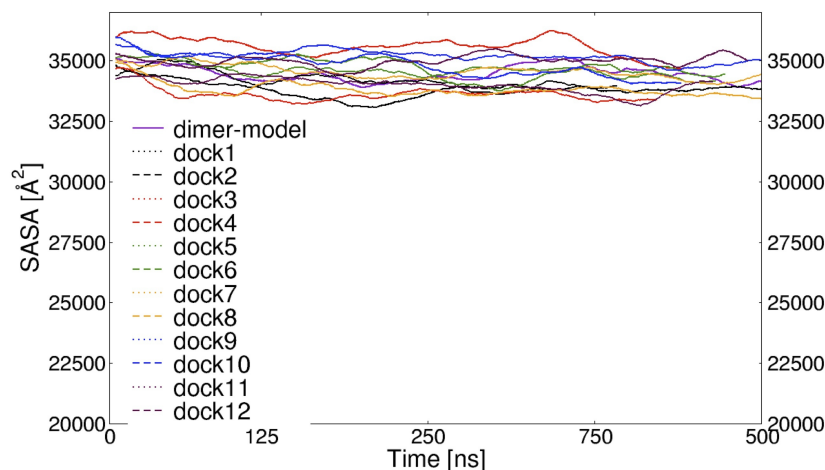
Figure 4.2.17: SASA comparison between the representative structure of the homology dimeric unit and the structures produced during the docking analysis.

4.3 Discussion

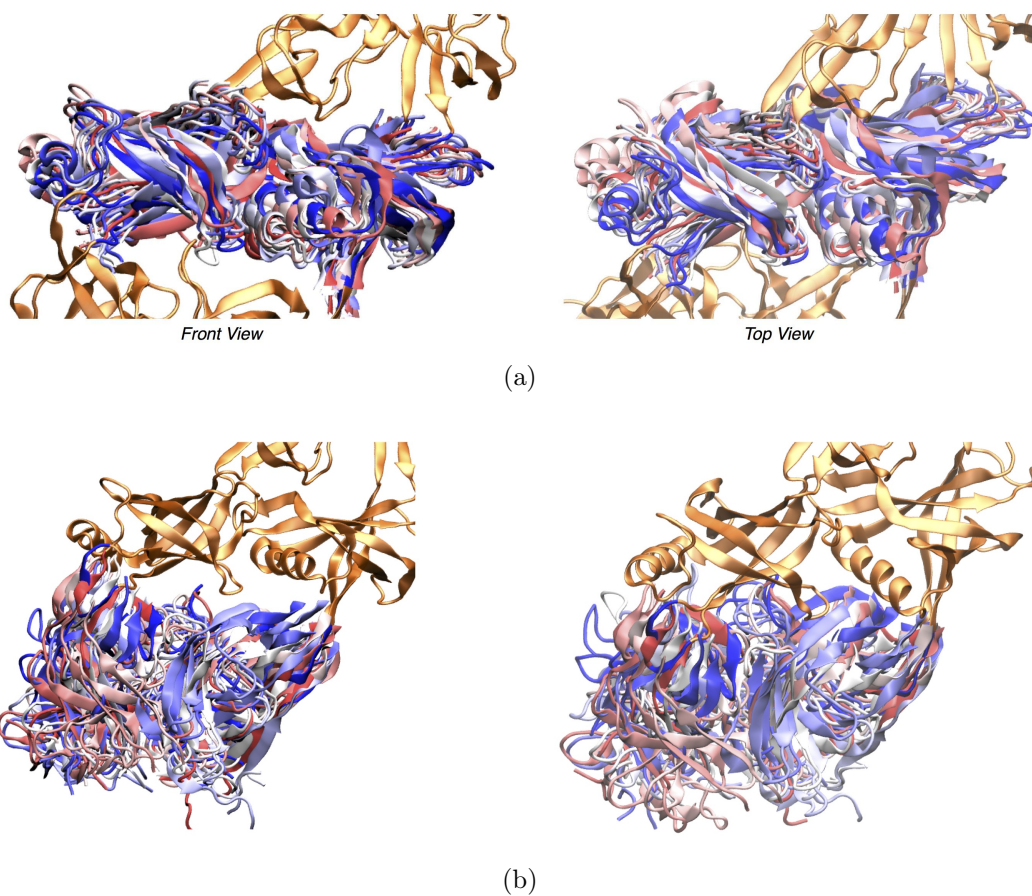The Membrane Fusion Protein MexA is known to be fundamental in the MexAB-OprM functionality, as its participation is needed to either help the OprM opening [86, 114, 282] and to confer stability to the complex. However the number of MexA proteins that participate in the assembly is still under debate as well as the exact role in the extrusion process [55, 86–88, 114, 273]. Despite the different models that describe the possible stoichiometry of the assembly, several experiments suggest the possibility for MexA, and for MFPs in general, to form dimeric units [86, 88, 92, 99, 274, 276–278, 282]. In order to investigate such possibility, in this chapter was studied the stability of monomeric MexA units in both wild type and mutant version. The mutation was chosen among many others as it has been postulated [92] to impair the MexA dimerization, leading to a not-functional assembly. The monomeric structures, studied with extended MD simulations, showed an high flexibility but only

for the wild type case the structure remained stable as the flexibility of the mutant led to large structural changes. These changes in the mutated monomer showed a bending of the α-*hairpin* and the *MP* domain in the β-*barrel* domain direction and arose together with secondary structure loss in domains that are involved in both MexA-OprM and MexA-MexB interactions. Despite the large differences between the wild type and the mutant, the dimerization was studied for both the systems with a protein-protein docking program. To validate the docking results, a dimeric homology model was produced and studied with a new extended MD simulation. A comparison between the homology model for the dimer and the docking results showed that only for the wild type good dimeric candidates were produced. As expected, the large structural changes in the mutant impaired the possibility to produce a reasonable dimeric structure.

The dimeric candidates, however, showed large differences in the α-*hairpin* and the *MP* domains. To validate such models and to investigate the possibility of a convergence to a common structure, new MD simulations were run for each candidate. The results showed that the dimeric candidates reached an equilibrium structure that is close to the representative structure of the dimeric homology model. All these results suggest that the MexA wild type monomer is able to form dimers and that, despite the initial differences in the possible dimers, they reached a stable conformation. It is worthwhile to point out that the main differences, after the equilibrium, are located in domains involved in the binding with the asymmetric structures of MexB and OprM. This suggests that the dimers, although less flexible than the monomers, still maintain the flexibility to adapt to the OprM and MexB surfaces.

Chapter 5

Conclusions and perspectives

In the context of multi-drug resistance in Gram-negative bacteria, this thesis focused on two main key-points *a)* a better understanding of the inhibition mechanisms of the RND transporter AcrB and *b)* a study of the possible dimerization of the membrane fusion protein MexA.

The first part of this thesis, described in Chapter 3, addressed the study of the RND transporter AcrB expressed in *S. enterica.* This study concerns the first case in which MDR was directly connected to a mutation in RND efflux pumps in clinical settings. In particular, during the treatments of *S.typhimurium* infection, this serovar developed resistance to the ciprofloxacin antibiotic. The analysis of the genome sequence of the isolates revealed that *S. typhimurium,* as consequence of the therapy, mutated the AcrB distal binding pocket by replacing a glycine residue with an aspartic acid. To study the effects of the G288D mutation, homology models for the mutant and wild type were built using the AcrB crystal structure for *E. coli* [140] as template. Subsequently, multiple MD simulations were performed for each system revealing local structural changes in the distal pocket. The presence of an aspartate,

also increased the polarity of the region, affecting its hydration, as confirmed with the analysis of the radial distribution function profiles. Furthermore, the possible effect of the mutation on ciprofloxacin, doxorubicin and minocycline efflux was studied. The comparison of the results revealed possible steric clashes with both doxorubicin and minocycline. On the other hand, the mutation should not impair the binding of ciprofloxacin, which may form H-bonds between its hydroxyl group, the aspartate residue and the resident waters around it. This possibility was suggested to explain its recognition with a consequent increased efflux of, and decreased susceptibility to, ciprofloxacin and other fluoroquinolones.

The second part of the work of this thesis, described in Chapter 4, focused on the role of the MFP component of the MexAB-OprM efflux pump in the assembly stoichiometry. The uncertainties in the composition of the assembly concern mainly the number of the membrane fusion proteins involved. The two main stoichiometry models predict the presence of three [55, 103, 273] or six [86, 88, 274, 283] MexA proteins, suggesting different roles for the periplasmic component of the pump in the extrusion mechanism. In fact, in the first model a direct connection between the RND transporter MexB and the outer membrane channel OprM is predicted with three MexA proteins that participate in the OprM opening and confer stability to the assembly. On the other hand, the second model predicts the participation of six MexA proteins with a more active role in the efflux mechanism, since the direct connection between MexB and OprM should not be established anymore. An hexameric funnel-like structure of MexA is in fact predicted to be inserted between the other two components of the pump [86, 88]. In order to gain insights in the assembly stoichiometry, the study focused on the (supposed) very first step of the MexA hexamer building, i.e. the dimerization, following two main strategies. The first, concerns the MexA monomeric structure for which the wild type and a mu-

tated MexA were studied. The mutation selected for this study is known to impair the MexAB-OprM efflux pump, presumably by affecting the MexA dimerization [92]. Representatives structures from MD simulations were used to study the dimerization process with a protein-protein docking program. The second strategy was to build an homology model for the MexA dimer and study its stability with an extended MD simulation. A comparison between the results of MexA-MexA docking and the homology dimer was then performed, confirming that only for the wild type case acceptable dimers could be obtained. Further MD simulations for the docked structures confirm a reasonable convergence to the homology model. In particular the *lipoyl* domain, supposed to be involved in the MexA-MexA multimerization, showed always the same structure with very low differences between the docked structures and the dimeric homology model in term of RMSD. The other domains, involved in MexA-OprM and MexA-MexB binding, showed almost the same structure, but with higher differences with respect the the dimeric homology model. These results suggest that the MexA dimerization seems to be possible and the resulting structures are very similar, with differences mainly located in domains that have already been proved to be very flexible [145, 148] in order to adapt to the asymmetric structures of the other components of the pump.

Thus, this study confirms the possibility of MexA dimerization but other aspects of the 3:6:3 model remain unclear like, as example, how three dimers could interact to form the hexameric funnel-like structure or how this structure could be involved in the OprM opening. Further investigations are, thus, necessary to deepen the knowledge about the assembly building and stability of the MexAB-OprM efflux pump as well as its homologues like, as example, AcrAB-TolC.

Acknowledgments

It is time to write one the most important Chapters of each thesis: the acknowledgments. The importance of this section is usually underestimated but reaching this point has several important meanings. First, it means that the thesis is finished and this is certainly a good news for the both the writer and the reader. Secondly, it implies that I finished my PhD and this is exciting and terrifying at the same time! Nevertheless, it also means that I can officially thank all the person who helped me during these years. However, since this Chapter should not be longer than the previous ones, I must sum many persons in categories, hoping not to be impolite.

First of all, I would like to thank my supervisors Paolo and Attilio because they are the best supervisors one student could have. I tried very hard to test their famous patience but I must say I have failed: they helped me anyway! I would like to say thank also to Giuliano, for his informal support, Giovanni and Andrea for their technical help and all the people of the research group: Andrea, Tommaso, Silvia, Venkata, Susruta, Debhia, Federico and Daniela.

The next thank goes to my colleagues of the XXVII cycle, in particular the ones that shared the office with me: Federica, Vasco and Mario. I would like to thank also all the friends I met during the PhD schools in Trieste and Amsterdam, because everyone taught me something important. How can't I say thank to all the old friends

from the very first year of University and the new ones that joined us?

Finally, the last part of the acknowledgments is for my family, which made all of this possible, supporting me in all my choices, the good ones and the bad ones. However, the very last and the most important acknowledgment is for my dear love Debby, which more than everyone else knows how difficult is to stand me! Despite my "not really good" character, she's always at my side and I think I will never understand why should I deserve such love.

Bibliography

[1] Orcutt BN, Sylvan JB, Knab NJ, and Edwards KJ. Microbial ecology of the dark ocean above, at, and below the seafloor. *Microbiology and Molecular Biology Review*, 75(2):361–422, 2011.

[2] Hauptmann AL, Stibal M, Bælum J, Sicheritz-Pontén T, Brunak S, Bowman JS, Hansen LH, Jacobsen CS, and Blom N. Bacterial diversity in snow on north pole ice floes. *Extremophiles*, 18(6):945–951, 2014.

[3] Bassil NM, Bryan N, and Lloyd JR. Microbial degradation of isosaccharinic acid at high ph. *The ISME Journal*, 9:310–320, 2015.

[4] Kim W, Tengra FK, Young Z, Shong J, Marchand N, Chan HK, Pangule RC, Parra M, Dordick JS, and Plawsky JL. Spaceflight promotes biofilm formation by pseudomonas aeruginosa. *PLoS ONE*, 8(4):1932–6203, 2013.

[5] Alberts B, Johnson A, Walter P, Lewis J, Ra M, and Roberts K. Molecular biology of the cell (5th edn). *Taylor and Francis*, 2007.

[6] Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA, Deal C, et al. The nih human microbiome project. *Genome research*, 19(12):2317–2323, 2009.

[7] Human Microbiome Project Consortium et al. Structure, function and diversity of the healthy human microbiome. *Nature*, 486(7402):207–214, 2012.

[8] Poole K. Multidrug resistance in Gram-negative bacteria. *Curr Opin Microbiol*, 4(5):500–508, January 2001.

[9] Nikaido H. Multidrug resistance in bacteria. *Annual review of biochemistry*, 78:119–46, January 2009.

[10] Howell L. Global risks 2013: An initiative of risk response newtork. *World Economic Forum*, 2013.

[11] Blair JMA, Richmond GE, and Piddock LJV. Multidrug efflux pumps in gram-negative bacteria and their role in antibiotic resistance. *Future microbiology*, 9(10):1165–1177, 2014.

[12] Livermore DM. The need for new antibiotics. *Clinical Microbiology and Infection*, 10:1–9, 2004.

[13] Levy SB and Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. *Nature medicine*, 10:S122–S129, 2004.

[14] World Health Organization. *Global status report on alcohol and health-2014*. World Health Organization, 2014.

[15] Piddock LJV. Multidrug-resistance efflux pumps? not just for resistance. *Nat Rev Micro*, 4(8):629–636, August 2006.

[16] Tenover FC. Mechanisms of antimicrobial resistance in bacteria. *Am J Med*, 119(6):S3–S10, 2006.

[17] Alekshun MN and Levy SB. Molecular mechanisms of antibacterial multidrug resistance. *Cell*, 128(6):1037–1050, 2007.

[18] Thanassi DG, Cheng LW, and Nikaido H. Active efflux of bile salts by escherichia coli. *J Bacteriol*, 179(8):2512–2518, 1997.

[19] Jones HE, Holland IB, Jacq A, Wall T, and Campbell AK. Escherichia coli lacking the acrab multidrug efflux pump also lacks nonproteinaceous, phb–polyphosphate ca 2+ channels in the membrane. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1612(1):90–97, 2003.

[20] Ma D, Cook DN, Hearst JE, and Nikaido H. Efflux pumps and drug resistance in gram-negative bacteria. *Trends Microbiol*, 2:489–493, 1994.

[21] Nikaido H. Multidrug efflux pumps of gram-negative bacteria. *J Bacteriol*, 178(20):5853, 1996.

[22] Nikaido H. Preventing drug access to targets: cell surface permeability barriers and active efflux in bacteria. In *Seminars in cell and developmental biology*, volume 12, pages 215–223. Elsevier, 2001.

[23] Paulsen IT. Multidrug efflux pumps and resistance: regulation and evolution. *Curr. Opin. Microbiol.*, 6:446–451, 2003.

[24] Hooper DC. Efflux pumps and noscocomial antibiotics resistance: a primer for hospita epidemiologits. *Clin. Infect. Dis.*, 40:1811–1817, 2005.

[25] Nikaido H and Pagès JM. Broad-specificity efflux pumps and their role in multidrug resistance of gram-negative bacteria. *FEMS Microbiology Reviews*, 36(2):340–363, 03 2012.

[26] Lomovskaya O, Lee A, Hoshino K, Ishida H, Mistry A, Warren MS, Boyer E, Chamberland S, and Lee VJ. Use of a genetic approach to evaluate the consequences of inhibition of efflux pumps in pseudomonas aeruginosa. *Antimicrob Agents and Chemother*, 43(6):1340–1346, 1999.

[27] Lomovskaya O and Watkins W. Inhibition of efflux pumps as a novel approach to combat drug resistance in bacteria. *J Mol Microbiol Biotechnol*, 3(2):225–236, 2001.

[28] Lomovskaya O, Warren MS, Lee A, Galazzo J, Fronko R, Lee M, Blais J, Cho D, Chamberland S, Renau T, et al. Identification and characterization of inhibitors of multidrug resistance efflux pumps in pseudomonas aeruginosa: novel agents for combination therapy. *Antimicrobial Agents and Chemotherapy*, 45(1):105–116, 2001.

[29] Martins M, Dastidar SG, Fanning S, Kristiansen JE, Molnar J, Pages JM, Schelz Z, Spengler G, Viveiros M, and Amaral L. Potential role of non-antibiotics (helper compounds) in the treatment of multidrug-resistant gram-negative infections: mechanisms for their direct and indirect activities. *Int J Antimicrob Agents*, 31(3):198–208, 2008.

[30] Pagès JM and Amaral L. Mechanisms of drug efflux and strategies to combat them: Challenging the efflux pump of gram-negative bacteria. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1794(5):826–833, 2009.

[31] Blair JMA and Piddock LJV. Structure, function and inhibition of RND efflux pumps in Gram-negative bacteria: an update. *Curr Opin Microbiol*, 12(5):512–9, October 2009.

[32] XZ Li and H Nikaido. Efflux-Mediated Drug Resistance in Bacteria: an Update. *Drugs*, 69(12):1555–1623, 2009.

[33] Tal N and Schuldiner S. A coordinated network of transporters with overlapping specificities provides a robust survival strategy. *Proc Natl Acad Sci USA*, 106(22):9051–9056, 2009.

[34] Luckey M. *Membrane structural biology: with biochemical and biophysical foundations.* Cambridge University Press, 2014.

[35] Poole K. Pseudomonas aeruginosa: Resistance to the max. *Frontiers in Microbiology*, 2(65), 2011.

[36] Ricci V, Tzakas P, Buckley A, Coldham NC, and Piddock LJV. Ciprofloxacin-resistant salmonella enterica serovar typhimurium strains are difficult to select in the absence of acrb and tolc. *Antimicrob Agents Chemother*, 50(1):38–42, 2006.

[37] Baucheron S, Tyler S, Boyd D, Mulvey MR, E Chaslus-Dancla, and Cloeckaert A. Acrab-tolc directs efflux-mediated multidrug resistance in salmonella enterica serovar typhimurium dt104. *Antimicrob Agents Chemother*, 48(10):3729–3735, 2004.

[38] Buckley AM, Webber MA, Cooles S, Randall LP, La Ragione RM, Woodward MJ, and Piddock LJV. The acrab-tolc efflux system of salmonella enterica serovar typhimurium plays a role in pathogenesis. *Cell Microbiol*, 8(5):847–856, 2006.

[39] Hudault S, Guignot J, and Servin AL. Escherichia coli strains colonising the gastrointestinal tract protect germfree mice againstsalmonella typhimuriuminfection. *Gut*, 49(1):47–55, 2001.

[40] Behrman RE, Boyer KM, Petersen NJ, Farzaneh I, Pattison CP, Hart MC, and Maynard JE. An outbreak of gastroenteritis due to e. coli 0142 in a neonatal nursery. *J Pediatr*, 86(6):919–927, 1975.

[41] Källenius G, Svenson SB, Hultberg H, Möllby R, Helin I, Cedergren B, and Winberg J. Occurrence of p-fimbriated escherichia coli in urinary tract infections. *The Lancet*, 318(8260):1369–1372, 1981.

[42] Besser RE, Lett SM, Weber JT, Doyle MP, Barrett TJ, Wells JG, and Griffin PM. An outbreak of diarrhea and hemolytic uremic syndrome from escherichia coli o157: H7 in fresh-pressed apple cider. *Jama*, 269(17):2217–2220, 1993.

[43] Calbo E, Romaní V, Xercavins M, Gómez L, Vidal CG, Quintana S, Vila J, and Garau J. Risk factors for community-onset urinary tract infections due to escherichia coli harbouring extended-spectrum β-lactamases. *J Antimicrob Chemother*, 57(4):780–783, 2006.

[44] CDC. Antibiotic resistance threats in the united states. 2013.

[45] Nikaido H. Outer membrane barrier as a mechanism of antimicrobial resistance. *Antimicrobial Agents and Chemotherapy*, 33(11):1831, 1989.

[46] Okusu H, Ma D, and Nikaido H. Acrab efflux pump plays a major role in the antibiotic resistance phenotype of escherichia coli multiple-antibiotic-resistance (mar) mutants. *J Bacteriol*, 178(1):306–308, 1996.

[47] Zgurskaya HI and Nikaido H. Multidrug resistance mechanisms: drug efflux across two membranes. *Mol Microbiol*, 37:219–225, 2000.

[48] Webber MA and Piddock LJV. Absence of mutations in marrab or soxrs inacrb-overexpressing fluoroquinolone-resistant clinical and veterinary isolates of escherichia coli. *Antimicrob Agents Chemother*, 45(5):1550–1552, 2001.

[49] Baucheron S, Imberechtxu H, Chaslus-Dancla E, and Cloeckaert A. The acrb multidrug transporter plays a major role in high-level fluoroquinolone resistance in salmonella enterica serovar typhimurium phage type dt204. *Microbial Drug Resistance*, 8(4):281–289, 2002.

[50] Yu EW, Aires JR, and Nikaido H. Acrb multidrug efflux pump of escherichia coli: composite substrate-binding cavity of exceptional flexibility generates its extremely wide substrate specificity. *J Bacteriol*, 185(19):5657–5664, 2003.

[51] Webber MA, Talukder A, and Piddock LJV. Contribution of mutation at amino acid 45 of acrr to acrb expression and ciprofloxacin resistance in clinical and veterinary escherichia coli isolates. *Antimicrob Agents Chemother*, 49(10):4390–4392, 2005.

[52] Kobayashi N, Tamura N, van Veen HW, Yamaguchi A, and Murakami S. β-lactam selectivity of multidrug transporters acrb and acrd resides in the proximal binding pocket. *J Biol Chem*, 289(15):10680–10690, 2014.

[53] Zgurskaya HI and Nikaido H. Bypassing the periplasm: reconstitution of the AcrAB multidrug efflux pump of Escherichia coli. *Proc Natl Acad Sci USA*, 96:7190–7195, 1999.

[54] Tikhonova EB and Zgurskaya HI. Acra, acrb, and tolc of escherichia coli form a stable intermembrane multidrug efflux complex. *J Biol Chem*, 279:32116–32124, 2004.

[55] Symmons MF, Bokma E, Koronakis E, Hughes C, and Koronakis V. The assembled structure of a complete tripartite bacterial multidrug efflux pump. *Proc Natl Acad Sci USA*, 106:7173–7178, 2009.

[56] Lomovskaya O and Zgurskaya HI. Efflux pumps from gram-negative bacteria:from structure and function to inhibition. *Emerging Trends in Antibacterial Discovery. Part II*, pages 77–105, 2011.

[57] Tseng TT, Gratwick KS, Kollman J, Park D, Nies DH, Goffeau A, and Saier MH. The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J Mol Microbiol Biotechnol*, 1:107–125, 1999.

[58] Murakami S, Nakashima R, Yamashita E, Matsumoto T, and Yamaguchi A. Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature*, 443:173–179, 2006.

[59] Seeger MA, Schiefner A, Eicher T, Verrey F, Diederichs K, and Pos KM. Structural asymmetry of acrb trimer suggests a peristaltic pump mechanism. *Science*, 313(5791):1295–1298, 2006.

[60] Pos KM. Trinity revealed: Stoichiometric complex assembly of a bacterial multidrug efflux pump. *Proc Natl Acad Sci USA*, 106(23):6893–6894, 2009.

[61] Poole K, Krebes K, McNally C, and Neshat S. Multiple antibiotic resistance in pseudomonas aeruginosa: evidence for involvement of an efflux operon. *J Bacteriol*, 175:7363–7372, 1993.

[62] Okusu H and Nikaido H. Acrab efflux pump plays a major role in the antibiotic resistance phenotype of escherichia coli multiple-antibiotic-resistance (mar) mutants. *J Bacteriol*, 178:306–308, 1996.

[63] Murakami S. Multidrug efflux transporter, AcrB-the pumping mechanism. 18:459–465, 2008.

[64] Olliver A, Vallé M, Chaslus-Dancla E, and Cloeckaert A. Role of an acrr mutation in multidrug resistance of in vitro-selected fluoroquinolone-resistant mutants of salmonella enterica serovar typhimurium. *FEMS microbiology letters*, 238(1):267–272, 2004.

[65] van der Straaten T, Janssen R, Mevius DJ, and van Dissel JT. Salmonella gene rma (rama) and multiple-drug-resistant salmonella enterica serovar typhimurium. *Antimicrob Agents Chemother*, 48(6):2292–2294, 2004.

[66] Kaatz GW, Thyagarajan RV, and Seo SM. Effect of promoter region mutations and mgra overexpression on transcription of nora, which encodes a staphylococcus aureus multidrug efflux transporter. *Antimicrob Agents Chemother*, 49(1):161–169, 2005.

[67] Warner DM, Shafer WM, and Jerse AE. Clinically relevant mutations that cause derepression of the neisseria gonorrhoeae mtrc-mtrd-mtre efflux pump system confer different levels of antimicrobial resistance and in vivo fitness. *Molecular microbiology*, 70(2):462–478, 2008.

[68] Blair JMA, Webber MA, Baylay AJ, Ogbolu DO, and Piddock LJV. Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*, 13(1):42–51, 2015.

[69] Mazzariol A, Cornaglia G, and Nikaido H. Contributions of the ampc β-lactamase and the acrab multidrug efflux system in intrinsic resistance of escherichia coli k-12 to β-lactams. *Antimicrob Agents Chemother*, 44(5):1387–1390, 2000.

[70] Elkins CA and Nikaido H. 3d structure of acrb: the archetypal multidrug efflux transporter of escherichia coli likely captures substrates from periplasm. *Drug resistance updates*, 6(1):9–13, 2003.

[71] Middlemiss JK and Poole K. Differential impact of mexb mutations on substrate selectivity of the mexab-oprm multidrug efflux pump of pseudomonas aeruginosa. *J Bacteriol*, 186(5):1258–1269, 2004.

[72] Schulz R, Vargiu AV, Collu F, Kleinekathöfer U, and Ruggerone P. Functional rotation of the transporter AcrB: insights into drug extrusion from simulations. *PLoS Comput Biol*, 6(6):e1000806, June 2010.

[73] Vargiu AV, Collu F, Schulz R, Pos KM, Zacharias M, Kleinekathöfer U, and Ruggerone P. Effect of the F610A mutation on substrate extrusion in the AcrB transporter: explanation and rationale by molecular dynamics simulations. *J Am Chem Soc*, 133(28):10704–7, 2011.

[74] Collu F, Vargiu AV, Dreier J, Cascella M, and Ruggerone P. Recognition of imipenem and meropenem by the RND-transporter MexB studied by computer simulations. *J Am Chem Soc*, 134(46):19146–19158, 2012.

[75] Vargiu AV and Nikaido H. Multidrug binding properties of the acrb efflux pump characterized by molecular dynamics simulations. *Proc Natl Acad Sci USA*, 109(50):20637–20642, 2012.

[76] Kinana AD, Vargiu AV, and Nikaido H. Some ligands enhance the efflux of other ligands by the escherichia coli multidrug pump acrb. *Biochemistry*, 52(46):8342–8351, 2013.

[77] Schulz R, Vargiu AV, Collu F, Kleinekathöfer U, and Ruggerone P. Functional rotation of the transporter acrb: insights into drug extrusion from simulations. *PLoS Comput Biol*, 6(6), 2010.

[78] Schulz R, Vargiu AV, Ruggerone P, and Kleinekathöfer U. Role of water during the extrusion of substrates by the efflux transporter AcrB. *J Phys Chem B*, 115:8278–8287, 2011.

[79] Feng Z, Hou T, and Li Y. Unidirectional peristaltic movement in multisite drug binding pockets of acrb from molecular dynamics simulations. *Molecular BioSystems*, 8(10):2699–2709, 2012.

[80] Fischer N and Kandt C. Porter domain opening and closing motions in the multi-drug efflux transporter acrb. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1828(2):632–641, 2013.

[81] Fischer N, Raunest M, Schmidt TH, Koch DC, and Kandt C. Efflux pump-mediated antibiotics resistance: insights from computational structural biology. *Interdisciplinary sciences, computational life sciences*, 6(1):1–12, March 2014.

[82] MH Saier, R Tam, A Reizer, and J Reizer. Two novel families of bacterial membrane proteins concerned with nodulation, cell division and transport. *Molecular microbiology*, 11(5):841–847, 1994.

[83] Dinh T, Paulsen IT, and Saier Jr MH. A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of gram-negative bacteria. *J Bacteriol*, 176:3825–3831, 1994.

[84] Ma D, Cook DN, Alberti M, Pon NG, Nikaido H, and Hearst JE. Genes acra and acrb encode a stress-induced efflux system of escherichia coli. *Molecular microbiology*, 16(1):45–55, 1995.

[85] Zgurskaya HI and Nikaido H. Acra is a highly asymmetric protein capable of spanning the periplasm. *J Mol Biol*, 285(1):409–420, 1999.

[86] Xu Y, Lee M, Moeller A, Song S, Yoon BY, Kim HM, Jun SY, Lee K, and Ha NC. Funnel-like hexameric assembly of the periplasmic adapter protein in the tripartite multidrug efflux pump in gram-negative bacteria. *J Biol Chem*, 286:17910–17920, 2011.

[87] Xu Y, Moeller A, Jun SY, Le M, Yoon BY, Kim JS, Lee K, and Ha NC. Assembly and channel opening of outer membrane protein in tripartite drug efflux pumps of gram-negative bacteria. *J Biol Chem*, 287:11740–11750, 2012.

[88] Du D, Wang Z, James NR, Voss JE, Klimont E, Ohene-Agyei T, Venter H, Chiu W, and Luisi BF. Structure of the AcrAB-TolC multidrug efflux pump. *Nature*, 509:512–515, 2014.

[89] Jin-Sik Kim, Hyeongseop Jeong, Saemee Song, Hye-Yeon Kim, Kangseok Lee, Jaekyung Hyun, , and Nam-Chul Ha. Structure of the tripartite multidrug

efflux pump acrab-tolc suggests an alternative assembly mode. *Mol. Cells*, 38(2):180–186, 2015.

[90] Mikolosko J, Bobyk K, Zgurskaya HI, and Ghosh P. Conformational flexibility in the multidrug efflux system protein AcrA. *Structure*, 14:577–587, 2006.

[91] Ge Q, Yamada Y, and Zgurskaya H. The c-terminal domain of acra is essential for the assembly and function of the multidrug efflux pump acrab-tolc. *J Bacteriol*, 191(13):4365–4371, 2009.

[92] Nehme D, Li XZ, Elliot R, and Poole K. Assembly of the MexAB-OprM Multidrug Efflux System of Pseudomonas aeruginosa: Identification and Characterization of Mutations in mexA Compromising MexA Multimerization and Interaction with MexB. *J Bacteriol*, 186(10):2973–2983, 2004.

[93] Stegmeier JF, Polleichtner G, Brandes N, Hotz C, and Andersen C. Importance of the adaptor (membrane fusion) protein hairpin domain for the functionality of multidrug efflux pumps. *Biochemistry*, 45(34):10303–10312, 2006.

[94] Lobedanz S, Bokma E, Symmons MF, Koronakis E, Hughes C, and Koronakis V. A periplasmic coiled-coil interface underlying tolc recruitment and the assembly of bacterial drug efflux pumps. *Proc Natl Acad Sci USA*, 104(11):4612–4617, 2007.

[95] Lee B and Richards F. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.*, 55:379–400, 1971.

[96] Lomovskaya O, Lewis K, and Matin A. Emrr is a negative regulator of the escherichia coli multidrug resistance pump emrab. *J Bacteriol*, 177(9):2328–2334, 1995.

[97] Buchanan SK. Type i secretion and multidrug efflux: transport through the tolc channel-tunnel. *Trends in biochemical sciences*, 26(1):3–6, 2001.

[98] Koronakis V, Sharff A, Koronakis E, Luisi B, and Hughes C. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature*, 405:914–919, 2000.

[99] Helen I. Zgurskaya, Jon W. Weeks, Abigail T. Ntreh, Logan M. Nickels, and David Wolloscheck. Mechanism of coupling drug transport reactions located in two different membranes. *Frontiers in Microbiology*, 6(100), 2015.

[100] Andersen C, Hughes C, and Koronakis V. Chunnel vision. Export and efflux through bacterial channel-tunnels. *EMBO reports*, 1(4):313–8, October 2000.

[101] Andersen C, Hughes C, and Koronakis V. Electrophysiological behavior of the tolc channel-tunnel in planar lipid bilayers. *J Membr Biol*, 185(1):83–92, 2002.

[102] Andersen C, Koronakis E, Hughes C, and Koronakis V. An aspartate ring at the tolc tunnel entrance determines ion selectivity and presents a target for blocking by large cations. *Molecular microbiology*, 44(5):1131–1139, 2002.

[103] Lobedanz S, Bokma E, Symmons MF, Koronakis E, Hughes C, and Koronakis V. A periplasmic coiled-coil interface underlying tolc recruitment and the assembly of bacterial drug efflux pumps. *Proc. Natl. Acad. Sci. U.S.A.*, 104:4612–4617, 2007.

[104] Bavro VN, Pietras Z, Furnham N, Pérez-Cano L, J Fernandez-Recio, Pei XY, Misra R, and Luisi B. Assembly and Channel Opening in a Bacterial Drug Efflux Machine. *Molecular Cell*, 30:114–121, 2008.

[105] Raunest M and Kandt C. Locked on one side only: ground state dynamics of the outer membrane efflux duct tolc. *Biochemistry*, 51(8):1719–1729, 2012.

[106] Pei XY, Hinchliffe P, Symmons MF, Koronakis E, Benz R, Hughes C, and Koronakis V. Structures of sequential open states in a symmetrical opening transition of the tolc exit duct. *Proc Natl Acad Sci USA*, 108(5):2112–2117, 2011.

[107] Andersen C, Koronakis E, Bokma E, Eswaran J, Humphreys D, Hughes C, and Koronakis V. Transition to the open state of the tolc periplasmic tunnel entrance. *Proc Natl Acad Sci USA*, 99(17):11103–11108, 2002.

[108] Schulz R and Kleinekathöfer U. Transitions between closed and open conformations of TolC: The effects of ions in simulations. *Biophysical Journal*, 96:3116–3125, 2009.

[109] Weeks JW, Bavro VN, and Misra R. Genetic assessment of the role of AcrB hairpins in the assembly of the TolC-AcrAB multidrug efflux pump of Escherichia coli. *Molecular Microbiology*, 91:965–975, 2014.

[110] Murakami S, Nakashima R, Yamashita E, and Yamaguchi A. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature*, 419:587–593, 2002.

[111] Bokma E, Koronakis E, Lobedanz S, Hughes C, and Koronakis V. Directed evolution of a bacterial efflux pump: adaptation of the e. coli tolc exit duct to the pseudomonas mexab translocase. *FEBS letters*, 580(22):5339–5343, 2006.

[112] Zgurskaya HI and Nikaido H. Cross-linked complex between oligomeric periplasmic lipoprotein acra and the inner-membrane-associated multidrug efflux pump acrb from escherichia coli. *J Bacteriol*, 182(15):4264–4267, 2000.

[113] Elkins CA and Nikaido H. Chimeric analysis of acra function reveals the importance of its c-terminal domain in its interaction with the acrb multidrug efflux pump. *J Bacteriol*, 185(18):5349–5356, 2003.

[114] Janganan TK, Bavro VN, Zhang L, Matak-Vinkovic D, Barrera NP, Venien-Bryan C, Robinson CV, Borges-Walmsley MI, and Walmsley AR. Evidence for the assembly of a bacterial tripartite multidrug pump with a stoichiometry of 3:6:3. *J Biol Chem*, 286:26900–26912, 2011.

[115] Hinchliffe P, Symmons MF, Hughes C, and Koronakis V. Structure and operation of bacterial tripartite pumps. *Annual review of microbiology*, 67:221–242, 2013.

[116] Su CC, Long F, Zimmermann MT, Rajashankar KR, Jernigan RL, and Yu EW. Crystal structure of the CusBA heavy-metal efflux complex of Escherichia coli. *Nature*, 470:558–562, 2011.

[117] Srinivasan A, Wolfenden LL, Song X, Mackie K, Hartsell TL, Jones HD, Diette GB, Orens JB, Yung RC, Ross TL, et al. An outbreak of pseudomonas aeruginosa infections associated with flexible bronchoscopes. *N Engl J Med*, 348(3):221–227, 2003.

[118] Pagani L, Colinon C, Migliavacca R, Labonia M, Docquier JD, Nucleo E, Spalla M, Bergoli ML, and Rossolini GM. Nosocomial outbreak caused by multidrug-resistant pseudomonas aeruginosa producing imp-13 metallo-β-lactamase. *J Clin Microbiol*, 43(8):3824–3828, 2005.

[119] Hidron AI, Edwards JR, Patel J, Horan TC, Sievert DM, Pollock DA, and Fridkin SK. Antimicrobial-resistant pathogens associated with healthcare-

associated infections: annual summary of data reported to the national health-care safety network at the centers for disease control and prevention, 2006–2007. *Infection Control*, 29(11):996–1011, 2008.

[120] Kerr KG and Snelling AM. Pseudomonas aeruginosa: a formidable and ever-present adversary. *J Hosp Infect*, 73(4):338–344, 2009.

[121] Mahar P, Padiglione AA, Cleland H, Paul E, Hinrichs M, and Wasiak J. Pseudomonas aeruginosa bacteraemia in burns patients: risk factors and outcomes. *Burns*, 36(8):1228–1233, 2010.

[122] Quinn JP, Dudek EJ, DiVincenzo CA, Lucks DA, and Lerner SA. Emergence of resistance to imipenem during therapy for pseudomonas aeruginosa infections. *J Infect Dis*, 154(2):289–294, 1986.

[123] Döring G, Conway SP, Heijerman HGM, Hodson ME, Høiby N, Smyth A, Touw DJ, et al. Antibiotic therapy against pseudomonas aeruginosa in cystic fibrosis: a european consensus. *European Respiratory Journal*, 16(4):749–767, 2000.

[124] Rossolini GM and Mantengoli E. Treatment and control of severe infections caused by multiresistant pseudomonas aeruginosa. *Clinical Microbiology and infection*, 11(s4):17–32, 2005.

[125] Paterson DL. The epidemiological profile of infections with multidrug-resistant pseudomonas aeruginosa and acinetobacter species. *Clinical Infectious Diseases*, 43(Supplement 2):S43–S48, 2006.

[126] Jones RN, Stilwell MG, Rhomberg PR, and Sader HS. Antipseudomonal activity of piperacillin/tazobactam: more than a decade of experience from the

sentry antimicrobial surveillance program (1997–2007). *Diagnostic microbiology and infectious disease*, 65(3):331–334, 2009.

[127] Zilberberg MD, Chen J, Mody SH, Ramsey AM, and Shorr AF. Imipenem resistance of pseudomonas in pneumonia: a systematic literature review. *BMC pulmonary medicine*, 10(1):45, 2010.

[128] Nikaido H. Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science*, 264(5157):382–388, 1994.

[129] Nikaido H and Hancock R. Outer membrane permeability of pseudomonas aeruginosa. *The bacteria*, 10:145–93, 2012.

[130] Gutiérrez O, Juan C, Cercenado E, Navarro F, Bouza E, Coll P, Pérez JL, and Oliver A. Molecular epidemiology and mechanisms of carbapenem resistance in pseudomonas aeruginosa isolates from spanish hospitals. *Antimicrob Agents Chemother*, 51(12):4329–4335, 2007.

[131] Rodríguez-Martínez JM, Poirel L, and Nordmann P. Molecular epidemiology and mechanisms of carbapenem resistance in pseudomonas aeruginosa. *Antimicrob Agents Chemother*, 53(11):4783–4788, 2009.

[132] Wang J, Zhou JY, Qu TT, Shen P, Wei ZQ, Yu YS, and Li LJ. Molecular epidemiology and mechanisms of carbapenem resistance in pseudomonas aeruginosa isolates from chinese hospitals. *Int J Antimicrob Agents*, 35(5):486–491, 2010.

[133] Higgins PG, Fluit AC, Milatovic D, Verhoef J, and Schmitz FJ. Mutations in gyra, parc, mexr and nfxb in clinical isolates of pseudomonas aeruginosa. *Int J Antimicrob Agents*, 21(5):409–413, 2003.

[134] Jacoby GA. Mechanisms of resistance to quinolones. *Clinical Infectious Diseases*, 41(Supplement 2):S120–S126, 2005.

[135] Lee JK, Lee YS, Park YK, and Kim BS. Alterations in the gyra and gyrb subunits of topoisomerase ii and the parc and pare subunits of topoisomerase iv in ciprofloxacin-resistant clinical isolates of pseudomonas aeruginosa. *Int J Antimicrob Agents*, 25(4):290–295, 2005.

[136] Drlica K, Hiasa H, Kerns R, Malik M, Mustaev A, and Zhao X. Quinolones: action and resistance updated. *Curr Top Med Chem*, 9(11):981, 2009.

[137] Poole K. Aminoglycoside resistance in pseudomonas aeruginosa. *Antimicrobial agents and Chemotherapy*, 49(2):479–487, 2005.

[138] Poole K. Efflux pumps as antimicrobial resistance mechanisms. *Annals of medicine*, 39(3):162–176, 2007.

[139] Akama H, Kanemaki M, Yoshimura M, Tsukihara T, Kashiwagi T, Yoneyama H, Narita SI, Nakagawa A, and Nakae T. Crystal structure of the drug discharge outer membrane protein, oprm, of pseudomonas aeruginosa dual modes of membrane anchoring and occluded cavity end. *J Biol Chem*, 279(51):52816–52819, 2004.

[140] Sennhauser G, Amstutz P, Briand C, Storchenegger O, and Grütter MG. Drug export pathway of multidrug exporter acrb revealed by darpin inhibitors. *PLoS Biol*, 5(1):e7, 2007.

[141] Sennhauser G, Bukowska MA, Briand C, and Grütter MG. Crystal structure of the multidrug exporter mexb from pseudomonas aeruginosa. *J Mol Biol*, 389(1):134–145, 2009.

[142] Tikhonova EB, Wang Q, and Zgurskaya HI. Chimeric analysis of the multicomponent multidrug efflux transporters from gram-negative bacteria. *J Bacteriol*, 184:6499–6507, 2002.

[143] Krishnamoorthy G, Tikhonova EB, and Zgurskaya HI. Fitting periplasmic membrane fusion proteins to inner membrane transporters: mutations that enable escherichia coli acra to function with pseudomonas aeruginosa mexb. *J Bacteriol*, 190(2):691–698, 2008.

[144] Akama H, Matsuura T, Kashiwagi S, Yoneyama H, Narita SI, Tsukihara T, Nakagawa A, and Nakae T. Crystal structure of the membrane fusion protein, MexA, of the multidrug transporter in Pseudomonas aeruginosa. *J Biol Chem*, 279:25939–25942, 2004.

[145] Wang B, Weng J, Fan K, and Wang W. Interdomain flexibility and pH-induced conformational changes of AcrA revealed by molecular dynamics simulations. *J Phys Chem B*, 116:3411–3420, 2012.

[146] Nehme D and Poole K. Interaction of the MexA and MexB components of the MexAB-OprM multidrug efflux system of Pseudomonas aeruginosa: Identification of MexA extragenic suppressors of a T578I mutation in MexB. *Antimicrobial Agents and Chemotherapy*, 49:4375–4378, 2005.

[147] Nehme D and Poole K. Assembly of the MexAB-OprM multidrug pump of Pseudomonas aeruginosa: Component interactions defined by the study of pump mutant suppressors. *J Bacteriol*, 189:6118–6127, 2007.

[148] Vaccaro L, Koronakis V, and Sansom MSP. Flexibility in a drug transport accessory protein: molecular dynamics simulations of MexA. *Biophysical journal*, 91:558–564, 2006.

[149] Gunsteren WFV, Luque FJ, Timms D, and Torda AE. Molecular mechanics in biology: From structure to function, taking account of solvation. *Annual Review of Biophysics and Biomolecular Structure*, 23(1):847–863, 1994.

[150] Anderson JA, Lorenz CD, and Travesset A. General purpose molecular dynamics simulations fully implemented on graphics processing units. *J Comput Phys*, 227(1):5342–5359, 2008.

[151] Giupponi G, Harvey MJ, and De Fabritiis G. The impact of accelerator processors for high-throughput molecular modeling and simulation. *Drug discovery today*, 13(23):1052–1058

[152] Friedrichs MS, Eastman P, Vaidyanathan V, Houston M, Legrand S, Beberg AL, Ensign DL, Bruns CM, and Pande VS. Accelerating molecular dynamic simulation on graphics processing units. *J Comput Chem*, 30(6):864–872, 2009.

[153] Harvey MJ and De Fabritiis G. An implementation of the smooth particle mesh Ewald method on GPU hardware. *J Chem Theory Comput*, 5:2371–2377, 2009.

[154] Harvey MJ, Giupponi G, and e Fabritiis G. ACEMD: Accelerating biomolecular dynamics in the microsecond time scale. *J Chem Theory Comput*, 5:1632–1639, 2009.

[155] Goetz AW, Wolfle TM, and Walker RC. Quantum chemistry on graphics processing units. *Annu Rep Comp Chem*, 6:21–35, 2010.

[156] Eastman P and Pande VS. Efficient nonbonded interactions for molecular dynamics on a graphics processing unit. *J Comput Chem*, 31(6):1268–1272, 2010.

[157] Goetz AW, Williamson MJ, Xu D, Poole D, Le Grand S, and Walker RC. Routine microsecond molecular dynamics simulations with amber on gpus. 1. generalized born. *J Chem Theory Comput*, 8(5):1542–1555, 2012.

[158] Dror RO, Dirks RM, Grossman JP, Xu H, and Shaw DE. Biomolecular simulation: A computational microscope for molecular biology. *Annual Review of Biophysics*, 41(1):429–452, 2012.

[159] Le Grand S, Goetz AW, and Walker RC. Spfp: Speed without compromise-a mixed precision model for gpu accelerated molecular dynamics simulations. *Computer Physics Communications*, 184(2):374–380, 2013.

[160] Lane TJ, Shukla D, Beauchamp KA, and Pande VS. To milliseconds and beyond: challenges in the simulation of protein folding. *Curr Opin Struct Biol*, 23(1):58–65, 2013.

[161] Alder BJ and Wainwright TE. Phase transition for a hard sphere system. *J Chem Phys*, 27:59–61, 1957.

[162] Adler BJ and Wainwright TE. Studies in molecular dynamics. i. general method. *J Chem Phys*, 31:459–466, 1959.

[163] Rahman A. Correlations in the motion of atoms in liquid argon. *Physical Review*, 136(2A):A405, 1964.

[164] Stirling A, Iannuzzi M, Laio A, and Parrinello M. Azulene-to-naphthalene rearrangement: The car-parrinello metadynamics method explores various reaction mechanisms. *Chem. Phys. Chem.*, 5:1558–1568, 2004.

[165] McCammon JA. Dynamics of folded proteins. *Nature*, 267:16, 1977.

[166] Berendsen HJC, van der Spoel D, and van Drunen R. Gromacs: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.*, 91:43–56, 1995.

[167] Case, DA, Cheatam III TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, and Woods RJ. The amber biomolecular simulation programs. *J Comput Chem*, 26(16):1668–1688, 2005.

[168] Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, and Schulten K. Scalable molecular dynamics with namd. *J Comput Chem*, 26(16):1781–1802, 2005.

[169] Brooks BR, Brooks CL, MacKerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, et al. Charmm: the biomolecular simulation program. *J Comput Chem*, 30(10):1545–1614, 2009.

[170] Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, and Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *J Chem Theory Comput*, 9:3878–3888, 2013.

[171] Karplus M and McCammon JA. Molecular dynamics simulations of biomolecules. *Nature Structural & Molecular Biology*, 9(9):646–652, 2002.

[172] Kandt C, Ash WL, and Tieleman DP. Setting up and running molecular dynamics simulations of membrane proteins. *Methods*, 41(4):475–488, 2007.

[173] Grossfield A, Feller SE, and Pitman MC. Convergence of molecular dynamics simulations of membrane proteins. *Proteins: structure, function, and bioinformatics*, 67(1):31–40, 2007.

[174] Frenkel D and Smit B. *Understanding Molecular Simulations: From Algorithms to Applications.* Academic Press, San Diego, CA, 2 edition, 2002.

[175] Lebowitz JL, McDonald IR, and Verlet L. Ensemble dependence of fluctuations with application to machine computation. *Phys. Rev.*, 153:250–254, 1967.

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

[177] Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, and Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys*, 81(8):3684–3690, 1984.

[178] Nosé S. A unified formulation of the constant temperature molecular dynamics method. *J Chem Phys*, 81:1511–519, 1984.

[179] Hoover WG. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, 1985.

[180] Andersen HC. Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys*, 72(4):2384–2393, 1980.

[181] Verlet L. Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical Review*, 159(1):98–103, 07 1967.

[182] Hockney R. The potential calculation and some applications. *Methods in Computational Physics*, 9:135–211, 1970.

[183] Ryckaert JP, Ciccotti Giovanni, and Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with contraints: Molecular dynamics of n-alkanes. *J Comput Phys*, 23:327–341, 1977.

[184] Tuckerman ME and Martyna GJ. Understanding modern molecular dynamics: Techniques and applications. *J. Phys. Chem. B*, 104(2):159–178, 2000.

[185] Tuckerman ME, Liu Y, Ciccotti G, and Martyna GJ. Non-hamiltonian molecular dynamics: Generalizing hamiltonian phase space principles to non-hamiltonian systems. *J. Chem. Phys.*, 115(4):1678–1702, 2001.

[186] Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, and Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc*, 117(19):5179–5197, 1995.

[187] Pearlman DA, Case DA, Caldwell JW, Ross WR, Cheatham III TE, DeBolt SE, Ferguson D, Seibel G, and Kollman P. Amber, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculationsto simulate the structural and energetic properties of molecules. *Comp. Phys. Comm.*, 91:1–41, 1995.

[188] Case DA, Babin V, Berryman JT, Betz RM, Cai Q, Cerutti DS, Cheatham TE III, Darden TA, Duke RE, Gohlke H, Goetz AW, Gusarov S, Homeyer N, Janowski P, Kaus J, Kolossváry I, Kovalenko A, Lee TS, LeGrand S, Luchko T, Luo R ad Madej B, Merz KM, Paesani F, Roe DR, Roitberg A, Sagui C,

Salomon-Ferrer R, Seabra G, Simmerling CL, Smith W, Swails J, Walker RC, Wang J, Wolf RM, Wu X, and Kollman PA. Amber14. 2014.

[189] Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL, and Woods RJ. Glycam06: a generalizable biomolecular force field. carbohydrates. *J Comput Chem*, 29(4):622–655, 2008.

[190] CJ Dickson, Madej BD, ÅA Skjevik, Betz RM, Teigen K, Gould IR, and Walker RC. Lipid14: the amber lipid force field. *J Chem Theory Comput*, 10(2):865–879, 2014.

[191] Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, and Karplus M. Charmm - a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.*, 4(2):187–217, 1983.

[192] Clark M, RD Cramer, and Vanopdenbosch N. Validation of the general-purpose tripos 5.2 force-field. *J. Comp. Chem.*, 10(8):982–1012, 1989.

[193] Van Gunsteren WF and Berendsen HJC. Groningen molecular simulation (gromos). *Library manual, Biomos, Groningen, The Netherlands*, pages 1–221, 1987.

[194] Christen M, Hünenberger PH, Bakowies D, Baron R, Bürgi R, Geerke DP, Heinz TN, Kastenholz MA, Kräutler V, Oostenbrink C, et al. The gromos software for biomolecular simulation: Gromos05. *J Comput Chem*, 26(16):1719–1751, 2005.

[195] Soares T A, Hünenberger PH, Kastenholz MA, Kräutler V, Lenz T, Lins RD, Oostenbrink C, and van Gunsteren WF. An improved nucleic acid parameter set for the gromos force field. *J Comput Chem*, 26(7):725–737, 2005.

[196] Marrink SJ and Tieleman DP. Perspective on the martini model. *Chemical Society Reviews*, 2013.

[197] Hansson T, Oostenbrink C, and van Gunsteren W. Molecular dynamics simulations. *Curr Opin Struct Biol*, 12(2):190–196, 4 2002.

[198] Steinbach PJ and Brooks BR. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J Comput Chem*, 15(7):667–683, 1994.

[199] Ewald PP. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der Physik*, 369(3):253–287, 1921.

[200] Eastwood JW and Hockney RW. Shaping the force law in two-dimensional particle-mesh models. *J Comput Phys*, 16(4):342–359, 1974.

[201] Hockney RW and Eastwood JW. *Computer Simulations Using Particles*. McGraw-Hill, New York, 1981.

[202] Cheatham III TE, Miller JL, Fox T, Darden TA, and Kollman PA. Molecular dynamics simulations on solvated biomolecular systems: the particle mesh ewald method leads to stable trajectories of dna, rna, and proteins. *J Am Chem Soc*, 117(14):4193–4194, 1995.

[203] Pollock E and Glosli J. Comments on p 3 m, fmm, and the ewald method for large periodic coulombic systems. *Computer Physics Communications*, 95(2):93–110, 1996.

[204] Deserno M and Holm C. How to mesh up ewald sums. i. a theoretical and numerical comparison of various particle mesh routines. *J Chem Phys*, 109(18):7678–7693, 1998.

[205] Hegyi H and Gerstein M. The relationship between protein structure and function: a comprehensive survey with application to the yeast genome. *J Mol Biol*, 288(1):147–164, 1999.

[206] Hvidsten TR, Laegrid A, Kryshtafovych A, Andersson G, Fidelis K, and Komorowski J. A comprehensive analysis of the structure-function relationship in proteins based on local structure similarity. *PLoS ONE*, 4(7), 2009.

[207] Engh RA and Huber R. Accurate bond and angle parameters for x-ray protein structure refinement. *Acta Crystallographica Section A: Foundations of Crystallography*, 47(4):392–400, 1991.

[208] Liapunov AM. *Stability of motion*. Academic Press, New York, 1966.

[209] Garman DF. Developments in x-ray crystallographic structure determination of biological macromolecules. *Science*, 314(6175):1102–1108, 2014.

[210] Drenth J. *Principles of protein X-ray crystallography*. Springer Science & Business Media, 2007.

[211] van Ingen H and Bonvin AMJJ. Information-driven modeling of large macromolecular assemblies using nmr data. *J Magn Reson*, 241:103–114, 2014.

[212] Harris SA, Gavathiotis E, Searle MS, Orozco M, and Laughton CA. Cooperativity in drug-dna recognition: a molecular dynamics study. *J. Am. Chem. Soc.*, 123:12658–12663, 2001.

[213] Jackman LM and Sternhell S. *Application of Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry: International Series in Organic Chemistry*. Elsevier, 2013.

[214] Orlova EV and Saibil HR. Structural analysis of macromolecular assemblies by electron microscopy. *Chemical Reviews*, 111(12):7710–7748, 12 2011.

[215] Goldstein JI, Newbury DE, Echlin P, Joy DC, Fiori C, Lifshin E, et al. *Scanning electron microscopy and X-ray microanalysis. A text for biologists, materials scientists, and geologists.* Plenum Publishing Corporation, 1981.

[216] Hayat MA et al. *Principles and techniques of scanning electron microscopy. Biological applications. Volume 1.* Van Nostrand Reinhold Company., 1974.

[217] Gordon R, Bender R, and Herman GT. Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography. *J Theor Biol*, 29(3):471–481, 1970.

[218] Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, I Phan, Pilbout S, and Schneider M. The swiss-prot protein knowledgebase and its supplement trembl in 2003. *Nucleic Acids Research*, 31(1):365–370, 2003.

[219] Westbrook J, Feng Z, Chen L, Yang H, and Berman HM. The protein data bank and structural genomics. *Nucleic Acids Research*, 31(1):489–491, 2003.

[220] Schwede T, Kopp J, Guex N, and Peitsch MC. Swiss-model: an automated protein homology-modeling server. *Nucleic Acids Research*, 31(13):3381–3385, 2003.

[221] Sali A and Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*, 234:779–815, 1993.

[222] Bower MJ, Cohen FE, and Dunbrack RL Jr. Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool1. *J Mol Biol*, 267(5):1268–1282, 4 1997.

[223] Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo Fr, and Šali A. Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure*, 29(1):291–325, 2000.

[224] Biasini M, Biener S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, and Schwede T. Swiss-model: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acid Research*, 42:252–258, 2014.

[225] Sali A, Potterton L, Yuan F, van Vlijmen H, and Karplus M. Evaluation of comparative protein modeling by modeller. *Proteins: Structure, Function, and Bioinformatics*, 23(3):318–326, 1995.

[226] Eswar N, Webb B, Marti-Renom M A, Madhusudhan MS, Eramian D, Shen M, Pieper U, and Sali A. Comparative protein structure modeling using modeller. *Curr protocols in bioinformatics*, 5, 10 2006.

[227] Rost B. Twilight zone of protein sequence alignments. *Protein Engineering*, 12(2):85–94, 1999.

[228] Fiser A and Sali A. Modeller: generation and refinement of homology-based protein structure models. *Methods Enzymol*, 374:461–91, 2003.

[229] Sali A and Overington JP. Derivation of rules for comparative protein modeling from a database of protein structure alignments. *Protein Science*, 3(9):1582–1596, 1994.

[230] MacKerell AD, Bashford D, Bellott MLDR, Dunbrack RL, Enseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B*, 102(18):3586–3616, 1998.

[231] Shen MY and Sali A. Statistical potential for assessment and prediction of protein structures. *Protein Sci*, 15:2507–2524, 2006.

[232] Alber F, Dokudovskaya S, Veenhoff LM, Wenzhu Z, Klipper J, Devos D, Suprapto A, Karni-Schmidt O, Williams R, Chait BT, Rout MP, and Sali A. Determining the architectures of macromolecular assemblies. *Nature*, 450:683–694, 2007.

[233] Alber F, Forster F, Korkin D, Topf M, and Sali A. Integrating diverse data for structure determination of macromolecular assemblies. *Annu Rev Biochem*, 77:443–477, 2008.

[234] Suchanova B and Tuma R. Folding and assembly of large macromolecular complexes monitored by hydrogen-deuterium exchange and mass spectrometry. *Microbial Cell Factories*, 7(12), 2008.

[235] Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, and Taylor RD. Improved protein–ligand docking using gold. *Proteins: Structure, Function, and Bioinformatics*, 52(4):609–623, 2003.

[236] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, and Olson AJ. Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *J. Comp. Chem.*, 30(16):2785–2791, 2009.

[237] Lang PT, Brozell SR ., Mukherjee S, Pettersen EF, Meng EC, Thomas V, Rizzo RC, Case DA, James TL, and Kuntz ID. Dock 6: Combining techniques to model rna–small molecule complexes. *RNA*, 15(6):1219–1230, 06 2009.

[238] Dominguez C et al. Haddock: a protein–protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.*, 125:1731–1737, 2003.

[239] Grosdidier A et al. Swissdock, a protein-small molecule docking web service based on eadock dss. *Nucleic Acids Res.*, 39:W270–W277, 2011.

[240] Zacharias M. Protein-protein docking with a reduced protein model accounting for side-chain flexibility. *Protein Sci*, 12(6):1271–82, June 2003.

[241] De Vries S and Zacharias M. Flexible docking and refinement with a coarse-grained protein model using ATTRACT. *Proteins*, 81(12):2167–74, December 2013.

[242] Chen YC. Beware of docking. *Trends in Pharmacological Sciences*, 36(2):78–95, 2015.

[243] Zacharias M. Attract: protein-protein docking in capri using a reduced protein model. *Proteins*, 60(2):252–256, 2005.

[244] Zacharias M and May A. Accounting for global protein deformability during protein-protein and protein-ligand docking. *Biochimica et biophysica acta*, 1754:225–231, 2005.

[245] Zacharias M and Fiorucci S. Binding site prediction and improved scoring during flexible protein-protein docking with attract. *Proteins*, 78(15):3131–3139, 2010.

[246] De Vries S and Zacharias M. Attract-em: a new method for the computational assembly of large molecular machines using cryo-em maps. *PLoS ONE*, 7(12):49733, 2012.

[247] Knapp B, Frantal S, Cibena M, Schreiner W, and Bauer P. Is an intuitive convergence definition of molecular dynamics simulations solely based on the root mean square deviation possible? *J Comput Biol*, 18(8):997–1005, 2011.

[248] Lobanov M, Bogatyreva NS, and Galzitskaia OV. Radius of gyration as indicator of compactness of protein structure. *Mol Biol (Mosk)*, 42(4):701–706, 2008.

[249] Karpen ME, Tobias DJ, and Brooks CL. Statistical clustering techniques for the analysis of long molecular dynamics trajectories: analysis of 2.2-ns trajectories of ypgdv. *Biochemistry*, 32(2):412–420, 1993.

[250] Shao J, Tanner SW, Thompson N, and Cheatham III TE. Clustering molecular dynamics trajectories: 1. characterizing the performance of different clustering algorithms. *J Chem Theory Comput*, 3(6):2312–2334, 2007.

[251] Roe DR and Cheatham TE III. Ptraj and cpptraj: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.*, 9(7):3084–3095, 2013.

[252] Humphrey W, Dalke A, and Schulten K. Vmd - visual molecular dynamics. *J. Mol. Graphics*, 14:33–38, 1996.

[253] Piddock LJV, Whale K, and Wise R. Quinolone resistance in salmonella: clinical experience. *The Lancet*, 335(8703):1459, 1990.

[254] Ricci V and Piddock LJV. Ciprofloxacin resistance in clinical isolates of salmonella typhimurium obtained from two patients. *Antimicrob Agents Chemother*, 37(4):662–666, 1993.

[255] Piddock LJV, White DG, Gensberg K, Pumbwe L, and Griggs DJ. Evidence for an efflux pump mediating multiple antibiotic resistance in salmonella enterica serovar typhimurium. *Antimicrob Agents Chemother*, 44(11):3118–3121, 2000.

[256] Piddock LJ Griggs DJ, Gensberg K. Ciprofloxacin resistance in clinical isolates of salmonella typhimurium obtained from two patients. *Antimicrob Agents Chemother*, 44(11):3118–, 1993.

[257] Nakashima R, Sakurai K, Yamasaki S, Nishino K, and Yamaguchi A. Structures of the multidrug exporter acrb reveal a proximal multisite drug-binding pocket. *Nature*, 480(7378):565–569, 2011.

[258] Eicher T et al. Transport of drugs by the multidrug transporter acrb involves an access and a deep binding pocket that are separated by a switch-loop. *Proc Natl Acad Sci USA*, 109(15):5687–5692, 2012.

[259] Yao X-Q, Kimura N, Murakami S, and Takada S. Drug uptake pathways of multidrug transporter acrb studied by molecular simulations and site-directed mutagenesis experiments. *J Am Chem Soc*, 135(20):7474–7485, 2013.

[260] Husain F, Bikhchandani M, and Nikaido H. Vestibules are part of the substrate path in the multidrug efflux transporter acrb of escherichia coli. *J Bacteriol*, 193(20):5847–5849, 2011.

[261] Bohnert JA, Schuster S, Seeger MA, Fahnrich E, Pos KM, and Kern WV. Site-directed mutagenesis reveals putative substrate binding residues in the

escherichia coli rnd efflux pump acrb. *J Am Chem Soc*, 190(24):8225–8229, 2008.

[262] Hornak V, Abel R, Okur A, Strockbin B, Roitberg A, and Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins*, 65(3):712–725, 2006.

[263] Vargiu AV, Ruggerone P, Opperman TJ, Nguyen ST, and Nikaido H. Inhibition of e. coli acrb multidrug efflux pump by mbx2319: Molecular mechanism and comparison with other inhibitors. *Antimicrob Agents Ch*, 58(10):6224–6234, 2014.

[264] Cha H-J, Muller RT, and Pos KM. Switch-loop flexibility affects transport of large drugs by the promiscuous acrb multidrug efflux transporter. *Antimicrob Agents Ch*, 58(8):4767–4772, 2014.

[265] Sterpone F, Ceccarelli M, and Marchi M. Dynamics of hydration in hen egg white lysozyme. *J Mol Biol*, 311(2):409–419, 2001.

[266] Collu F, Ceccarelli M, and Ruggerone P. Exploring binding properties of agonists interacting with a ÎŽ-opioid receptor. *PLoS ONE*, 7(12):e52633, 2012.

[267] et al. Nakashima R. Structural basis for the inhibition of bacterial multidrug exporters. *Nature*, 500:102–106, 2013.

[268] Lawler AJ, Ricci V, Busby SJW, and Piddock LJV. Genetic inactivation of acrab or inhibition of efflux induces expression of rama. *PLoS ONE*, 68(7):1551–1557, 2013.

[269] Li XZ, Nikaido H, and Poole K. Role of mexa-mexb-oprm in antibiotic efflux in pseudomonas aeruginosa. *Antimicrob Agents Chemother*, 39(9):1948–1953, 1995.

[270] Ziha-Zarifi I, Llanes C, Köhler T, Pechere JC, and Plesiat P. In vivo emergence of multidrug-resistant mutants ofpseudomonas aeruginosa overexpressing the active efflux system mexa-mexb-oprm. *Antimicrobial Agents and Chemotherapy*, 43(2):287–291, 1999.

[271] Xian-Zhi Li, Li Zhang, and Keith Poole. Interplay between the mexa-mexb-oprm multidrug efflux system and the outer membrane barrier in the multiple antibiotic resistance of pseudomonas aeruginosa. *J Antimicrob Chemother*, 45(4):433–436, 2000.

[272] Reffay M, Yann G, Benabdelhak H, Phan G, Taulier N, Ducruix A, Hodges RS, and Urbach W. Tracking membrane protein association in model membranes. *PLoS ONE*, 4(4):e5035, 2009.

[273] Trepout S, Taveau JC, Benabdelhak H, Grace ER, Ducruix A, Frangakis AS, and Lambert O. Structure of reconstituted bacterial membrane efflux pump by cryo-electron tomography. *Biochimica et biophysica acta*, 1798:1953–1960, 2010.

[274] Ferrandez Y, Monlezun L, Phan G, Benabdelhak H, Benas P, Ulryck N, Falson P, Ducruix A, Picard M, and Broutin I. Stoichiometry of the MexA-OprM binding, as investigated by blue native gel electrophoresis. *Electrophoresis*, 33(8):1282–7, April 2012.

[275] Greene NP, Hinchliffe P, Crow A, Ababou A, Hughes C, and Koronakis V. Structure of an atypical periplasmic adaptor from a multidrug efflux pump of the spirochete borrelia burgdorferi. *FEBS letters*, 587(18):2984–2988, 2013.

[276] Kobayashi N, Nishino K, and Yamaguchi A. Novel macrolide-specific abc-type efflux transporter inescherichia coli. *J. of Bacteriol*, 183(19):5639–5644, 2001.

[277] Yum S, Xu Y, Piao S, Sim SH, Kim HM, Jo WS, Kim KJ, Kweon HS, Jeong MH, Jeon H, Lee K, and Ha NC. Crystal structure of the periplasmic component of a tripartite macrolide-specific efflux pump. *J Mol Biol*, 387(5):1286–1297, 4 2009.

[278] Long F, Su CC, Zimmermann MT, Boyken SE, Rajashankar KR, Jernigan RL, and Yu EW. Crystal structures of the cusa efflux pump suggest methionine-mediated metal transport. *Nature*, 467(7314):484–488, 09 2010.

[279] Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC andz Zhang W, Merz KM, Roberts B, Hayik S, Roitberg A, Seabra G, Swails J, Goetz AW, Kolossváry I, Wong KF, Paesani F, Vanicek J, Wolf RM, Liu J, Wu X, SR Brozell, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh MJ, Cui G, Roe DR, Mathews DH, Seetin MG, Salomon-Ferrer R, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, and Kollman PA. Amber12. 2012.

[280] Porollo A, Adamczak R, and Meller J. Polyview: A flexible visualization tool for structural and functional annotations of proteins. *Bioinformatics*, 20:2460–2462, 2004.

[281] Sokal RR and Michener CD. A statistical method for evaluating systematic relationships. *Univ Kans Sci Bull*, 38:1409–1438, 1958.

[282] Bavro VN, Pietras Z, Furnham N, Pérez-Cano L, Fernandez-Recio J, Pei XY, Misra R, and Luisi B. Assembly and Channel Opening in a Bacterial Drug Efflux Machine. *Molecular Cell*, 30:114–121, 2008.

[283] Janganan TK, Bavro VN, Zhang L, Matak-Vinkovic D, Barrera NP, Venien-Bryan C, Robinson CV, Borges-Walmsley MI, and Walmsley AR. Evidence for the assembly of a bacterial tripartite multidrug pump with a stoichiometry of 3:6:3. *J Biol Chem*, 286:26900–26912, 2011.