Università degli Studi di Cagliari

PHD DEGREE

Physics

Cycle XXXIII

Computational Studies on Pharmaceutical Targets in Human Diseases

Scientific Disciplinary Sector:

FIS/07

PhD Student: Chiara Fais

Coordinator of the PhD Programme: Prof. Paolo Ruggerone

Supervisors: Prof. Paolo Ruggerone,
Dr. Attilio V. Vargiu

Final exam. Academic Year 2019 – 2020
Thesis defence: February 2021 Session

To my grandparents

Chiara Fais gratefully acknowledges Sardinian Regional Government for the financial support of her PhD scholarship (P.O.R. Sardegna F.S.E. - Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2014-2020 - Axis III Education and training, Thematic goal 10, Investment Priority 10ii), Specific goal 10.5. Chiara also would like to acknowledge HPC-Europa3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme, for providing financial support and state-of-the-art high-performance computing for her research during her scientific visit at Queen Mary University of London (London, UK).

Contents

List of Figures

List of Tables

Nomenclature

ABC	ATP-binding cassette
$AcrB_{WT}$	wild type AcrB
$AcrB_{G288D}$	G288D variant of AcrB
AMY	amitriptyline
cryo-EM	cryo-electron microscopy
cryo-ET	cryo-electron tomography
CCF	cross-correlation function
CIP	ciprofloxacin
CPZ	chlorpromazine
CH1-4	channel 1-4
DBP	distal binding pocket
DBP_T	distal binding pocket of monomer T
dRMSD	distance root mean square deviation
*Ec*AcrB	*Escherichia coli* AcrB
EPI	efflux pump inhibitor
HT	hydrophobic trap
IMP	inner membrane protein
HDX-MS	hydrogen-deuterium exchange mass spectrometry
MATE	multidrug and toxic compounds extrusion
MD	molecular dynamics
MDR	multidrug resistance
MFP	membrane fusion protein
MFS	major facilitator superfamily

MIC	minimum inhibitory concentration
MM/GBSA	molecular mechanics with generalized Born surface area
MM/PBSA	molecular mechanics with Poisson-Boltzmann surface area
NMP	1-(1-napthylmethyl)-piperazine
NOR	norfloxacin
OMP	outer membrane protein
PAβN	phenylalanyl-arginine-β-naphtylamide
PBP	proximal binding pocket
PBP_L	proximal binding pocket of monomer L
PME	particle mesh Ewald
RMSD	root mean square deviation
RMSF	root mean square fluctuation
RND	resistance-nodulation-cell division
SMR	small multidrug resistance
*S*TmAcrB	*Salmonella* Typhimurium AcrB
TM	transmembrane

Abstract

Bacterial multidrug resistance (*i.e.* the ability of some bacterial species to survive in presence of various drugs) has become a primary challenge at a global level. Due to various factors, such as the overuse of antibiotics in human activities like health care and farming or inadequate diagnostic, many bacteria have indeed evolved acquiring novel and highly efficient resistance mechanisms. Some species, in particular, have become resistant to almost all in-use drugs.

Among the several mechanisms of resistance, efflux pumps of the RND superfamily (resistance-nodulation-cell division) play a major role. These complexes span the cell wall and are able to expel a wide range of noxious compounds, including antibiotics of many different classes. In order to reinvigorate the action of these drugs, a viable route is to hinder their transport out of the cell through co-administration of efflux pumps inhibitors (EPIs). At present several EPIs have been identified, but none of them is usable in clinical therapies due to adverse effects. Moreover, several questions are still open regarding the mode of action of known EPIs as well as the functioning mechanism of RND efflux pumps. Further research in this field is thus needed.

In order to characterize the mode of action of several EPIs of this pump, we applied computational techniques such as molecular docking and molecular dynamics (MD) simulations. Specifically, we focused on the EPIs: (i) amitriptyline and chlorpromazine, repurposed drugs which were proven to act as inhibitors against AcrB; (ii) PAβN, a known inhibitor of the pump whose mode of action is not fully understood.

This thesis focuses on the inhibition of the AcrB efflux pump, the best known representative of the RND superfamily. High-resolution structural data are indeed available for this protein (specifically, for its *Escherichia coli* orthologue). Moreover, a fluoroquinolone resistant variant of this pump has been detected in clinical environments.

With regard to amitriptyline and chlorpromazine, our *in silico* investigations revealed that both compounds tend to occupy a known binding pocket of AcrB. Their binding mode presents considerable similarities with that of several substrates and other EPIs of the pump, indicating that amitriptyline and chlorpromazine may inhibit the AcrB pump through competitive binding.

In the case of PAβN, MD simulations were compared with experimental data from hydrogen-deuterium exchange mass spectrometry. From these analyses, it emerged that PAβN can significantly restrain the conformational dynamics of AcrB and its fluoroquinolone resistant variant. This EPI, therefore, may act by preventing conformational changes that are functional for AcrB. Importantly, our MD simulations revealed that PAβN and the antibiotic ciprofloxacin can simultaneously occupy the same binding pocket, suggesting that the EPI does not act by competitive binding.

Further computational analyses were conducted on structural models of *Salmonella* Typhimurium AcrB. Experimental structural data on this wt protein are indeed missing, while

the structure of its fluoroquinolone resistant variant has recently been solved through cryo-electron microscopy (cryo-EM). In order to assess the structural differences between the two proteins, we derived their structural models through homology modelling and MD simulations (modeling of the fluoroquinolone resistant variant was integrated with cryo-EM data). Structural analyses were then performed, with focus on the binding pockets of the protein. Considerable differences were detected regarding the volume as well as the hydration properties of the pockets. Although not strictly related to EPI development, this information may be valuable for the design of novel drugs and/or inhibitors of AcrB from *Salmonella*.

Chapter 1

Introduction

1.1 The phenomenon of multidrug resistance

Bacterial resistance to antibiotics is a widespread phenomenon in nature [1, 2]. Resistant species have indeed been detected in diverse pristine environments (*i.e.* free from antibiotics of anthropogenic origin), including Antarctic soils [65] or isolated caves [66]. This phenomenon is thought to have originated at prehistorical times [155]. It could be the result of a natural selection process [67], related to the development of antibiotic-secreting systems by some microorganisms like fungi and bacteria [155]. Among them are, for example, Actinobacteria, a taxonomic group [69] which includes several antibiotic-secreting species, like *Streptomyces griseus* (the producer of the clinical antibiotic streptomycin) [155]. Phylogenetic analyses of these bacteria [69] indicated that antibiotic production and resistance mechanisms were present over 150 million years ago, suggesting that they might have evolved contemporaneously.

Despite the ancient origin of natural antibiotics, their presence in the environment does not seem to have exerted a strong selective pressure. Antibiotic-susceptible bacteria are indeed commonly found in ecosystems [2, 5], where they can co-exist with antibiotic-producing as well as resistant species [155]. Such equilibrium conditions, however, underwent a deep alteration in the past century [16], upon the discovery of antibiotics by mankind [9]. Starting with the identification of Salvarsan by Paul Ehrlich (1909) and of penicillin by Alexander Fleming (1929) [9], indeed, the 1900s were characterized by the rapid discovery of a high variety of antibiotics [10]. Many classes of in-use drugs were discovered between 1940 and 1960 (the 'golden age' of antibiotic discovery) [42] and were rapidly made available (Fig. 1.1) [11]. They found an application not only in human medicine, but also in agriculture and animal farming [13]: nowadays, it is estimated that 4 to 400 mg of antimicrobials are used to produce 1 kg of meat in European countries [16]. Such an intensive use of antibiotics and bactericidal compounds has determined a much stronger selective pressure than that present in pristine ecosystems [13]. This, in turn, has favoured the acquisition of resistance mechanisms by previously susceptible bacteria [70].

At present, several species are known to have developed resistance against more therapeuticals [5, 12]. This phenomenon, known as multidrug resistance (hereafter MDR), has become a global concern according to the World Health Organization [12]. It is estimated [10], indeed, that MDR is responsible for approximately 20% of deaths worldwide, being involved in over 15% nosocomial infections. Without a proper counteraction, this balance could become even more dramatic, with MDR causing 10 million deaths per year in 2050 [10, 13]. In this regard,

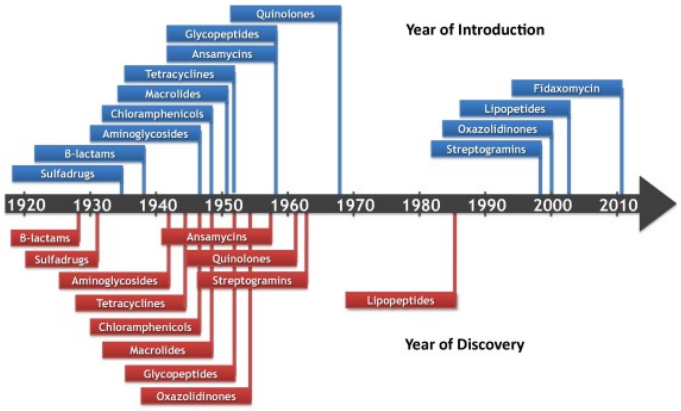

Figure 1.1: Timeline of the discovery and introduction to clinical use of the main antibiotic classes. Image from [11].

pathogens of the ESKAPE group (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumanii, Pseudomonas aeruginosa* and *Enterobacteriaceae*) are of considerable importance [110], since they have developed resistance to most in-use antibiotics [16]. A possibility for a new treatment comes from a recent study by Smith *et al.* [19], which proved that semi-synthetic derivatives of arylomycins (a class of natural antibiotics) can be effective against ESKAPE pathogens. However, further research is needed in order to develop a new class of antibiotics.

In addition to the identification of new drugs, several actions are needed to prevent further spreading of MDR. These include important investments on research [29]. Indeed, significant contributions may come from the development of alternative clinical approaches, such as vaccines or antibody therapies [12]. Moreover, we need to deepen our knowledge of bacterial resistance mechanisms [29], about which many aspects are not clear. This could lead to innovations in the identification of targets for drug design [15], or provide hints for the inhibition of MDR mechanisms [28].

Importantly, efforts in research should be paired to the adoption of new practices on a global scale [37]. Antibiotic use in agriculture and farming should be limited, in order to avoid their spreading in the environment [35]. At the same time, diagnostic protocols should be improved to prevent errors in therapeutical approaches [37]. In this regard, efforts should be made also in the identification of sub-standard drugs, especially in low income countries [18]. These drugs are indeed not adequate for the treatment of infections (for example, because of poor absorption levels) and their use may ease the diffusion of MDR [18]. Due to the complexity of the problem, a 'One health approach', *i.e.* a coordinated agenda to account for intervention on human and veterinary medicine, research and environmental sciences, has been proposed by the World Health Organization to counteract the threaten of MDR worldwide [29].

1.2 Resistance mechanisms in bacteria

1.2.1 The bacterial cell wall: a first defence against external agents

In bacteria, a first and important defence against toxins and other external agents is represented by the cell wall, the most external layer of the cell [20]. This region constitutes a barrier, preserves the cell shape and regulates exchanges with the environment at the same time [20]. Moreover, specific mechanisms allow to regulate its permeability [21], as described in the following paragraphs.

The cell wall does not have the same structural characteristics in all bacterial species. Specifically, two main architectures have been identified, according to which bacterial species are classified as follows [20]:

- gram-positive bacteria, whose cell wall (Fig. 1.2a) is composed by a cytoplasmic (or inner) membrane and an outer layer of peptidoglycan, a polymer which plays a protective and structural role similar to that of an exoskeleton. Peptidoglycan also contains several acid molecules that are functional for cell mobility, and is anchored to the inner membrane through lipoproteins. In addition to them, the inner membrane contains various proteins with different functioning, such as uptake of nutrients from the environment;

- gram-negative bacteria, in which the cell wall consists of an inner membrane, a peptidoglycan layer and an outer membrane (Fig. 1.2b). In this case, peptidoglycan is enclosed in the region delimited by the two membranes (named periplasmic space) and is linked to both of them through lipoproteins. As in gram-positive bacteria, exchange of compounds with the environment is possible thanks to several proteins, which are present in both membranes as well as in the periplasmic space. Moreover, the cell wall is enriched with many proteins and other biomolecules playing different roles, such as lipopolysaccharides, toxins present in the outer membranes that are released during host attacks. Lipopolysaccharides also lower the permeability of the outer membrane, strongly limiting the penetration of toxic agents inside the cell [21].

In gram-negative bacteria, therefore, the outer membrane constitutes an additional protection against external agents, including several drugs [21]. This represents an important contribution to the occurrence of MDR in gram-negative pathogens, which include the majority of the ESKAPE species [22]. Additional resistance mechanisms, common to gram-positives and gram-negatives, are described in the following paragraph.

1.2.2 Classification of MDR mechanisms

Bacterial resistance mechanisms can be devided in four main categories [21, 23]:

- modification of the drug target: drugs often act by stably binding specific bacterial enzymes in their active sites, thus compromising their functionality. In this cases, resistance mechanisms can consist in mutations inside or in proximity of the active site, in such a way to reduce the binding affinity of the drug without altering the enzyme activity;

- drug inactivation: with these mechanisms, drugs are modified inside the bacterial cell, through degradation or addition of chemical groups by specific enzymes;

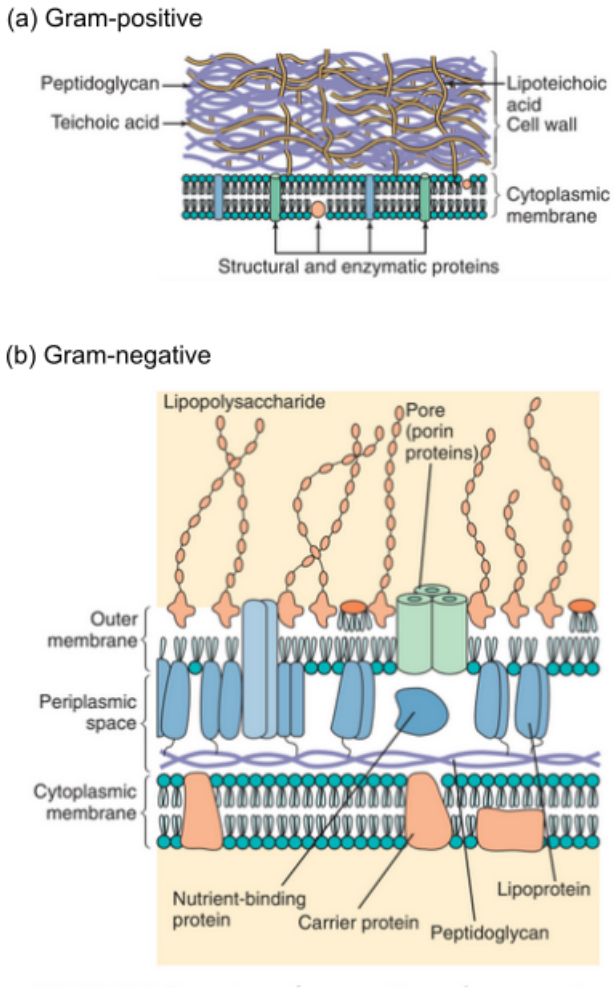

Figure 1.2: Structure of the cell wall in gram-positive (a) and gram-negative (b) bacteria. Exchange of ions and other molecules with the environment is possible through various proteins located in the cell wall, such as porins, carriers and nutrient-binding proteins. Image adapted from [20].

- limiting the uptake of compounds from the environment: this can be achieved in multiple ways. In gram-negative bacteria, as previously mentioned, a first barrier is consituted by the outer membrane. Moreover, if necessary, bacteria (both gram-positive and gram-negative) can decrease the number of proteins dedicated to the uptake of specific compounds and representing a key entrance path for antibiotics;

- active efflux of drugs: bacterial cells express a wide variety of efflux systems (or pumps). These are protein complexes located in the cell wall that are able to actively expel diverse toxic compounds from the cell. These include dyes, detergents and a wide variety of drugs, which need to reach a threshold concentration in the cytoplasm to become effective. Additionally, bacteria can vary the expression rate of efflux pumps on need, as in the case of proteins responsible for uptake of compounds.

Such mechanisms can be intrinsic (*i.e.* they naturally occur in the bacterial species) or acquired (they are expressed upon mutation in the DNA of the cell, or after the transfer of genetic material from other bacteria) [21]. Among the intrinsic mechanisms, one of the main examples is the expression of efflux pumps, which are found in all species [23]. These machineries play a peculiar role in MDR. Indeed, due to their broad spectrum of substrates, active efflux alone can allow the survival of the cell in the presence of various drugs [24]. This initial protection, in turn, can lead the cell to acquire additional and more specific resistance mechanisms [25, 26], through mutations or possibly the uptake of genetic material from the environment.

Efflux mechanisms are especially effective in gram-negative bacteria, where they are coupled to the low-permeability outer membrane [21]. In this bacteria, which include ESKAPE pathogens like *P. aeruginosa* [25], efflux pumps from a complex network that allows uptake of compounds from both the cytoplasmic and periplasmic region [24].

1.3 Efflux systems

1.3.1 Classification of efflux systems

Due to their relevance for the phenomenon of MDR, bacterial efflux systems have been an important matter of study in the past decades [28]. In order to shed light on their structure and functioning mechanism, several approaches have been used. They include structural characterization techniques (such as X-ray crystallography, cryo-electron microscopy and cryo-electron tomography - hereafter cryo-EM and cryo-ET, respectively) [31], molecular genetic studies and diverse methods to investigate their transport mechanism, like site-directed mutagenesis [28]. Significant contributions to our knowledge of these machineries have also been provided by computational approaches (like molecular docking and molecular dynamics simulations, described in Chapter 2), which have been used to investigate several features of the transport process at the molecular level [34].

A structural classification of efflux systems can be done in terms of the number and the topology of their components [31]. They can indeed consist of a single transmembrane protein (named transporter), which is located in the cytoplasmic membrane, or in tripartite complexes that span the cell wall [31]. While the former are present in both gram-positive and gram-negative bacteria [31], tripartite systems are characteristic of gram-negatives [32], whose cell wall, as described in Subsection 1.2.1, has a more articulated structure.

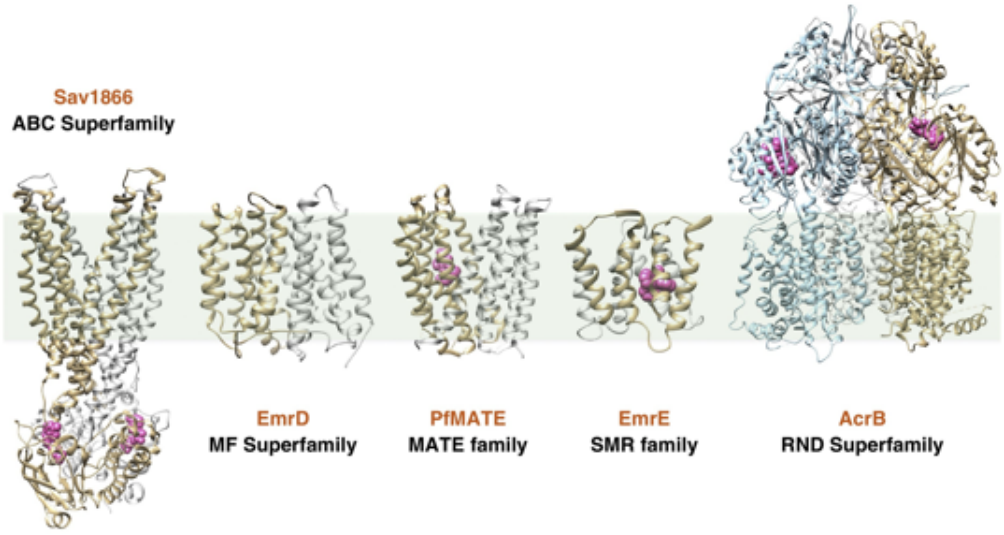

Figure 1.3: X-ray crystal structures of transporters of the identified superfamilies. The name of the crystallized protein and its family of belonging are reported in brown and black, respectively. In each structure, protomers are represented in different colours, while co-crystallized substrates are shown as pink spheres. PDB-IDs: Sav1866: 2HYD; EmrD: 2GFP; PfMATE: 3VVP; EmrE: 3B5D; AcrB: 3AOD. Image adapted from [27].

1.3.2 Transporters

Research on transporters has revealed a considerable diversity under many aspects. At present, five main classes have been identified according to several criteria, such as structural characteristics, phylogeny and the energy source for substrate extrusion [27, 28]. The identified classes are: (i) ATP-binding cassette superfamily (ABC), (ii) small multidrug resistance family (SMR), (iii) major facilitator superfamily (MFS), (iv) multidrug and toxic compounds extrusion family (MATE) and (v) resistance-nodulation-cell division superfamily (RND) [27].

With regard to the characterization of each class, high-resolution structures of some representatives (Fig. 1.3) have had a fundamental importance [31]. Comparison of some members of the MFS, MATE and SMR classes reveals similarities in their general architecture [27]. These transporters, indeed, consist in transmembrane (hereafter TM) helical bundle [27], whose size greatly varies depending on the considered class. Bundles composed by 12 helices have been frequently detected in MFS and MATE transporters, while SMR transporters composed by 4 helices have been identified [31]. Functional studies on these classes have lead to transport models in which the protein switches between two states, named inward-open and outward-open (Fig. 1.4a) [27, 31]. In the inward-open state, only the cytoplasmic end of the transporter is open, allowing the binding of the substrate. The binding event is thought to trigger the transition to the outward-open state, where the substrate is released on the opposite side of the cytoplasmic membrane. Upon substrate release, the protein goes back to the inward-open state. The energy source of this process frequently consists in the transport of ions (H^+ or Na^+) across the cytoplasmic membrane, through antiport or symport mechanisms (although some uniports have also been identified in the MFS class) [31].

A partially similar functional mechanism, based on the switch between an inward-open

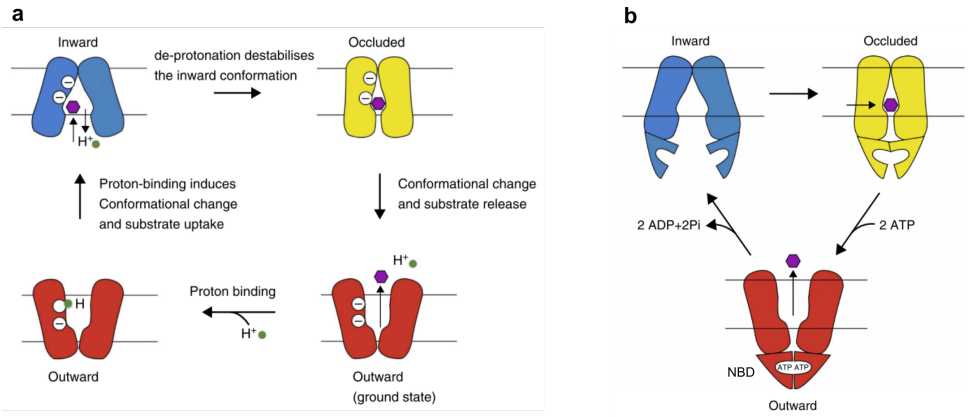

Figure 1.4: Functioning models of MFS (a) and ABC (b) transporters. Similarly to MFS transporters, SMR and MATE representatives expel their substrates through the switch between an inward-open and an outward-open conformation, coupled to the transport of ions (H^+ or Na^+) across the membrane. Image adapted from [27].

and an outward-open state, has been proposed for ABC transporters [27, 31]. These proteins, however, possess a peculiar energy source for substrate extrusion, which is ATP hydrolysis [31].

In contrast to the transporters described so far, members of the RND superfamily do not extrude their substrates through the described two-states mechanism [31]. These proteins are indeed characterized by a large domain located above the cytoplasmic membrane, which contains several binding pockets and channels for substrate transport [33]. The TM domain does not directly interact with substrates [33], and is involved in the transport of protons through antiport mechanisms [35]. In the best studied RND transporters, which belong to gram-negative bacteria, the extrusion process is thought to consist in three steps [31]. Firstly, the substrate accesses the extra-membrane domain through dedicated channels (step 1). In steps 2 and 3, it is progressively pushed towards the upper end of the extra-membrane domain, while ions are transported across the TM domain [31]. This functioning mechanism, named functional rotation [27], is described in greater detail in Section 1.4.

1.3.3 Tripartite efflux pumps

1.3.4 Structure

In gram-positive bacteria, transporters operate as isolates [27]. gram-negatives, on the other hand, possess transporters as well as tripartite efflux pumps, as mentioned previously. The latter extend for the whole length of the cell wall (roughly 30 nm) [36], and are composed by [27, 32]:

- a transporter (in tripartite pumps, this component is also named inner membrane protein, or IMP). The transporter determines the class of the tripartite pump, which typically is ABC, RND or MFS;
- an outer membrane protein (OMP), whose characteristic structure consists in a TM β-barrel (located in the outer membrane) and a helical tube-shaped domain that extends for roughly 10 nm in the periplasm;

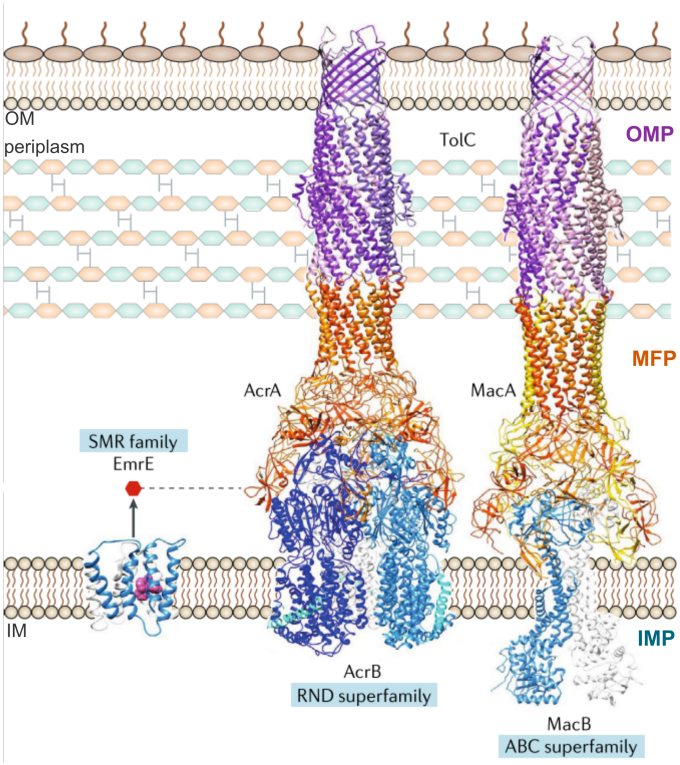

Figure 1.5: Examples of efflux systems in gram-negative bacteria. While ABC multi-component pumps bind their substrates in the cytoplasm and expel them, RND pumps uptake substrates in the periplasmic region, thus cooperating with isolated transporters that push toxins across the inner membrane (such as those of the SMR family). Proteins of the multi-component pumps are represented in different colours, with each protomer shown in a different shade. Transporters (labeled IMP) are represented in blue and gray, MFP in yellow and orange, OMP in pink and purple. Image adapted from [31].

- a membrane fusion protein (MFP), which is enterely located in the periplasm but anchored to the inner membrane. It connects the transporter to the OMP. Similarly to the OMP, the MFP possess a peculiar structure, composed by a chain of two or three globular domains and an helical hairpin. Its size varies according to the considered system.

Structures of the assembled complex have been obtained through cryo-EM for two tripartite pumps of *E. coli*. These are AcrA-AcrB-TolC [36, 37] and MacA-MacB-TolC [38] (also named AcrAB-TolC and MacAB-TolC), which belong to the RND and ABC classes, respectively [31]. These systems (Fig. 1.5) are composed by the proteins AcrB and MacB (transporters), AcrA and MacA (MFPs) and TolC (OMP), which is the same in the two pumps [35, 36, 37]. In both of them, six copies of the MFP are present, which form a tube-shaped homohexamer. The helilcal hairpins of each MFPs are in contact with the OMP, while part of the globular domains interacts with the transporter [35, 36, 37].

1.3.5 Functioning mechanism

Although several structural data are available for several tripartite pumps [27, 32], many aspects of their functioning are still a matter of debate [32]. As mentioned, substrate extrusion models have been formulated for transporters of all classes (Section 1.3), through studies conducted on their representatives [31]. However, less is known about the functioning of the MFP and the OMP [32].

With regard to *E. coli* AcrAB-TolC, a possible mechanism has recently been proposed by Shi *et al.* [39] on the basis of *in situ* cryo-ET structures of the assembly. According to the cryo-ET data, AcrAB exists as isolate in the absence of substrates (Fig. 1.6a). When a substrate (puromycin) is introduced, the assembly of the complete AcrAB-TolC pump is observed more frequently (62% of the sample). Under this condition, however, the AcrA-TolC interface is not open (Fig. 1.6b,c,e), as previously reported in cryo-EM studies [36, 37], but rather presents an occlusion. The AcrA-TolC channel is widely open only in the presence of an inhibitor of the pump, MBX3132 (*i.e.* a small molecule that directly interacts with the pump, preventing substrate extrusion - see Section 1.5) (Fig. 1.6d). On the basis of these data, the authors have hypothesized that the binding of TolC to the AcrAB complex is triggered by substrate binding. In the tripartite complex, however, the AcrA-TolC interface opens only transiently to allow the passage of the substrate. It is possible that the opening event involve interactions of both AcrA and TolC with peptidoglycan, which lacks in cryo-EM samples. This may explain why this configuration has not been detected in previous studies. The open state of AcrA-TolC may be stabilized in presence of the inhibitor MBX3132. This compound could indeed lock the pump in the final stage of the transport process, thus affecting its functionality. Although this work offers interesting insights on the functioning of AcrAB-TolC, and on the mode of action of its inhibitor MBX3132, further studies are needed to support the proposed mechanism. Several aspects, indeed, need to be cleared, such as the dynamics of the AcrA-TolC complex during the extrusion process and the possible role of peptidoglycan.

In addition to the transport processes *per se*, another relevant aspect of the functioning of tripartite pumps is their cooperation with isolated transporters. As stated previously, while ABC and MFS transporters can sequester their substrates from the cytoplasm, the transport pathway of RND transporters is enterely located in their extra-membrane domain [33]. RND tripartite pumps alone, therefore, can extrude noxious compounds from the periplasm, but not from the cytoplasm. However, these machineries typically possess a very broad substrate

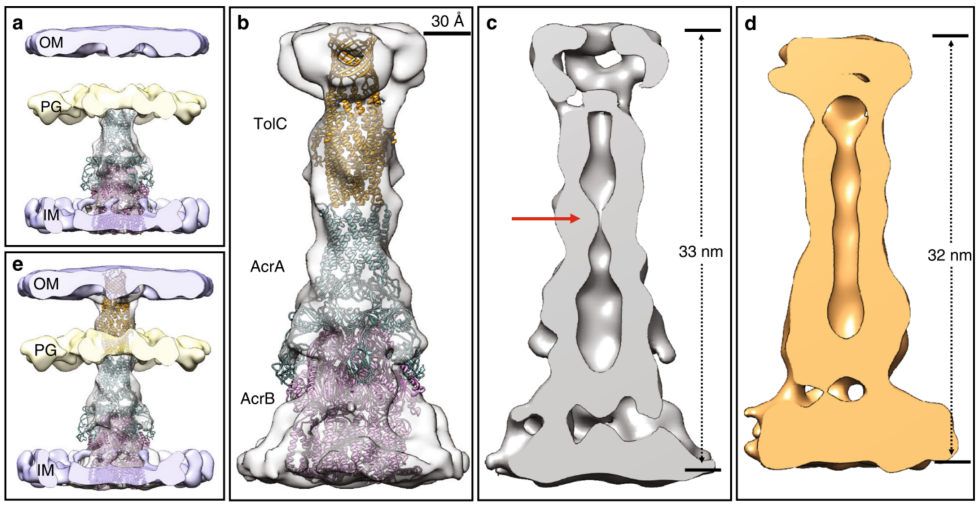

Figure 1.6: (a) *In situ* cryo-ET structure of the AcrAB complex. The outer and inner membranes (OM and IM, respectively) are also represented in purple, while the peptidoglycan layer (PG) is shown in yellow). (b) Structural model of the AcrAB-TolC complex fitted in cryo-ET data. (c) Section of the cryo-ET structure of AcrAB-TolC obtained in presence of a substrate (puromycin). The restriction at the AcrA-TolC interface is indicated. (d) Section of the cryo-ET structure of AcrAB-TolC in presence of the inhibitor MBX3132 (see Section 1.5). (e) Cryo-ET structure of the assembled AcrAB-TolC pump. The outer an inner membrane as well as the peptidoglycan layer are shown (same colour code as (a)). Image adapted from [39].

range, part of which is in common with one or more transporters expressed by the same bacterium [31]. This allows expulsion of toxins from the cytoplasm in two steps [32]. Firstly, the toxin is fed into the periplasm by a transporter. Once here, it enters the RND tripartite machine, which complete the extrusion process [32]. Until structural data on the MacAB-TolC pump became available, this double-step mechanism had been thought to be exclusive of RND pumps [31]. However, analyses of the cryo-EM structure of MacAB-TolC lead to the detection of a putative substrate binding site at the interface between MacB and MacA [38]. This suggests that uptake of substrates from the periplasm could be possible also for ABC tripartite pumps [38], and maybe for machineries of the remaining superfamilies (for which structural data on the assembly are missing). In this regard, it must be pointed out that tripartite pumps able to extrude toxins from the periplasm can be very beneficial for the cell. Indeed, in addition to cooperating with transporters, this machineries can also lead back to the cell exterior compounds that penetrated the outer membrane, preventing them from reaching the cytoplasm [43].

1.3.6 Clinical relevance of efflux pumps

From the clinical point of view, relevant efflux pumps have been identified in all classes [31]. In gram-positives, the main contributions to MDR are provided by transporters of the MFS superfamily [40]. A paradigmatic case is that of *S. aureus*, a gram-positive member of the ESKAPE group (see Section 1.1), in which the MFS representatives NorA and QacAB provide resistance to hydrophilic fluoroquinolones, such as ciprofloxacin, and various biocides and antiseptics [31]. In the same bacterium, a significant role is also played by the ABC transporter

MrsA, which is involved in the efflux of macrolides and streptogramins [31].

Clinically relevant transporters have also been identified in gram-negative species. These include MdfA, an MFS transporter of *E. coli* involved in the extrusion of fluoroquinolones, tetracyclines and aminoglycosides [40, 41]. Additional examples are EmrE and NorM, which are expressed by the gram-negative *Neisseria gonorrheae* [31]. EmrE, which belongs to the SMR family, can transport drugs of different classes, like macrolides, aminoglycosides and β-lactams [31]. NorM, a MATE transporter, can confer resistance to hydrophilic fluoroquinolones. In addition to transporters, several multi-component efflux pumps provide major contributions to MDR in gram-negatives [31]. The aforementioned AcrAB-TolC and MacAB-TolC, expressed by *E. coli* and other *Enterobacteriaceae*, confer resistance to a very wide range of antibiotics [32]. Substrates of AcrAB-TolC, indeed, involve β-lactams, fluoroquinolones, novobiocin, tetracycline, erythromycin and chloramphenicol [31, 42], while MacAB-TolC is involved in the efflux of macrolides [31]. Additional examples of relevant multi-component pumps include the RND pumps AdeABC of *Acinetobacter baumannii* and MexAB-OprM of *P. aeruginosa*, whose range of transported drugs present considerable similarities with that of AcrAB-TolC [31].

My thesis is mainly focused on the RND transporter AcrB, expressed by *Enterobacteriaceae* and part of the AcrAB-TolC pump. A short description of the structural characteristics and the functioning mechanism of this transporter is provided in the following.

1.4 AcrB: a paradigm of RND transporters

1.4.1 Structure

The AcrB transporter of *E. coli* is, at present, the best known representative of the RND superfamily [42]. Indeed, it was the first transporter of this class to be crystallized in 2002 [44] and, since then, it has been the subject of numerous studies [32, 42]. At present, several crystal structures are available for this protein with resolution below 2.5 Å [45, 46, 47]. Moreover, cryo-EM and cryo-ET maps have recently been obtained for the isolated AcrB [48] as well as for the assembled AcrAB-TolC pump [36, 37, 39] (see Section 1.3).

Studies on AcrB revealed that this protein is a homotrimer (Fig. 1.7) [44]. Each monomer is composed by 1049 amino acids, and has a height of 120 Å [44]. It consists in a TM domain (50 Å high) and an extra-membrane portion characteristic of RND transporters (see Section 1.3). It is composed by the pore domain (40 Å high, immediately above the TM domain) and the docking domain (30 Å high, at the upper extremity of the protein) (Fig. 1.7a) [44].

The topology and structure of the monomers are represented in Fig. 1.8a and b, respectively. In each of them, the TM domain is composed by 12 helices [44], which are arranged in two pseudo-symmetric subdomains named R1 (helices TM1-6) and R2 (TM7-12) (Fig. 1.8a,b) [42, 49]. These subdomains are linked by an extra-membrane helix, named I-α, which is located in the cytoplasm (Fig. 1.8a,b) [44]. On top of the TM region, the pore domain consists of four subdomains (labeled PN1, PN2, PC1, PC2), each composed by two α-helices linked to anti-parallel β-strands [44]. Similarly, two subdomains have been identified in the docking domain, named DN and DC and mainly composed by anti-parallel β-strands [44]. The DN subdomain presents a peculiar structure, being characterized by a protruding loop (or connecting loop) that is involved in interactions with the adjacent monomer (Fig. 1.8b) [44, 50].

The assembled protein presents a considerable number of cavities and channels, which constitute the transport pathway (Fig. 1.8c,d). The TM domains of the three monomer enclose

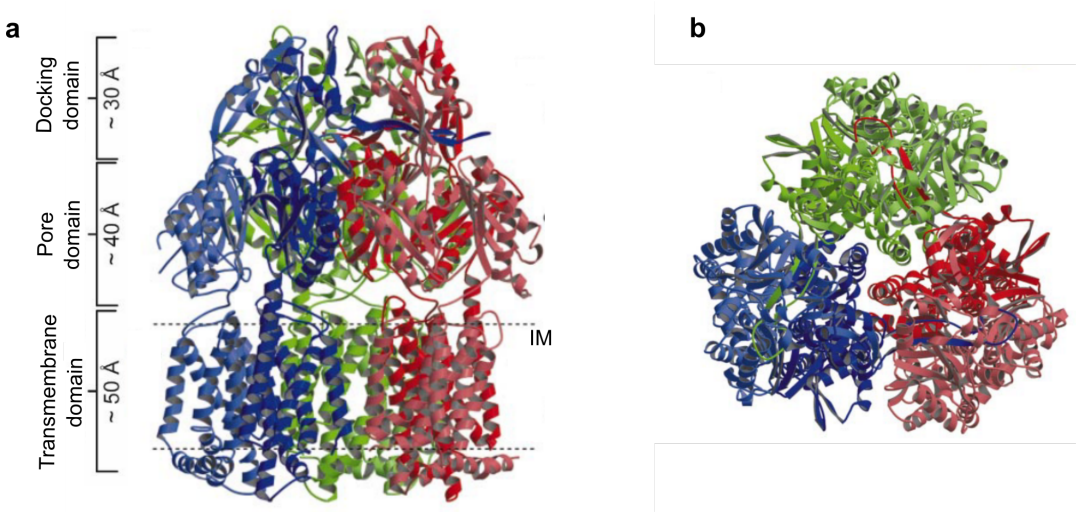

Figure 1.7: Side view (a) and top view (b) of *E. coli* AcrB (PDB ID: 1IWG). Monomers are represented in different colours. Adapted from [44].

the so-called central cavity (Fig. 1.8c), which has a diameter of 30 Å [44] and is partially filled with lipids *in vivo* [53]. It is connected to the exterior of the protein thrugh three vestibules (Fig. 1.8c). These are inter-monomer cavities located right above the membrane, which extend for roughly 15 Å in height [42, 44].

In the pore domain of every monomer, two binding pockets have been identified, named proximal binding pocket (PDP) and distal binding pocket (DBP) [45] (Fig. 1.8c). The PDP, located between subdomains PC1 and PC2, is the more external pocket. A flexible and glycine-rich loop (named switch loop or G-loop) separates it from the DBP, which is enclosed between subdomains PN1, PN2 and PC1 [42, 45]. From the DBP, a channel (or exit gate) extends towards the cavity formed by the docking domain of the three monomers, also referred to as central funnel (30 Å in diameter) (Fig. 1.8c) [44].

Structural studies on AcrB have lead to the identification of several access route to the binding pockets. With regard to the PBP, two possible entries have been detected, named channel 1 and channel 2 (hereafter CH1 and CH2, respectively) (Fig. 1.8c) [42, 51]. CH1 is located in proximity of the vestibule, being formed by the PC2 subdomain and the TM domain. CH2, or periplasmic cleft, is instead formed by the most external part of the PC1 and PC2 subdomains. A study on the transport pathway in AcrB revealed that mutations in CH1 and CH2 affect the efflux of minocycline, doxorubicin and erythromycin, suggesting that these drugs access the PBP through the mentioned entries [51].

In addition to CH1 and CH2, channels linking the DBP to the external environment have also been identified. These include channel 3 (CH3) [51], which connects the DBP to the central cavity (Fig. 1.8c). According to a mutagenesis study [51], this is the preferred access route for planar aromatic cations with low molecular mass (less than 500 Da). Compounds with this characteristics are, for example, the antimicrobial benzalklodium or the fluorescent dye ethidium bromide. Moreover, another entry to the DBP, named channel 4 (CH4), has recently been identified (Fig. 1.8d) [52]. It is located at the interface between the PN2 subdomain and the TM domain, and is thought to be mainly involved in the transport of drugs like β-lactams and fusidic acid. Investigations on this access route have revealed that substrates may bind

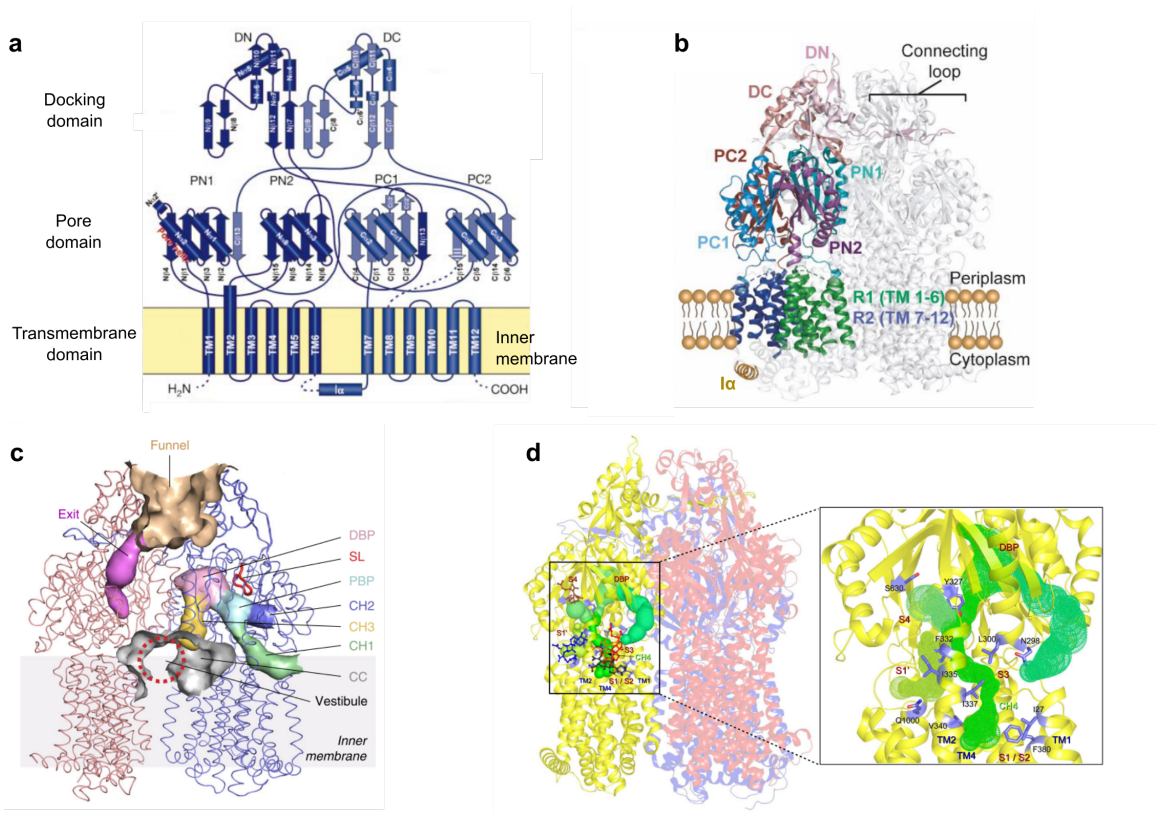

Figure 1.8: (a) Topology diagram of an AcrB monomer. Secondary structure elements are represented as cylinders (helices) and arrows (β-strands). Dotted lines correspond to unstructured segments. Image adapted from [44]. (b) Structure of an AcrB monomer. Subdomains are labeled and represented in different colours. Moreover, the I-α helix (linking the R1 and R2 subdomains, as shown in (a)) and the connecting loop are indicated. (c) Cavities and channels in the AcrB transporter. Truncated labels refer to: central cavity (CC), channel 1-3 (CH1-3), proximal binding pocket (PBP), switch loop (SL) and distal binding pocket (DBP). Image from [51]. (d) Substrate pathways constituting channel 4 (CH4). The considered pathways (labeled S1-4, S1') are shown as green meshed surfaces. Nearby residues affecting AcrB activity according to functional analysis are represented as blue sticks. Image from [52].

AcrB on the top of the TM domain (helices TM1 and TM2), and from here they may reach the DBP through multiple pathways, which constitute CH4.

1.4.2 Transport mechanism

At present, the transport mechanism of AcrB, and of RND transporters in general, is not completely understood. A widely accepted model, named functional rotation, has been formulated on the basis of experimental data (such as crystal structures) [45, 54] as well as computational investigations [42].

According to this hypothesis, the transport of substrates by AcrB requires each of its monomers to undergo a sequence of conformational changes [42]. The involved conformations (for which crystal structures are available [45]) are named loose (L, or access), tight (T, or binding) and open (O, or extrusion) states [42].

Significant differences between such states regard the configuration of the porter domain [54]. Indeed, in the L state the PBP and its access routes (CH1 and CH2, see Subsection 1.4.1) are widely open, due to the marked separation of the PC1 and PC2 subdomains. However, subdomain PC1 is in contact with PN1 and PN2, determining the occlusion of the DBP and the exit gate. In the T state all subdomains in the porter domain are instead well distanced, so that both the PBP and the DBP are accessible. The entry channels to the DBP (CH3 and CH4, see Subsection 1.4.1) are open as well, while the exit gate remains closed. Passage through the latter channel is possible only in the O state. In this configuration, indeed, the inclination of PN1 varies of roughly 12° away from PN2, determining the opening of the gate. On the contrary, the PBP, DBP and their entry channels are closed, due to the fact that subdomains PC1, PC2 and PN2 are in close contact with each other.

The functional rotation model [54] has originally been formulated for substrates that access the AcrB monomers through the CH1 and CH2 entries, which lead to the PBP. According to it, substrates firstly access the AcrB monomer in the L state, reaching the PBP. This event triggers the switch from the L to the T state, in which the substrate enters the DBP. With a second conformational change, from T to O, the substrate is expelled through a sort of peristaltic motion. This is created by the occlusion of the DBP, coupled to the opening of the exit gate. Once the substrate has been extruded, the monomer returns to the initial L state. The completion of the transport process, therefore, requires a cycle of conformational changes, which is L→T→O→L. This process is coupled to the protonation of charged residues located in the TM domain, at the interface between the R1 and R2 subdomains (helices TM6 and TM10) [54]. Indeed, according to experimental and *in silico* investigations, protons access and bind the involved residues in the T→O transition, and are then released in the cytoplasm through the following O→L transition [55]. This influx mechanism provides the required energy for the completion of the functional rotation [46, 55].

Studies based on X-ray crystallography have shown that AcrB can exist in asymmetric states, its monomers being at different stages of the functional rotation process [45, 54]. These states include, for example, LTO or LLT [42]. Moreover, the recent identifications of CH3 [51] and CH4 [52] suggest that some substrates can directly access the DBP through these entries, bypassing the PBP. This is supposed to happen in the T state, when the DBP and its channels are open [51, 52].

Additional investigations have shed further light on the transport mechanism and the properties of the PBP and DBP. It has indeed been shown that the DBP presents a higher percentage of hydrophobic residues than the PBP (52% and 41%, respectively) [42]. Since

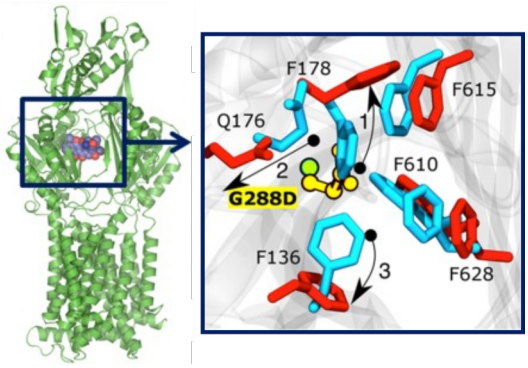

Figure 1.9: Overall view of an AcrB monomer, with zoom on the surroundings of residue 288 (located in the DBP). In the close view, the DBP of the wt AcrB (crystal structure, PDB ID: 4DX7 [45]) and its G288D variant (MD simulation data) are superimposed. Residue 288 is shown in green (wt protein) and yellow (G288D variant). Moreover, nearby residues are represented in blue (wt) and red (G288D). Relevant changes in the orientation of the represented residues are highlighted through black arrows. Image adapted from [143].

substrates of AcrB share a certain degree of lipophilicity, this characteristic is thought to favour the displacement of substrates towards the DBP in the T→O transition [42]. In this regard, an important role is also played by the switch loop (see Subsection 1.4.1), which possess a considerable degree of flexibility because of the presence of several glycine residues. Mutation of such residues restrain the dynamics of the loop and have been shown to affect the functionality of the transporter [56]. This is possibly due to the fact that a more rigid switch loop does not allow the passage of substrates from the PBP to the DBP [56].

The presence of hydrophobic residues in the binding pockets is thought to be of importance also for the substrate specificity of AcrB. Indeed, studies conducted on this transporter and on several homologues suggest that lipophilicity of the pockets, together with additional features (such as shape, electrostatic potential and hydration), plays an important role for the substrate specificity of the transporter [118, 135]. These findings are consistent with our present knowledge of the fluoroquinolone resistant variant of AcrB, bearing the G288D substitution. This mutation, detected during the treatment of a clinical patient with *Salmonella* infection, determines an increased resistance to fluoroquinolones (such as ciprofloxacin) as well as an enhanced sensitivity to doxorubicin and minocycline [144, 145]. Such alteration is coupled to significant variations in the structure and hydration properties of the DBP, where residue 288 is located. The G288D substitution has indeed been proven to determine a net increase in the hydration of the DBP, causing the reorientation of several residues of the hydrophobic trap (a niche of the DBP enclosed between PC1 and PN2, rich in hydrophobic and aromatic residues) (Fig. 1.9) [143]. These variations, in turn, may considerably affect the binding affinities of the AcrB substrates [143].

1.5 Efflux pump inhibitors

A viable route to contrast the action of RND tripartite pumps, such as AcrAB-TolC, is the development of efflux pump inhibitors (EPIs) [57]. These are compounds capable of preventing the extrusion of substrates by the pump. They may thus be co-administrated with antibiotics,

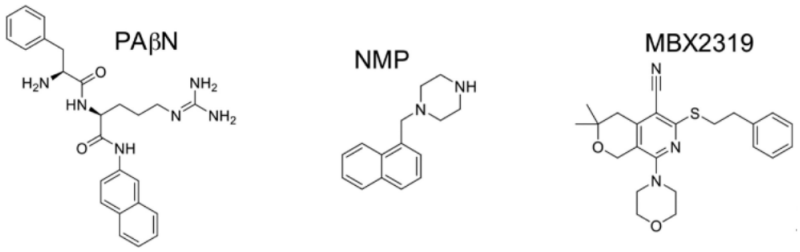

Figure 1.10: Chemical structures of the EPIs PAβN, NMP and of an inhibitor of the MBX series (MBX2319). Image adapted from [113].

in order to restore their efficiency [57].

Research on this topic, which started over 20 years ago [58], has lead to the identification of several EPIs [47, 57]. At present, however, none of them is suitable for clinical use, mainly because of cytotoxic effects that emerged in preclinical tests [58]. Further investigations are thus needed to overcome this difficulty, *e.g.* through the identification of new EPIs or the design of more efficient derivatives of the available ones [57].

With regard to *E. coli* and *Enterobacteriaceae*, the main EPIs discovered so far are phenylalanyl-arginine-β-naphtylamide (PAβN), 1-(1-napthylmethyl)-piperazine (NMP) and members of the MBX series of pyranopyridine derivatives [47, 57] (see Fig. 1.10). In *E. coli*, these compounds have been proven to be effective against the AcrAB-TolC efflux pump [47, 57]. Specifically, investigations conducted through mutagenesis [63], X-ray crystallography [47, 59] and *in-silico* [47, 60] approaches revealed that they can interact with the AcrB transporter and bind the DBP.

PAβN and NMP were discovered in 1999 [61] and 2005 [62], respectively. The former [61] was firstly identified through a screening for EPIs of the RND pump MexAB-OprM, expressed by *P. aeruginosa*. It was later proven to be effective also against *E. coli* AcrAB-TolC, increasing the strain sensitivity to drugs like levofloxacin and tetracycline [57]. NMP was similarly discovered with a screening procedure, having AcrAB-TolC as its target [62]. It is known to potentiate a considerable range of antibiotics, such as levofloxacin, chloramphenicol and ciprofloxacin [57]. Both PAβN and NMP, however, are effective only at high concentrations (50 μM or higher), and are thus not usable as therapeuticals [57]. Such high doses would indeed cause accumulation in tissues, with consequent toxic effects [58]. It is nonetheless important to investigate their mode of action, which is not fully understood at present. This could indeed lead to a deeper knowledge of the functioning of tripartite RND pumps, and possibly provide hints for the design of novel EPIs [57].

In this regard, structural data on the interactions of PAβN and NMP with AcrB are lacking, possibly due to the low binding affinity of these EPIs for the binding pockets of AcrB [37, 57]. At present, indeed, only one crystal structure of PAβN bound to the AcrB N109A variant is available [59], which shows that the EPI binds the PBP of the transporter. Significant information on the binding poses of both PAβN and NMP have however been obtained by mean of *in silico* approaches [60], through the combination of computational docking and molecular dynamic simulations. These investigations [60] have revealed that both PAβN and NMP can stably bind the DBP of AcrB, interacting with the switch loop and the surrounding region. Additional stabilizing interactions are formed by the aromatic groups of the EPIs

with the hydrophobic trap. For both EPIs, interactions with the switch loop may play a fundamental role in the inhibitory process, preventing the passage of substrates from the PBP to the DBP (see Section 1.4). Indeed, as mentioned in Subsection 1.4.2, mutagenesis studies [56] revealed that a more restrained switch loop affects the functionality of the transporter. Moreover, it has been hypothesized that the presence of PAβN and NMP in the DBP could not allow substrate binding, because of sterical hindrance [57].

In addition to PAβN and NMP, EPIs of the MBX series [47, 64] have been demostrated to potentiate a very wide range of substrates. The first member of this series (EPI MBX2319) was discovered in 2014 [64] and was proven to increase the sensitivity of the bacterial strain to various AcrB substrates, such as luvofloxacin, piperacillin and chloramphenicol. Importantly, MBX2319 and its derivatives are effective at concentrations of 3 μM or lower [47], *i.e.* at least one order of magnitude lower than those required by PAβN and NMP. Due to this characteristic and their broad efficacy, MBX EPIs are highly promising.

Interactions of these compounds with AcrB have been investigated by mean of X-ray crystallography as well as in silico molecular dynamics simulations [47]. Both approaches demonstrated that MBX EPIs can bind the DBP, forming interactions similar to those described for PAβN and NMP. Similarly, indeed, MBX EPIs interact stably with the switch loop or nearby residues, as well as with the hydrophobic trap. Their mode of action could thus be analogous to that of the formerly described EPIs. Differences in the required concentration of the MBX EPIs with respect to PAβN and NMP may be related to different binding affinities [47]. Indeed, in silico estimations of the binding free energies of these EPIs provided considerably higher values for the MBX compounds (roughly -50 kcal/ml [47]) than for PAβN and NMP (roughly -20 kcal/mol in both cases [60]).

Recently, important insights on the mode of action of the MBX EPIs have come from structural investigations, through cryo-EM [37] and cryo-ET [39] techniques. Cryo-EM data on AcrAB-TolC [37], sampled in different conditions, have indeed revealed that AcrB tends to adopt the TTT conformation in presence of the MBX3132 EPI. Although the EPI has not been solved in the structure, such stabilization has not been detected in the apo AcrB or in presence of substrates [37]. This finding is in good agreement with cryo-ET data [39] (see Subsection 1.3.3), which showed that, in presence of the same EPI, the AcrA-TolC channel stably adopts an open conformation. On the basis of structural data obtained in the same study, this channel has instead been proposed to open only transiently in presence of substrates. These data, therefore, suggest that MBX3132 may inhibit the pump by strongly restraining the dynamics of the transporter, which cannot complete its transport cycle. The stabilization of AcrB, in turn, could prevent functional motions of AcrA and TolC, which are blocked in the open conformation [39]. Further investigations are required to fully understand this mechanism. Indeed, important aspects need to be explained, such as the stabilization of the TTT state by MBX3132. Moreover, similar studies on the remaining EPIs could be conducted, to understand whether they share a similar mode of action and analogous restraining properties.

1.6 Thesis rationale

This thesis focuses on two relevant lines of research on EPIs, *i.e.* the investigation of the inhibitory mechanism of known inhibitors as well procedures for the identification of novel EPIs that are suitable for clinical use. Specifically, we considered the following compounds:

- the in-use antipsychotics amitriptyline and chlorpromazine, which were proven to be

effective against AcrAB-TolC in *E. coli* and *Salmonella* Typhimurium. These compounds are therefore potential candidates for the synthesis of new EPIs through drug repurposing. Their utilisation, however, requires the understanding of their inhibitory mechanism;

- the known EPI PAβN, whose effectiveness against AcrAB-TolC has been demonstrated. Nonetheless, several questions are still open with regard to its functioning.

In this thesis, the inhibition mechanisms of such compounds have been investigated through computational methods. Moreover, the obtained results have been integrated with experimental data as part of multidisciplinary collaborations.

1.7 Thesis outline

After this introduction, the thesis is organized as follows:

Chapter 2: description of the theoretical basis of the computational methods applied in this thesis;

Chapter 3: study of the inhibition mechanism of amitriptyline and chlorpromazine against *E. coli* and *Salmonella* Typhimurium AcrB;

Chapter 4: study of the inhibition mechanism of PAβN against *E. coli* AcrB (wt and fluoroquinolone resistant G288D variant). Ternary complexes formed by AcrB bound to PAβN and the antibiotic ciprofloxacin were also considered, in order to assess potential variations in the action of the inhibitor due to the co-presence of a substrate;

Chapter 5: *in silico* structural characterization of *Salmonella* Typhimurium AcrB (wt and fluoroquinolone resistant G288D variant), aimed at evaluating the impact of the G288D mutation on the structure and hydration of the binding pockets.

Chapter 2

Methods

2.1 Introduction

The *in silico* studies presented in this thesis have been performed in complementarity to diverse experimental approaches, including well diffusion essays (Chapter 3), hydrogen-deuterium exchange-mass spectrometry (HDX-MS, Chapter 4) and cryo-EM (Chapter 5). In general, experimental techniques have provided fundamental contributions in the investigation of sub-cellular processes [65]. Specifically, our understanding of the structure and functioning of AcrB and analogous transporters is greatly due to techniques like X-ray crystallography, cryo-EM and susceptibility tests (see Chapter 1, Section 1.4) [31].

Despite their fundamental contributions, these techniques are often not suitable for the investigation of the dynamical aspects of protein functioning, which are crucial for the understanding of most cellular phenomena. Indeed, only a few techniques like NMR or fluorescence spectroscopy can be applied for the investigation of structural dynamics. However, these methods can require difficult interpretation, or do not allow the observation of unstable states [66]. Other techniques like HDX-MS provide insights on changes in the protein dynamics upon specific events, like substrate binding (see Chapter 5). Nonetheless, information on the binding event or triggering mechanisms in the protein are lost [66].

For many sub-cellular systems, like the AcrB transporter, an atomic-level description of functional dynamics can be achieved through *in silico* techniques, like molecular modeling or molecular dynamics simulations [66]. These techniques have indeed proven to be useful for the investigation of various phenomena, including binding events, conformational changes and substrate recognition [34].

In the works presented in this thesis, we mainly focused on the binding of substrates and/or inhibitors to AcrB and on how such interactions affect the structural features of the transporter. We made use of homology modeling to build structural models of *S.* Typhimurium AcrB and of a fluoroquinolone resistant AcrB variant, whose high-resolution structure is currently not available. Putative binding sites of substrates and inhibitors were identified through molecular docking. Moreover, in all work we conducted all-atom MD simulations to investigate the dynamical properties of the system of interest. The theoretical background of the mentioned computational techniques is discussed in Section 2.2, while Section 2.3 is dedicated to analysis methods.

2.2 Theoretical background

2.2.1 Homology modelling

In proteins, structural features are strongly related to biological function [67]. Knowledge of the three dimensional structure of such systems is thus crucial.

Since the second half of the 20^{th} century, several experimental techniques have been applied to the resolution of protein structures [68]. These include X-ray crystallography, NMR and electron microscopy [65]. Studies based on such techniques have lead to the charachterization of a high number of proteins: over 165,000 three dimensional structures are currently available in the PDB database [69]. Application of experimental methods, however, is not always successful. Especially in the case of membrane proteins, like transporters, considerable difficulties may be encountered in the purification process [68].

In parallel to the development of experimental methods, diverse computational approaches have been proposed to predict the structure of proteins [70]. Among them, homology modeling is considered as the most accurate [68]. This technique is based on the empirical observation that proteins with similar amino acid sequence tend to fold into similar structures [71]. Its application requires at least one protein whose structure has been experimentally resolved (template) [68]. Prediction of realistic structural models requires a significant sequence similarity between the template and the protein of unknown structure (target) [72]. Typically, false negatives or inaccurate predictions are likely to be obtained with a sequence similarity below 10% [72, 73]. In this work, we selected *E. coli* AcrB (PDB IDs 2J8S [112], 4DX5 and 4DX7 [45]) as template for the modeling of the *S.* Typhimurium orthologue (94.7% sequence identity, see Chapter 5) and its fluoroquinolone resistant variant. Cryo-EM data were used for a further refinement of the structural model of the fluoroquinolone resistant variant (Chapter 5). Sequences of the target and template proteins were derived from the Uniprot database [74]. Models were generated through the dedicated software Modeller [75], whose protocol is described in the following (see Fig. 2.1 for a schematic representation).

In Modeller [75], the first step for structural prediction consists in the optimal alignment of the target and template sequences, aimed at minimizing the number of sequence gaps. Once this procedure has been completed, structural restraints necessary for model building are defined. These include:

- spatial restraints, such as the correlation between equivalent $C\alpha - C\alpha$ distances. Such restraints are defined through probability density functions whose expression has been obtained empirically, from the statistical data analysis of similar protein structures;
- stereochemical features, such as bond lenghts, bond angles and dihedral angles, which are defined by the CHARMM22 force field [77].

The mentioned restraints are combined in the so-called objective function. The structural model is obtained by optimizing such function in Cartesian space. This is achieved through conjugate gradient methods and molecular dynamics with simulated annealing.

In addition to this protocol, additional steps were performed through the Flex-EM software [78] for the refinement of our model of the fluoroquinolone resistant variant in its cryo-EM map. Firstly, the three-dimensional structure generated by Modeller [75] is fitted into the map. The protein is treated as a rigid body, its position and orientation in the map being optimized through Monte Carlo and conjugate gradient methods. Secondly, the model undergoes an MD

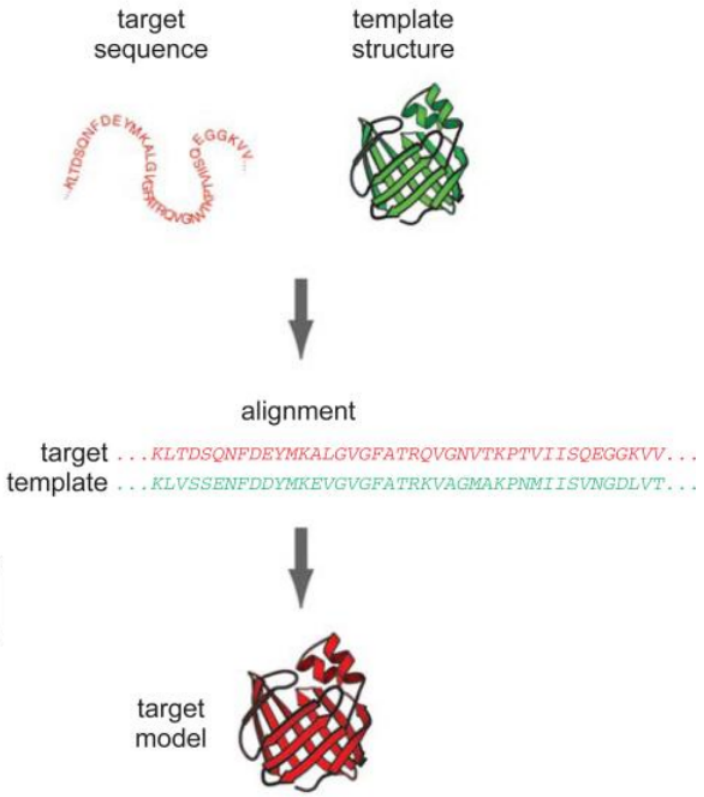

Figure 2.1: Schematic representation of the structure prediction steps in Modeller. Image adapted from [76].

simulated annealing procedure. During this procedure, the cross-correlation function (CCF) of the model is computed:

$$CCF = \frac{\sum_{i=0}^{Vox} \rho^{EM} \left(\sum_{j=1}^{N} \rho_{ij}^{P} \right)}{\sqrt{\sum_{i=0}^{Vox} (\rho^{EM})^2 \left(\sum_{j=1}^{N} \rho_{ij}^{P} \right)^2}} \tag{2.1}$$

where ρ^{EM} is the density of the cryo-EM map, ρ_{ij}^{P} is the density of grid points of the structural model and Vox is the number of voxels of the cryo-EM map located within two times the map resolution from any atom of the structural model. Values of the CFF range between 0 and 1, with 1 indicating an optimal fit in the cryo-EM map. The optimization through simulated annealing is terminated if the change in the CCF is below 0.001.

2.2.2 Molecular dynamics

Molecular dynamics (MD) is a powerful computational technique for the description of microscopic systems [79]. This approach is widely used for the study of biological systems [79], *e.g.* to evaluate the stability of structural models of biomolecules (see Subsection 2.2.1) or to investigate processes such as ligand binding and conformational changes in proteins [80, 81]. In the case of AcrB and omologous RND transporters, MD simulations have provided significant insights on their transport mechanism and on the substrate specificity of the DBP [40].

Several MD approaches have been developed over the years, which include *ab initio* MD (which accounts for quantum effects) as well as classical MD (in which atoms are treated as classical particles, requiring less computational resources) [82]. Although quantum effects are relevant for the study of specific biomolecules, such as metalloproteins [83], they can safely be neglected in most biological systems [82]. For this reason, their simulations are typically conducted through the classical MD approach [79].

2.2.2.1 Classical description of molecular systems

A fundamental assumption at the basis of this approach is the Born-Oppenheimer approximation [82], regarding the decoupling of nuclear and electronic motions. In molecular systems, indeed, nuclear motions are typically much slower than electronic motions. Therefore, electronic relaxation upon a nuclear displacement can be considered as instantaneous. It is thus possible to treat nuclear motions separately, and to compute the electronic energies for a given set of nuclear coordinates. A proper evaluation of the electronic energy would require the resolution of the Schroedinger equation, thus accounting for quantum effects. In the classical MD approach, in which such effects are assumed to be negligible, electronic interactions are expressed through an effective potential, written as a function of the coordinates of the nuclei. At present, several forms of the effective potential have been derived, mainly through semi-empirical methods [84].

On the basis of these approximations, nuclear dynamics can be described through the Newton's equation of motion [82, 85]. Specifically, indicating as m_i and r_i the mass and position of particle (nucleus) i in the system, Newton's equation can be written as:

$$m_i \frac{\partial^2 r_i}{\partial t^2} = \sum_{j=1}^{N} F_i(t) \tag{2.2}$$

where N is the number of atoms in the system and $F_i(t)$ is the resultant force acting on particle i at time t. This term accunts for the interactions with the remaining particles of the system. It is related to the effective potential, indicated as $U_i(t)$, by the equation:

$$F_i(t) = -\nabla U_i(t). \tag{2.3}$$

As previously mentioned, several expressions for the effective potential $U_i(t)$ (referred to as force fields) have currently been derived [84]. Among the most widely used force fields are CHARMM (used for homology modelling, see Section 2.1) [77] and the AMBER [86] force fields, which have been used for the MD simulations described in this thesis. Although specific parameters and functional forms can differ depending on the considered force field, some general aspects regarding the description of the interatomic interactions can be identified. Typically, force field are written as the combination of two terms [85], which account for the interactions between atoms connected by chemical bonds (U^{bonded}) and unconnected atoms ($U^{unbonded}$):

$$U = U^{bonded} + U^{unbonded}. \tag{2.4}$$

In most force fields, the term U^{bonded} is written in the following form:

$$U^{bonded} = U^{bonds} + U^{angles} + U^{torsions} \tag{2.5}$$

where U^{bonds} and U^{angles} account for variations in bond lenghts and bond angles, respectively, while $U^{torsions}$ is related to torsions around chemical bonds, described by dihedral angles (see Fig. 2.2 for a schematic representation). Possible expressions for these terms are:

$$U^{bonds} = \sum_{bonds} k_r(b - b_{eq})^2 \tag{2.6}$$

$$U^{angles} = \sum_{angles} k_\vartheta(\vartheta - \vartheta_{eq})^2 \tag{2.7}$$

$$U^{torsions} = \sum_{diehdrals} \frac{V}{2}[1 + cos(n\varphi - \gamma)]. \tag{2.8}$$

In Equations 2.6 and 2.7, variations in bond lenghts and angles are treated as harmonic oscillations around the equilibrium values b_{eq} and ϑ_{eq}. In the case of bond lenghts, a more accurate description of the associated energy is given by the Morse potential (Fig. 2.3). Eq. 2.6, however, provides a good approximation for small oscillations with respect to the equilibrium value, and has a much lower computational cost. With regard to the expression of $U^{torsions}$, dihedral angles are indicated as φ, while γ is a phase angle and n is the number of minima.

Similarly to U^{bonded}, term $U^{unbonded}$ (related to unbonded interactions) can be written as the sum of two contributions. These are given by electrostatic ($U^{electrostatic}$) and Van der Waals interactions (U^{VdW}):

$$U^{unbonded} = U^{electrostatic} + U^{VdW}. \tag{2.9}$$

The energy associated to electrostatic interactions, which involve charged particles, is given by the following expression:

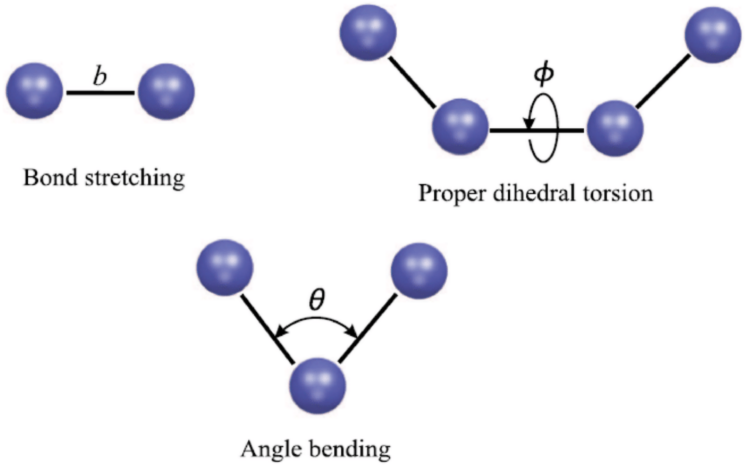

Figure 2.2: Representation of common bonded terms in force fields. Image adapted from [87].

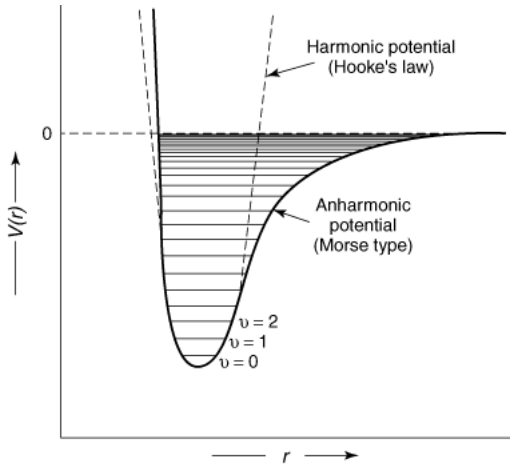

Figure 2.3: Behaviour of the Morse and harmonic potential with respect to the bond length. The harmonic potential represents a good approximation of the Morse potential in proximity of the equilibrium bond length (absolute minimum of the curve). Image from [88].

$$U^{electrostatic} = \sum_{i<j} \frac{q_i q_j}{4\pi\varepsilon r_{ij}} \tag{2.10}$$

where q_i and q_j are the charges of the considered particles, r_{ij} is the distance between them (notice that such distance does not necessarily correspond to a bond lenght, as in eq. 2.6) and ε is the dielectric constant.

Interactions between transient dipoles, due to oscillations in the charge distribution in neutral particles, can be described through the Lennard-Jones potential:

$$U^{VdW} = \sum_{i<j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \tag{2.11}$$

where ε_{ij} is the depth of the potential well, σ_{ij} is the zero-potential distance between the considered particles and r_{ij} is the distance between them, as in eq. 2.10. U^{VdW} is thus given by the contribution of repulsive short-distance interactions (term $(\sigma_{ij}/r_{ij})^{12}$ in eq. 2.11), which become dominant in case of atomic clashes, and attractive long-distance interactions (term $(\sigma_{ij}/r_{ij})^{6}$).

2.2.2.2 Periodic boundary conditions

In addition to the evaluation of interatomic interactions, a relevant aspect of MD simulations is related to the size of the simulated system [82, 85]. The system of interest is indeed enclosed in a box (typically a parallelepiped), in which edge effects cannot be neglected. In order to overcome this problem, periodic boundary conditions are used [82, 85]. With this method, identical copies (or images) of the simulation box are placed in each direction of a three-dimensional grid. In this way an infinite system is created, whose evolution is not affected by artifacts in proximity of the boundaries. Importantly, the use of periodic boundary conditions does not affect the conservation of the number of particles in the system. During the simulation, indeed, all periodic images evolve in the same way. Therefore, some atoms leave box i to enter its adjacent image (box $i+1$), the same number of equivalent atoms will enter box i from box $i-1$, and so on (see Fig. 2.4). The total number of atoms in each box is thus conserved along the whole simulation.

A possible issue in the application of periodic boundary conditions is the interaction between adjacent images through long-range forces, such as electrostatic ones, which cause considerable artifacts in the simulation. In principle, such interactions could become negligible provided the box is sufficiently large. However, an increase in the size of the system necessarily increments the computational and time cost of the simulation. In order to efficiently compute long-range interactions, several approaches have been developed. Among them, the particle mesh Ewald (PME) is widely used [85]. This approach is based on the idea of splitting the electrostatic potential (Eq. 2.10) in two terms, using the following identity:

$$\frac{1}{r} = \frac{f(r)}{r} - \frac{1-f(r)}{r} \tag{2.12}$$

where $f(r)$ is a generic function of the interatomic distance r. The efficiency of this method strongly depends on the choice of such function. The first term of the equation should indeed be negligible for large values of r (typically, a cutoff value between 9 and 10 Å is used), while the second term should vary slowly, to avoid divergences at large interatomic distances.

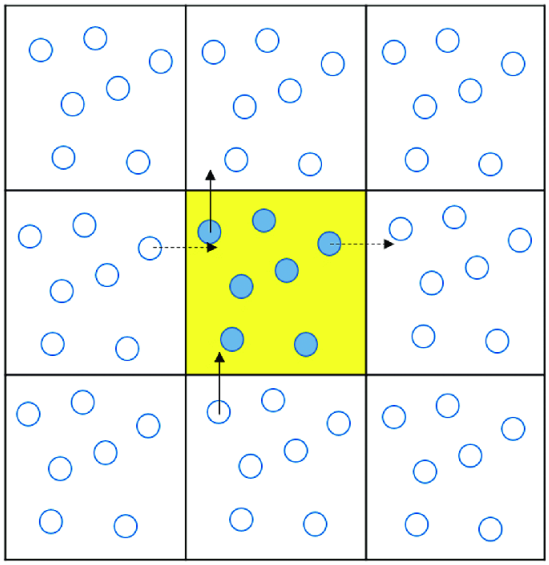

Figure 2.4: Periodic boundary conditions for a two-dimensional system. The number of particle in each periodic image is conserved along the whole simulation. Image from [89].

Several forms of $f(r)$ that satify these requirements have been identified, such as the Gaussian screening function:

$$\rho(r) = -q\left(\frac{\alpha}{\pi}\right)^{3/2}\exp(-\alpha r^2) \tag{2.13}$$

where q is the screened electrical charge and α is a parameter related to the width of the Gaussian.

2.2.2.3 Microscopic description and macroscopic observables: the ergodic hypothesis

Our knowledge of physical systems is often based on macroscopic observables. The description of a system at atomic level, such as that provided by MD simulations, is instead related on the evaluation of the position and momentum (r_i and p_i, respectively) of each particle of the system. Specifically, a given set of values for the variables (r_i, p_i) corresponds to a specific configuration of the system, referred to as microstate in statistical mechanics. The collection of all the possible microstates of the system of interest is named ensemble.

Macroscopic observables, which can be measured in experiments, are given by an average over the system ensemble:

$$< A >= \frac{\iint dr_i dp_i A(r_i, p_i)\exp(-H(r_i, p_i)/k_B T)}{\iint dr_i dp_i \exp(-H(r_i, p_i)/k_B T)} \tag{2.14}$$

where $A(r_i, p_i)$ is the observable of interest, $H(r_i, p_i)$ is the Hamiltonian of the system, T is the temperature and k_B is the Boltzmann constant. The integral, as mentioned, is calculated over all the possible microstates of the system.

Since the evaluation of such integrals is very demanding, a possible approach is that of generating all the possible microstates of the system, which could be used to compute the average

of the observable of interest. However, this requires the exploration of the whole ensamble, which can have a very high computational cost. Indeed, it is not possible to guarantee that all microstates have been generated, even with multiple MD simulations of the same system.

The accurate measurement of macroscopic observables through MD simulations relies on the satisfaction of the ergodic hypothesis. According to it, a system free to evolve in an indefinite amount of time will explore all its possible microstates. Under this condition, therefore, the ensemble average of a given observable (Eq. 2.14) is equivalent to its time average. This can be written as:

$$< A >= \lim_{T\to\infty} \frac{1}{T} \int dt A(r_i(t), p_i(t)). \tag{2.15}$$

Provided MD simulations are sufficiently long, therefore, macroscopical observables can be calculated through averages over the simulation time steps. Clearly, such procedure can lead to accurate results only if the simulation time is sufficient to explore the representative states of the system.

2.2.2.4 MD simulation protocol

In each work presented in this thesis, MD simulations have been performed using the dedicated software AMBER. The following protocol has been followed:

1. The topology and the initial coordinate files of the system have been created using the LeAP module of AMBER18 [86]. Force fields ff14SB [114] and protein.fb15 [137] have been used to represent the protein, while the TIP3P model has been used for water [115]. When present, lipids have been represented through the lipid17 force field (*http://ambermd.org/GetAmber.php*), while GAFF parameters [117] for generic compounds were taken from Malloci *et al.* [90].

2. The system has been subjected to a multi-step structural relaxation using the pmemd module of AMBER18. Relaxation has been achieved by gradually releasing positional restraints on the system.

3. The system has been heated from 0 to 310 K. The heating procedure differs according to the considered work. In studies in which a truncated model of AcrB has been simulated, which only included the periplasmic portion (see Chapters 3 and 5), heating has been performed in 1.25 ns under constant pressure (set to 1 atm), with positional restraints on the Cα within 5 Å from the bottom of the protein. In investigations on systems composed by the complete AcrB protein embedded in a membrane (see Chapter 4), heating has instead been performed in two steps, to further stabilize the system: (i) from 0 to 100 in 1 ns under constant volume conditions, (ii) from 100 to 310 K in 5 ns under constant pressure, set to 1 atm. Positional restraints have been imposed on the heavy atoms of AcrB and on the phosphorous atoms of lipids to allow the membrane to relax.

4. A short MD simulation has been performed to equilibrate the system. Equilibration was conducted under isotropic pressure scaling and at constant temperature. In the case of the truncated AcrB protein, it has been performed for 10 ns, while for simulations of the whole AcrB protein 20 equilibration steps have been performed, each of 500 ps in duration (10 ns in total).

5. NPT MD simulations have been performed for each system. The truncated AcrB protein, whose investigations were aimed at evaluating interactions with substrates and/or inhibitors (see Chapter 3) and assessing structural properties (see Chapter 5), has been simulated for 150 ns. The system composed by AcrB embedded in a membrane has instead been simulated for 1 μs, in order to evaluate its hydration properties and structural flexibility (see Chapter 4). For all simulations, a time step of 4 fs was used under hydrogen mass repartitioning [98]. The PME algorithmwas used to evaluate electrostatic interactions with a distance cutoff of 9 Å.

2.2.3 Molecular docking

Molecular docking is a widely used computational method aimed at predicting the structure of a complex (typically composed by a small ligand and a protein, or receptor) starting from its unbound components [91]. It is based on the generation of a very high number of configurations for the complex of interest, named poses. For each pose, the scoring function is evaluated, which consists in the combination of several terms (electrostatic and Van der Waals interactions, buried surface, etc) and provides a rough estimation of the binding affinity. In principle, therefore, poses with the highest values of the scoring function should represent the native structure of the complex, *i.e.* its structure in physiological conditions. The accuracy of the results are thus strongly related to the accuracy of the scoring function [91, 93].

Importantly, the docking approach is very different from that of MD simulations [92]. Indeed, while MD is aimed at reproducing the time evolution of a given system, molecular docking has been developed to explore and rank the possible (meta)stable conformations assumed by two or more interacting biomolecules. Docking, therefore, provides a collection of static snapshots representing putative structures of the bound complex. Such snapshots may then by used as starting structures for further MD simulations, in order to evaluate the stability of the complex and its time evolution. In principle, bound conformations of the complex could also be generated through molecular dynamics simulations, by placing the ligand and the receptor in the same box and waiting for them to bind. However, such procedure would be much more time consuming, expecially in cases where the binding site of the ligand is not known (blind docking). Docking algorithms, on the contrary, have been spefically developed to generate a large number of possible configurations at low computational costs.

A relevant aspect of molecular docking is related to the structural flexibility of both the ligand and the receptor [92]. By accounting for the full flexibility of both components, indeed, docking algorithms could provide more accurate results. For example, they could mimic small rearrangements in the binding site due to interactions with the ligand. On the other hand, evaluations of this kind would cause considerable increments in time and computational cost. For this reason, both the receptor and the ligand are often treated as semi-rigid bodies. Most docking softwares indeed allow to define specific rotamers, *i.e.* chemical bonds around which torsions are allowed. An alternative approach is that of ensemble docking, in which docking is not performed using a single three-dimensional structure of the receptor and/or the ligand, but rather a collection of structures representing different conformations. This approach allows to partially take into account the flexibility of the involved molecules. This can significally increase the accuracy of the algorithm predictions.

In the works presented in this thesis, docking calculations have been performed with the dedicated software Autodock VINA [94], which is among the most used in this field. It is based on the use of an empirical scoring function, and performs an iterated optimization of

the generated poses that significantly improves the accuracy of the predictions. In all the performed docking calculations, ligands have been treated as semi-rigid bodies through the definition of rotamers, while structural ensembles have been provided for the AcrB protein, used as receptor. In the case of the *E. coli* wild type AcrB, the structural ensemble was composed by high resolution X-ray crystal structures, while ensembles of *S.* Typhimurium AcrB and the fluoroquinolone resistant variant were composed by homology models (see Chapters 3 and 4; see also Subsection 2.2.1).

2.3 Methods for the analysis of MD trajectories

In the works presented in this thesis, several techniques have been applied for the analysis of MD trajectories. The following Subsection is dedicated to their description.

2.3.1 Root Mean Square Deviation (RMSD)

The RMSD analysis is the most common method to evaluate the similarity between three-dimensional structures. It can be used to compare crystal structures of the same protein or, in the case of MD simulations, protein structures extracted from different frames of the trajectory.

Typically, its evaluation is performed upon the alignment of the structures. In this phase, the superposition of the coordinates of equivalent atoms is optimized with dedicated algorithms. The RMSD is then calculated through the following expression:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} d_i^2} \tag{2.16}$$

where N is the number of equivalent atoms and d_i is the distance between the i-th couple of equivalent atoms upon the structural alignment. This calculation is often performed on a subset of atoms in the protein. A frequent choice is that of the Cα atoms, whose RMSD allows to assess structural variations of the backbone.

In the analysis of MD trajectories, RMSD calculations are typically performed using the first frame (*i.e.* the initial configuration of the system) as reference. For each of the remaining frames, the protein structure is extracted and aligned to the reference one, and the RMSD is calculated. This allows to evaluate if the protein conformation has significally changed along the simulation. Specifically, the presence of marked variations in the RMSD is indicative of the fact that the simulation has not reached the convergence.

2.3.2 Root Mean Square Fluctuation (RMSF)

The RMSF analysis provides a quantitative measurement of the flexibility of a molecule. It is defined through the expression:

$$RMSF = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(x_i(t) - x_i^{ref})^2} \tag{2.17}$$

where T is the time length of the MD trajectory, $x_i^{ref}$is the position of atom i in the reference frame (for example the initial configuration) and $x_i(t)$ is the position of the same atom

at time t. Typically, such calculation is performed for each residue of the simulated protein (per-residue RMSF) in order to detect its most flexible regions. In the present thesis, such calculation has been used to detect variations in the flexibility of AcrB under different conditions, related to interactions with the known inhibitor PAβN and the substrate ciprofloxacin (see Chapter 1). The obtained results have been compared with watershell analysis and experimental data from HDX-MS, which can assess variations in the protein hydration and flexibility (see Chapter 4).

2.3.3 Watershell analysis

Watershell analysis consists in the evaluation of the number of water molecules within the first and second watershells (within 3.4 and 5.0 Å, respectively) of the protein. In this thesis, this calculation has been performed using the *cpptraj* module of AMBER18 [86]. In order to compare the obtained results with experimental data from HDX-MS (see Chapter 4), the calculation has been restricted to the backbone nitrogen atoms. Analogous calculations have been performed on residue 288 of AcrB, in order to compare its hydration levels in the wt protein and in the fluoroquinolone resistant G288D variant (see Chapter 5).

2.3.4 Cluster analysis

Cluster analysis is a data mining technique widely used for the analysis of MD trajectories. It consists in the classification of a collection of objects in different groups (clusters) on the basis of specific similarity criteria, named metric. In a proper classification, objects belonging to the same cluster are more similar to each other (according to the selected metric) than to objects belonging to other clusters. In MD simulations of biomolecules, this method is often applied to detect the most representative conformations of the system of interest. In such cases, the RMSD can be used as metric to evaluate structural similarities.

At present, several algorithms for cluster analysis have been created. In this thesis, the average-linkage hierarchical agglomerative method has been used, which has proven to be among the most useful for the analysis of MD trajectories. With this technique, the distance between two clusters is defined as the average of the distances between individual points of the two clusters.

In MD simulations of AcrB in the presence of substrates and/or inhibitors (see Chapters 3 and 4), cluster analysis has been performed on ligand trajectories using the distance RMSD (dRMSD, *i.e.* the RMSD of all pairs of internal distances) as metric. The cluster representatives were used to detect the representative binding modes of each ligand. With regard to the characterization of the structural properties of AcrB (see Chapter 5), three clustering procedures have been performed, in each of which the dRMSD has been applied to the DBP of a different monomer. The obtained results were used to perform several analyses on the structure of the DBP, including the estimation of its volume and radius of gyration.

2.3.5 Measurement of the volume of the binding pocket

Measurement of the volume of binding pockets is of considerable importance for the study of conformational changes in receptors, often related to their biological functionality. For this reason, several algorithms for volume measurement have been created. In the present thesis, the dedicated software POVME [95] has been used to measure the volume of the DBP in AcrB, in order to characterize its structure and the effects of the G228D mutation (associated

to fluoroquinolone resistance, see Chapter 1) on the pocket geometry. POVME was used for this analysis due to the characteristics of its algorithm, which can efficiently be used to evaluate the volume of large and flexible pockets such as the DBP.

In POVME, the binding pocket of interest must be enclosed by the user in an inclusion region. Volume can be added or removed from the inclusion region through inclusion and exclusion spheres, respectively. The inclusion region is then filled with equidistant points, spaced 1 Å apart by default (in this thesis, a grid spacing of 0.5 Å was used). All points found within a cutoff distance of the protein atoms are deleted. Such cutoff distance is given by the sum of the Van der Waals radius of the protein atom and the Van der Waals radius of an hydrogen atom (1.09 Å). The volume of the pocket is thus measured by evaluating the number of remaining points. Such procedure can be applied to a single three-dimensional structure as well as to MD trajectories. In the latter case, a single inclusion region can be defined for the whole trajectory, provided the shape of the binding pocket does not change considerably. In this thesis, volume calculations were performed on the most populated clusters of each trajectory. The clusters have been obtained by using the dRMSD of the DBP as metric (see Subsection 2.3.4).

2.3.6 Radius of gyration

Together with the volume measurement, the calculation of the radius of gyration has been applied for the structural characterization of the DBP in AcrB (see Chapter 5). For a molecule composed by N atoms, this parameter is defined as:

$$R_g = \sqrt{\sum_{i=1}^{N} r_i^2} \tag{2.18}$$

where r_i is the distance of the i-th atom from the center of mass of the molecule. The radius of gyration, therefore, provides a measurement of the compactness of the considered structure. It can also be calculated for specific regions of a molecule, such as a protein domain or a binding site. In this thesis, we calculated the radius of gyration for three different regions of the DBP, in order to better characterize its geometry. These regions are the whole DBP, the hydrophobic trap and the upper part of the binding site. A list of the residues included in each region, together with the obtained results, is provided in Chapter 5.

2.3.7 Free energy of binding

In MD simulations of proteins in presence of their substrates or inhibitors, an important aspect to be evaluated is the the binding affinity, which is the propensity of the ligand to bind the protein (or receptor). In such cases, the binding strength can be evaluated through the variation in free energy associated to the reaction:

$$L + R \rightleftharpoons LR \tag{2.19}$$

where L and R are the unbound ligand and unbound receptor, respectively, while LR is the binary ligand-receptor complex.

Several methods have been developed for the calculation of the binding free energy from MD trajectories. Among the most popular are the molecular mechanics with Poisson-Boltzmann or generalized Born surface area (MM/PBSA and MM/GBSA, respectively) [96]. These methods

calculate the difference in free energy between the bound (LR) and unbound ($L + R$) states in Eq. 2.19. Such calculation is performed by decomposing the free energy (ΔG_b) in three terms:

$$\Delta G_b = \Delta E_{MM} + \Delta G_{solv} - T\Delta S \tag{2.20}$$

where ΔE_{MM} is the difference in energy related to electrostatic, Van der Waals and bonded interactions, ΔG_{solv} is the solvation free energy, T is the absolute temperature of the system and ΔS is the variation in conformational entropy. The term ΔE_{MM} is thus evaluated through the force field (typically without any cutoff for the electrostatic and Van der Waals interactions). Term ΔG_{solv} is instead given by the sum of polar and non-polar contributions:

$$\Delta G_{solv} = \Delta G^{p}_{solv} + \Delta G^{np}_{solv}. \tag{2.21}$$

Polar contributions are obtained by solving the Poisson-Boltzmann equation (this can be achieved using the Generalized Born method, giving the MM/GBSA approach), while non-polar contributions are related to the variation of the solvent accessible area. Finally, the entropic term in Eq. 2.20 is given by rototranslational and vibrational contributions. The former are evaluated through classical statistical mechanics (Sakur-Tetrode), while vibration terms are estimated by normal-mode analysis.

In this thesis, free energy calculations have been performed for putative inhibitors and substrates of AcrB (see Chapter 3). The obtained results have been compared with those reported for other inhibitors in previous literature [97], which have been estimated through the same methods.

Chapter 3

In silico investigation on the inhibitory mechanisms of first-generation antipsychotics against the AcrB transporter

3.1 Introduction

Efflux pumps of the RND superfamily play a key role for the resistance of gram-negative pathogens [21, 24]. As mentioned in Chapter 1, development of inhibitors is a viable route to contrast their action [58]. At present, several compounds with inhibitory properties have been identified. None of them is however suitable for clinical use [64], mainly due adverse effects and cytotoxicity (see Chapter 1, Section 1.5).

A possible alternative to *de novo* design of inhibitors is the repurposing of marketed drugs that are not in use for the treatment of infectious diseases [58]. Therapeuticals with inhibitory properties could indeed be used as a starting point for the design of new EPIs, *e.g.* through targeted modifications of their functional groups. Moreover, since their toxicity and pharmacokynetics have already been described, the time and cost required by drug development could be significantly reduced.

Among the drugs considered for repurposing, there is evidence that the first-generation antipsychotic medications chlorpromazine and amitriptyline behave as EPIs [99, 100]. Chlorpromazine has also been shown to possess antimicrobial activities [100, 101, 102]. While these activities occur at concentrations greater than those clinically achievable and/or desirable, chlorpromazine is able to potentiate the activities of many antibiotics at sub-inhibitory concentrations [100, 102, 103, 104] and increase the accumulation of ethidium bromide and other AcrB substrates [100, 105, 106]. However, the mechanism by which this occurs is not understood. Less is known about the efflux inhibitory effects of amitriptyline. However, like chlorpromazine, amitriptyline potentiates antibiotic activity [100]. In *Salmonella* Typhimurium, indeed, exposure to amitriptyline results in the induction of *ramA* [100], a gene associated with the lack of efflux [107]. On the other hand, hypersusceptibility to amitriptyline occurs when ramA is deleted [100].

Recently, experimental investigations [108] on the interactions between these compounds and AcrAB-TolC revealed that both amitriptyline and chlorpromazine are substrates of the

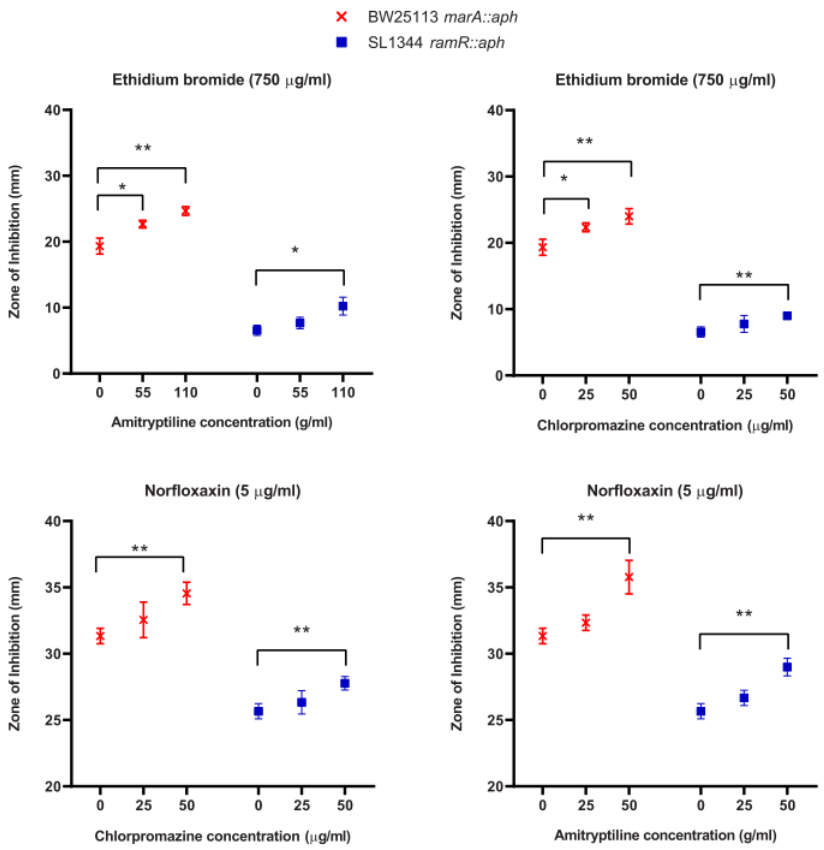

Figure 3.1: Comparisons of the zones of inhibition obtained for well diffusion assay with ethidium bromide and norfloxacin when used in combination with chlorpromazine and amitriptyline. *E. coli* strain: BW25113 *marR::aph*; S. Typhimurium strain: SL1344 *ramR::aph*. Image from [108].

pump. Chlorpromazine can indeed induce the upregulation of the AcrAB-TolC pump. Moreover, exposure to both chlorpromazine and amitriptyline resulted in the reversion of the non-functional AcrB D408A variant from *S.* Typhimurium [108]. This suggests that the two compounds are actively transported by the pump. Additional disk and well diffusion assays (Fig. 3.1) indicated that both chlorpromazine and amitriptyline are able to potentiate the activities of AcrB substrates, such as norfloxacin and ethidium bromide, against *S.* Typhimurium and *E. coli* [108].

These data shed further light on the inhibitory activity of chlorpromazine and amitriptyline, indicating that both compounds are involved in direct interactions with AcrAB-TolC. Knowledge of such interactions at atomic level is crucial for the proper characterization of these EPIs and for understanding their mode of action. Significant contributions in this sense can be provided by computational methods that, as mentioned in Chapter 2, can accurately describe dynamical processes like ligand binding [109]. In this light, we performed an *in silico* investigation on the binding of chlorpromazine and amitriptyline to AcrB in *E. coli* and *S.* Typhimurium. This work, which was mainly based on blind molecular docking and MD simulations (see Chapter 2 for a description of these techniques), is described in the following. The applied methods and the obtained results are reported in Section 3.2 and Section 3.3,

PDB ID	Resolution (Å)	Co-crystallized compound
2J8S [112]	2.5	-
4DX5 [45]	1.9	Minocycline
4DX7 [45]	2.3	Doxorubicin
5NC5 [37]	3.2	Puromycin
5EN5*	2.3	-
5ENO*	2.2	MBX2319 (inhibitor)
5ENP*	1.9	MBX2931 (inhibitor)
5ENQ*	1.8	MBX3132 (inhibitor)
5ENR*	2.3	MBX3135 (inhibitor)
5ENS*	2.8	Rhodamine-6G

*Truncated structures containing only the periplasmic domain of the protein [37].

Table 3.1: X-ray structures of *E. coli* AcrB used as structural templates to generate an ensemble of putative conformations of *S.* Typhimurium AcrB by homology modelling.

respectively. The results are further discussed in Section 3.4.

3.2 Methods

3.2.1 Homology modeling of AcrB from *Salmonella* Typhimurium

To perform ensemble docking calculations on *E. coli* and *S.* Typhimurium AcrB, several homology models of the latter were built using Modeller 9.21 [110] and several *E. coli* AcrB X-ray structures as the templates (see Table 3.1). Among these, the *E. coli* AcrB structures labeled 5ENx were truncated at the transmembrane (TM) region, and the protein assumed the LLT conformation. Therefore, we first generated their full structural models in the LTO conformation via homology modeling with multiple templates, as follows: chains A (in the L state) and C (in the O state) of the model were built using the corresponding chains of 4DX5 as the templates; chain B of the model was built using the TM of chain B of 4DX5 and chain C of the corresponding 5ENx structure (both in the T state) as the templates. For the modeling procedure, the amino acid sequences of the *E. coli* and *S.* Typhimurium AcrB transporters were first retrieved from the UniProt database (UniProt identifiers [IDs] P31224 and Q8ZRA7, respectively) [74]. The sequences were aligned using Clustal Omega [111] in order to determine the percentage of identical residues (94.7%) and verify the absence of gaps. Next, Modeller 9.21 was used to build the homology models. The variable target function method was used to perform the optimization and the models with the highest MOLPDF were used for molecular docking as described below.

3.2.2 Molecular docking

Blind ensemble docking calculations were performed for amitriptyline, chlorpromazine, ethidium bromide and norfloxacin on *E. coli* and *S.* Typhimurium AcrB structures using AutoDock VINA [94]. As we were interested in binding poses (preferred orientation of a ligand to a protein) in the periplasmic region of AcrB, docking was performed within a rectangular search space of size 125 Å by 125 Å by 110 Å enclosing that portion of the protein, as in reference 60. The exhaustiveness parameter was set to 8,192 (∼1,000 times the default 8) in order to

improve the sampling within the large box used (∼64 times the suggested volume of 30 Å by 30 Å by 30 Å). The flexibility of the receptor was considered indirectly by employing ensembles of conformations: 10 structures for each AcrB protein (*E. coli* and *S.* Typhimurium), while the flexibility of the ligands was considered by activating torsional angles in AutoDock VINA and using a starting structure that was optimized at the quantum-level of theory available at *www.dsf.unica.it/translocation/db* [90].

3.2.3 Molecular dynamics simulations

To select a tractable number of AcrB-ligand complexes on which to perform MD simulations, a cluster analysis was carried out on all the docking poses of each system, using the distance root mean square deviation (dRMSD) of the ligand as a metric to select their different orientations. The hierarchical agglomerative clustering algorithm implemented in the cpptraj module of the AMBER18 package [86] was used with a 3-Å dRMSD cutoff. Selected docking poses (namely, those featuring different orientations among the top ranked ones according to the AutoDock VINA scoring function) were subjected to all-atom MD simulations using the truncated model of AcrB [47, 60, 113], which includes only the periplasmic domain (residues 32 to 335 and 564 to 860 of each monomer). The AcrB-ligand complexes were inserted in a truncated octahedral box ensuring a minimum distance of 16 Å between the complex and the border of the box. The box was filled with a 0.15 M KCl aqueous solution. The topology and the initial coordinate files of the systems were created using the LEaP module of AMBER18. The AMBER force field ff14SB [114] was used to represent the protein systems; the TIP3P model was employed for water [115], and the parameters for the ions were obtained from reference [116]. The parameters of amitriptyline and chlorpromazine, obtained from the GAFF force field [117] or generated using the tools of the AMBER18 package are available at *www.dsf.unica.it/translocation/db* [90]. To improve the stability of the periplasmic region at the border with the TM domain, harmonic positional restraints (k=1 kcal mol^{1}Å^{2}) were imposed on Cα atoms of residues within 5 Å from the bottom region of the structure. Each system was first subjected to a multistep structural relaxation via a combination of steepest descent and conjugate gradient methods using the pmemd module of AMBER18, as described previously [113, 118]. The systems were then heated from 0 to 310 K in 1.25 ns under constant pressure (set to a value of 1 atm) and with restraints on the Cα atoms found within 5 Å from the bottom of the protein. Next, a 10-ns-long MD simulation was performed to equilibrate the box dimensions, applying to the system the same restraints used for the heating procedure. This equilibration step was carried out under isotropic pressure scaling using the Berendsen barostat, whereas an Andersen thermostat (with randomization of the velocities every 500 steps) was used to maintain a constant temperature. Finally, 150-ns-long production MD simulations were performed for each system. A time step of 4 fs was used during these runs, after the protein was subjected to hydrogen-mass repartitioning [119]; R-H bonds were constrained with the SHAKE algorithm. Coordinates were saved every 100 ps. The particle mesh Ewald algorithm was used to evaluate long-range electrostatic forces with a nonbonded cutoff 9 Å.

3.2.4 Postprocessing of MD trajectories

MD trajectories were analyzed using either in-house tcl and bash scripts or the cpptraj tool of AMBER18. Figures were prepared using gnuplot 5.0 [120] and VMD 1.9.3 [121].

(i) Cluster analysis. Clustering of the trajectories to select nonequivalent binding poses

Complex[a]	ΔG_{max}(kcal/mol)
AMI-$AcrB_{EC}$	-11.6
AMI-$AcrB_{ST}$	-12.1
CPZ-$AcrB_{EC}$	-9.2
CPZ-$AcrB_{ST}$	-9.3

[a]All corresponding poses are localized within the DBP_T. CPZ, chlorpromazine; AMI, amitriptyline.

Table 3.2: (Pseudo)binding free energies evaluated through the scoring function of AutoDock VINA for the top ranked poses of both amitriptyline and chlorpromazine on *Ec*AcrB and *S*TmAcrB.

of the ligands was carried out using the average-linkage hierarchical agglomerative method implemented in cpptraj and employing a dRMSD cutoff of 2.5 Å on all the nonhydrogenous atoms of the ligand.

(ii) Binding free energy calculations. The MM/GBSA approach [96] implemented in AMBER18 was used to calculate the solvation free energies following the same protocol used in previous studies [47, 60, 113, 122, 123]. This approach provides an intrinsically simple method for decomposing the free energy of binding into contributions from single atoms and residues [124]. The solute conformational entropy contribution ($T\Delta S$; see Chapter 2, Section 2.3) was not evaluated [86]. Calculations were performed on 50 different conformations of each complex, which were extracted from the most populated conformational cluster (representing the most sampled conformation of the complex along the production trajectories).

(iii) Ligand flexibilities. The root mean square fluctuations (RMSFs) of the ligands were calculated using cpptraj after structural alignment of each trajectory.

3.3 Results

3.3.1 Molecular docking

In order to investigate whether the mode of action of EPIs chlorpromazine and amitriptyline could be related with their interaction with AcrB, we firstly evaluated their propensity to bind the transporter in both *E. coli* and *S.* Typhimurium. This was assessed by means of a blind docking campaign, which resulted in over 200 poses per ligand. From the analyses of the putative binding poses in *E. coli* and *S.* Typhimurium AcrB (*Ec*AcrB and *S*TmAcrB, respectively), it emerges that chlorpromazine and amitriptyline display similar docking scores in the two systems (Table 3.2). Moreover, comparison of the distributions of the docking poses reveals a good overlap (Fig. 3.2). Significantly, in each system a large number of poses is located within the DBP of monomer T (hereafter DBP_T; see Chapter 1, Section 1.4 for a description of the AcrB translocation pathway), in tight interaction with residues of the hydrophobic trap (lined by phenylalanine residues F136, F178, F610, F615, and F628 in both *Ec*AcrB and *S*TmAcrB - see Table 3.3). This is known to be the preferred binding site of several EPIs effective against AcrB, such as PAβN, NMP and the EPIs of the MBX series (see Chapter 1, Section 1.5).

Moreover, to provide molecular-level insights on the possible mechanism by which chlorpromazine and amitriptyline interfere with the efflux of ethidium bromide and alter the intracellular accumulation of norfloxacin, we performed blind ensemble docking calculations of norfloxacin and ethidium bromide on both *Ec*AcrB and *S*TmAcrB. Importantly, the distri-

Region	Lining residues
Distal Binding Pocket (DBP)	46 89 128 130 134 136 139 176 177 178 179 180 273 274 276 277 327 573 610 612 615 617 620 628
Hydrophobic trap	136 178 610 615 628
CH3*	33 37 100 296 298
G-loop	616 617 618 619

* residue 296 was not included in the definition given in [51].

Table 3.3: Residues lining the regions of interest of AcrB. The same definitions can be used for *Ec*AcrB and *S*TmAcrB, due to the lack of gaps between their sequences.

Substrate	MD simulations	
	E. coli	*S.* Typhimurium
AMI	3	3
CPZ	3	3
NOR	3	3
EtBr	3	3

Table 3.4: MD simulations performed in this work. The starting configurations of each substrate were selected among the clusters of the docking poses localized within the DBP_T. AMI, amitriptyline; CPZ, chloropromazine; NOR, norfloxacin; EtBr, ethidium bromide.

butions of preferred putative binding sites of these AcrB substrates significantly overlapped those obtained for chlorpromazine and amitriptyline. Moreover, most of the highest affinity poses were localized within the DBP_T (Fig. 3.2).

3.3.2 Molecular dynamics simulations

Binding of the considered compounds within the DBP_T was further investigated through all-atom MD simulations and binding free energy calculations, in order to evaluate its thermodynamics and dynamical features. For each system, three initial structures for MD simulations were selected. This choice was performed through a cluster analysis on the docking binding poses within the DBP_T. In the case of amitriptyline and chlorpromazine, three clusters were obtained, which together grouped roughly 50% of all docking poses located in the DBP_T of *Ec*AcrB and *S*TmAcrB. In the case of norfloxacin and ethidium bromide, the same coverage was obtained through one cluster. The initial structures for MD simulations of amitriptyline and chlorpromazine were thus selected among the representatives of cluster 1 to 3 (sorted by population), while for norfloxacin and ethidium bromide three different binding poses were selected from the most populated cluster. In total, 24 MD simulations were performed (see Table 3.4), each having a duration of 150 ns. As reported in Subsection 3.2.3, all simulations were performed using a truncated structure of AcrB, which was validated in previous works [47, 60].

MD trajectories were firstly analyzed by evaluating the flexibility of the ligands and their displacement inside the protein (see Subsection 3.2.3). From this analysis, it emerged that all compounds could stably occupy the DBP_T, in both *Ec*AcrB and *S*TmAcrB. For the sake of clarity, only the most stable trajectories of each system (Fig. 3.3) will be discussed in the following.

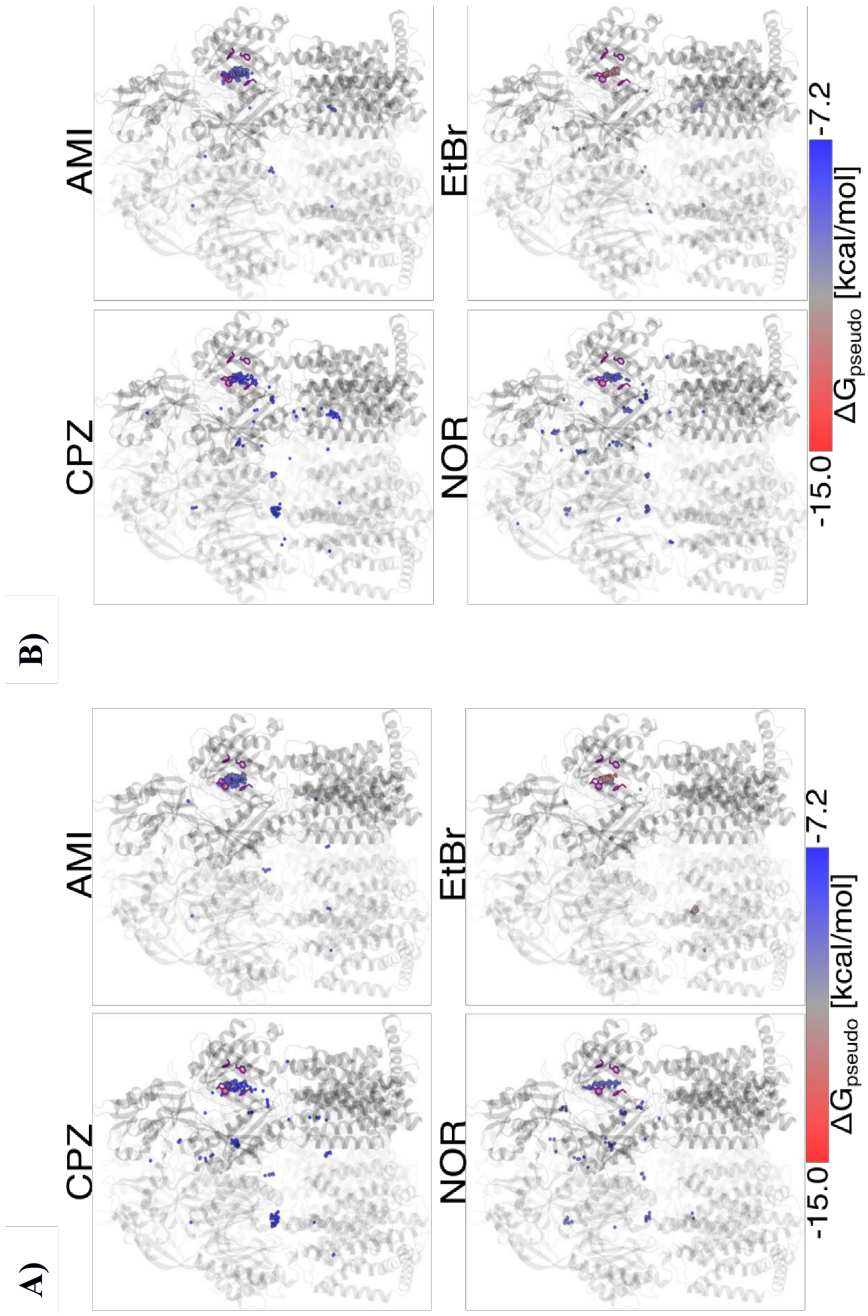

Figure 3.2: Distribution of the top docking poses obtained from blind ensemble docking calculations of chlorpromazine (CPZ), amitriptyline (AMI), norfloxacin (NOR), and ethidium bromide (EtBr) on $AcrB_{EC}$ (A) and on $AcrB_{ST}$ (B) (see Methods for details). The picture shows the distribution of the centers of mass of the poses, colored according to scoring function of AutoDock VINA (ΔG_{pseudo}). The protein is shown in transparent ribbons, with monomers L and T in the front, on the left and right side of the central intermonomer vestibule. The transparency increases going from monomer T to L to O. The sidechains of phenylalanines lining the hydrophobic trap of monomer T are shown as magenta sticks.

Evaluation of the binding region of amitriptyline and chlorpromazine revealed that both interact with several residues of the hydrophobic trap (Fig. 3.3). Specifically, such residues form 8 and 6 hydrophobic contacts with amitriptyline in *Ec*AcrB and *S*TmAcrB, respectively (a contact being counted when the minimum ligand-residue distance was less than 3.5 Å). Chlorpromazine is also involved in 10 hydrophobic contacts with residues of the same region, in both *Ec*AcrB and *S*TmAcrB. Moreover, the two ligands also form stabilizing interactions with segment 133-135 through their dimethylamine group (Table 3.5). In chlorpromazine, this group is also involved in cation-π interactions with residues of the hydrophobic trap of *S*TmAcrB, such as F136 and F617 (Table 3.5).

The differences in the interactions of the two ligands with AcrB are mirrored in the contributions of the DBP_T and the hydrophobic trap to the binding free energy. In *Ec*AcrB, indeed, contributions from both sites are higher for chlorpromazine than for amitriptyline (Table 3.6). In *S*TmAcrB the two ligands present similar contributions from the DBP_T, although a marked difference between contributions from the hydrophobic trap is retained (Table 3.6). In addition to the contributions from the regions of interest, the total binding free energies of the various compounds is reported in Table 3.6. However, this should be considered as an approximate estimate of the binding affinity. This is partly due to limitations of molecular mechanics in combination with the MM/GBSA method [96] and to the inability to obtain converged values of the conformational entropy of binding, which when combined with the solvation free energies, should provide a more realistic estimate of the true affinities (see reference [113]). For this reason, we focused on the structural analyses of the binding poses and on protein-ligands interactions.

In this regard, comparison between the binding regions of amitriptyline and chlorpromazine with those of known AcrB inhibitors, such as MBX3132 [47], revealed that considerable analogies exist, especially in the case of chlorpromazine. Accordingly, upon superposition of the complex structures of amitriptyline and chlorpromazine with the MBX3132-*Ec*AcrB, it emerged that MBX3132 forms a higher number of clashes with chlorpromazine than with amitriptyline (Table 3.7).

A similar comparison of the binding regions was made with norfloxacin and ethidium bromide. A significant overlap of ethidium bromide with both amitriptyline and chlorpromazine was observed in *Ec*AcrB as well as *S*TmAcrB (Fig. 3.3; see also Table 3.7). A considerably lower number of clashes was instead obtained in the case of norfloxacin (Table 3.7), which is found on top of chlorpromazine in both *Ec*AcrB and *S*TmAcrB, above amitriptyline in *Ec*AcrB, and below amitriptyline in *S*TmAcrB (Fig. 3.3).

In this regard, additional hints came from the analysis of our blind docking results. Indeed, it revealed that in both *Ec*AcrB and *S*TmAcrB, chlorpromazine, but not amitriptyline, binds just beneath the CH3 channel (see Table 3.3; see also Chapter 1, Section 1.4 for a description of this region). The binding poses in this region would clash with several poses found for norfloxacin and ethidium bromide in the same region of AcrB (Fig. 3.5). Notably, for both substrates, the numbers of poses behind this entry gate were greater in *Ec*AcrB than in *S*TmAcrB for corresponding monomers (L or T), while the numbers of chlorpromazine or amitriptyline poses in the proximity of CH3 were fairly similar.

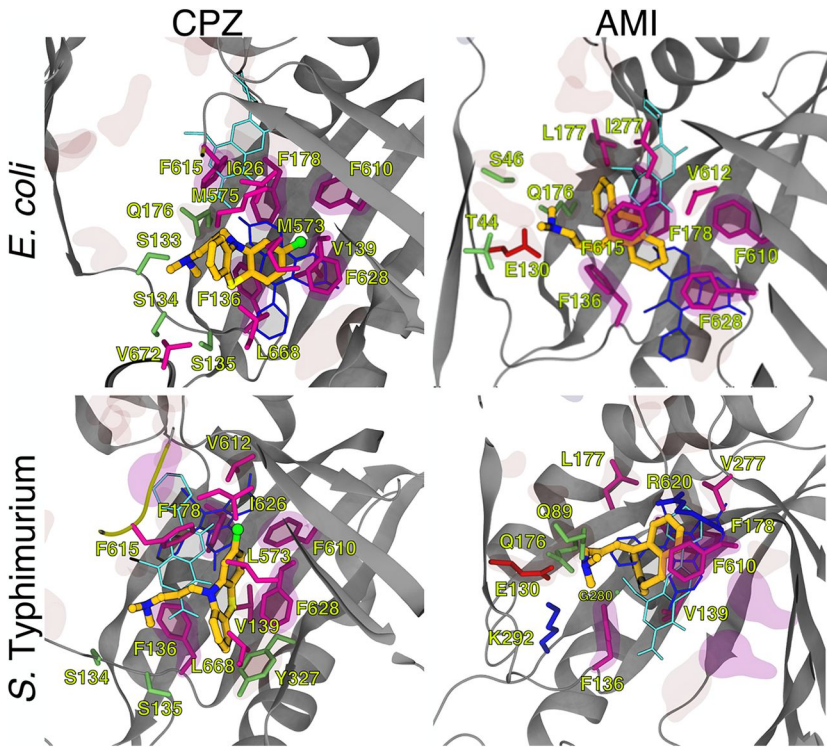

Figure 3.3: Representative conformations of the most stable binding modes of chlorpromazine and amitriptyline within the DBP_T of *Ec*AcrB and *S*TmAcrB, as obtained from all-atom MD simulations of the periplasmic portion of the transporter in explicit solvent (see Subsection 3.2.3). The protein is shown as gray ribbons, the inhibitors as CPK colored by element (C, N, S, and Cl in dark yellow, blue, light yellow, and green, respectively). Side chains of residues within 3.5 Å of the inhibitors are also shown as sticks colored by residue type (hydrophobic, polar, acid, and basic in purple, lime, red, and blue, respectively) and labeled. Side chains of residues defining the DBP_T and the phenylalanines lining the hydrophobic trap (see Table 3.3 for the definition of different protein regions) are also shown in transparent red and magenta surfaces, respectively. The most stable conformations of norfloxacin and ethidium bromide as obtained also from all-atom MD simulations are shown for comparison in cyan and blue sticks, respectively.

AMI - *Ec*AcrB	E130 (69.3%), Q176 (60.0%), water-mediated interactions (4.7%)
AMI - *S*TmAcrB	E130 (67.1%), Q176 (64.8%), water-mediated interactions (4.6%)
CPZ - *Ec*AcrB	S133 (82.7%), S134 (58.9%), water-mediated interactions (17.5%)
CPZ - *S*TmAcrB	S134 (52.0%), S135 (44.0%), F136* (10.0%), F617* (9.8%), water-mediated interactions (10.0%)

*cation-π interactions

Table 3.5: Residues of AcrB interacting with the polar tail of amitriptyline and chlorpromazine along the MD trajectories. CPZ, chlorpromazine; AMI, amitriptyline.

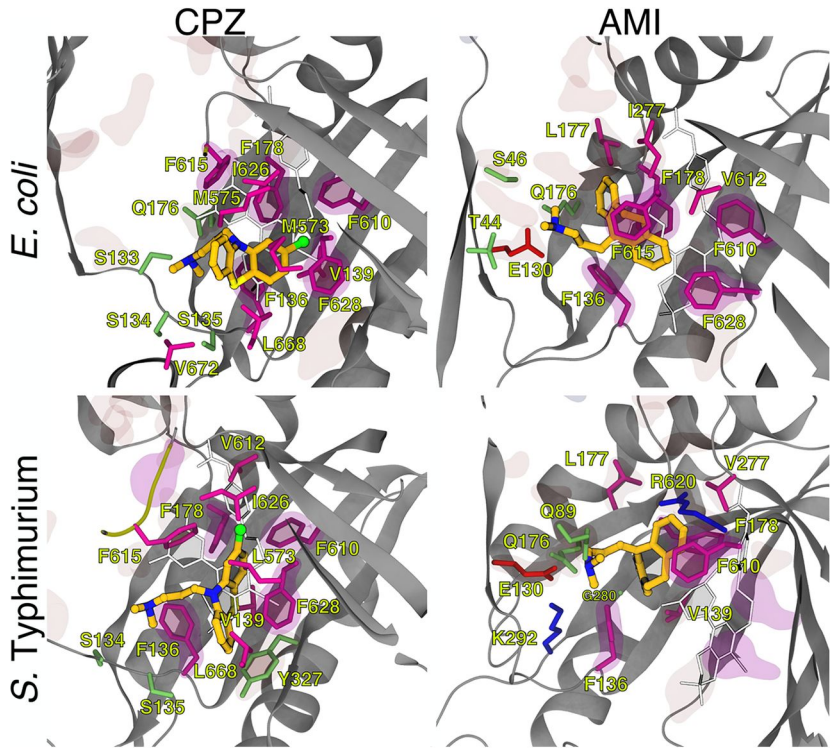

Figure 3.4: Comparison between representative conformations of the most stable binding modes of chlorpromazine and amitriptyline. Drugs are shown within the DBP_T of $AcrB_{EC}$ and $AcrB_{ST}$ and the experimental structure (shown as CPK colored by element) of the pyranopyrimidine inhibitor MBX3132 in $AcrB_{EC}$ (shown as white sticks). See Fig. 3.3 for details.

Organism	Compound	ΔG_b (kcal/mol)	DBP	HT
E. coli	CPZ	-31.9 (4.0)	-13.9	-8.9
	AMI	-25.6 (3.4)	-9.1	-6.6
	NOR	-36.4 (5.2)	-10.0	-6.5
	EtBr	-43.5 (2.9)	-14.6	-10.9
	MBX3132	-51.7	-19.6	-13.4
S. Typhimurium	CPZ	-25.7 (3.1)	-10.3	-8.4
	AMI	-27.7 (3.2)	-10.8	-5.7
	NOR	-29.8 (3.3)	-12.7	-10.1
	EtBr	-34.8 (2.9)	-14.0	-8.7

[a]The absolute values of $\Delta G_b$are reported with standard errors in parentheses together with the contribution to stabilization of the complexes from residues lining the DBP and the hydrophobic trap (HT). For comparison, data for MBX3132 bound to *Ec*AcrB are also reported [47]. CPZ, chlorpromazine; AMI, amitriptyline; NOR, norfloxacin; EtBr, ethidium bromide.

Table 3.6: Binding free energies to the DPT of *Ec*AcrB and *S*TmAcrB, calculated with the MM/GBSA approach[a].

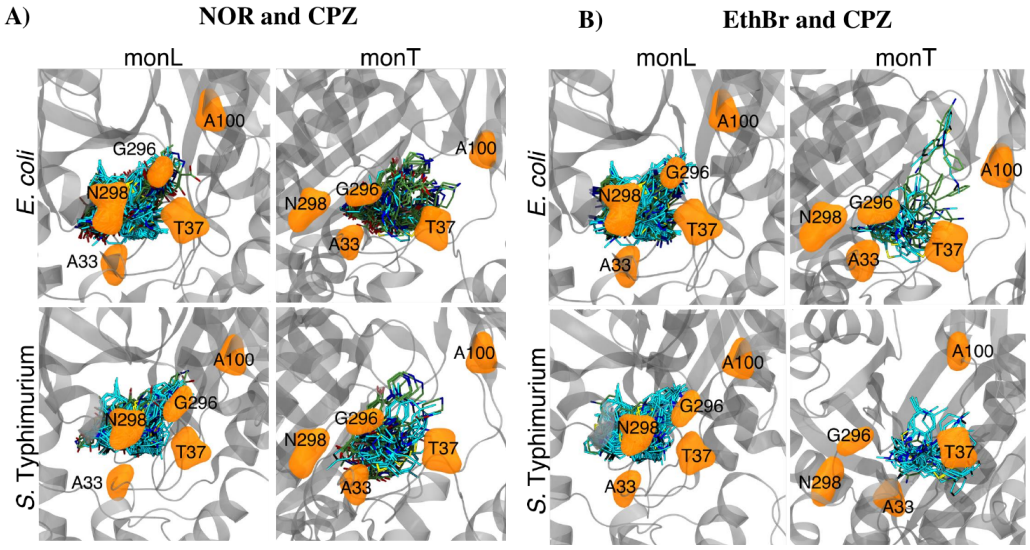

Figure 3.5: Overlap between docking poses of norfloxacin and ethidium bromide with chlorpromazine beneath the CH3 entry gates of monomers L and T in $AcrB_{EC}$ and $AcrB_{ST}$. The conformations of the substrate and the inhibitor are shown as sticks, with C atoms in lime and cyan colors, respectively. Sidechains of residues lining the CH3 entry (with the addition of residue 296, possibly involved in the recognition of carboxylated compounds [1]) are shown as orange semitransparent surfaces. (A) Overlap between the docking poses of norfloxacin and chlorpromazine. (B) Overlap between the docking poses of ethidium bromide and chlorpromazine. CPZ, chlorpromazine; AMI, amitriptyline; EthBr, ethidium bromide; NOR, norfloxacin.

Compound[b]	No. of atomic clashes		
	Norfloxacin	Ethidium Bromide	MBX3132
AMI - $AcrB_{EC}$	3	0	3
AMI - $AcrB_{ST}$	5	4	0[c]
CPZ - $AcrB_{EC}$	4	11	14
CPZ - $AcrB_{ST}$	6	3	15[c]

[a]The calculation was performed on the representative structure of the most populated cluster extracted from each MD trajectory (in the case of amitriptyline and chlorpromazine, we selected the trajectories associated with the more negative binding free energies among those displaying a stable position of the ligand in the last 50 ns of the production run). In addition, we used the crystal structure of *E. coli* AcrB in which the inhibitor MBX3132 has been cocrystallized (PDB ID 5ENQ). To evaluate the number of clashes, these structures were superimposed, and the number of heavy atoms of amitriptyline/chlorpromazine that overlap the other compounds was recorded.
[b]CPZ, chlorpromazine; AMI, amitriptyline.
[c]Under the hypothesis that MBX3132 binds to *S*TmAcrB similarly to the mode found in the X-ray structure 5ENQ of *Ec*AcrB.

Table 3.7: Number of atom clashes between atoms of chlorpromazine and amitriptyline and those of substrates norfloxacin and ethidium bromide and those of the inhibitor MBX3132 bound to $AcrB_{EC}$ (PDB ID: 5ENQ, [47])[a].

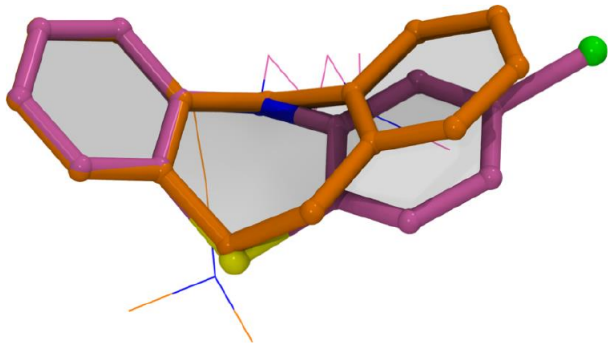

Figure 3.6: Comparison between equilibrium three-dimensional (3D) structures of chlorpromazine and amitriptyline. The rings building the molecular core of the two compounds are shown in CPK representation, with C atoms colored mauve and orange for chlorpromazine and amitriptyline, respectively. The tails are shown as lines colored with the same scheme.

3.4 Discussion

On the basis of our *in silico* investigations, chlorpromazine and amitriptyline can bind AcrB in its DBP. Specifically, both compounds interact fully or partly with the hydrophobic trap, which is known to be the preferred binding site of other EPIs active against AcrB. In *Ec*AcrB, in particular, chlorpromazine partly overlaps with the experimental binding pose of MBX3132, obtained through X-ray crystal structure (PDB ID: 5ENQ, [47]). Considerable overlaps between the two compounds would also be retained in *S.* Typhimurium, under the hypothesis that MBX3132 occupies the same binding region in *S*TmAcrB.

Such analogies between the binding modes suggest that chlorpromazine and MBX3132 may have a similar inhibition mechanism. As described in Chapter 1, EPIs of the MBX series are thought to act by competitive binding with other AcrB substrates or to restrain the concerted motions of the protein, associated with the functional rotation (see Chapter 1, Section 1.3) [47]. Chlorpromazine, therefore, may present a similar mode of action.

With regard to amitriptyline, smaller overlaps were detected with MBX3132 in *E. coli*, while in *S.* Typhimurium the two compounds are not superimposed. This is due to the fact that amitriptyline binds slightly upward with respect to MBX3132 and chlorpromazine (Fig. 3.4), interacting to a lower extent with the hydrophobic trap. This difference may be due to the additional chlorine atom in chlorpromazine, which can establish tight C-Cl$\cdots\pi$ interactions [125] with two or even three phenylalanine residues of this region. This hypothesis is consistent with the lower inhibitory effect of amitriptyline with respect to chlorpromazine (exception made for the impact on the accumulation of ethidium bromide in *S.* Typhimurium, which was comparable for the two compounds [108]). Weaker interactions with the hydrophobic residues of the DBP (particularly with the hydrophobic trap) may indeed result in a weaker competitive binding with other substrates, or in smaller effects on the dynamics of the protein.

Moreover, the results of our blind docking campaing suggest that chlorpromazine, but not amitriptyline, could interfere with the uptake of norfloxacin and ethidium bromide from the CH3 entry (see Chapter 1, Section 1.4; see also Fig. 3.5). In this regard, although CH3 was suggested to be the preferred binding site for the class of planar, aromatic and cationic com-

pounds, (i) both chlorpromazine and amitriptyline are cationic but not planar compounds; however, the phenothiazine ring of chlorpromazine confers the molecular core of this molecule a flatter conformation than that assumed in amitriptyline (see Fig. 3.6). (ii) Despite that ethidium bromide, but not norfloxacin, belongs to the class of compounds for which the CH3 entry was suggested as the preferred binding site, triple (A33W/T37W/N298W) and quadruple (A33W/T37W/A100W/N298W) mutants with amino acid substitutions in this channel resulted in 3- and 2-fold changes in the MICs of ethidium bromide and norfloxacin, respectively (see Table 1 in reference [51]). We speculate that the larger increase in the accumulation of norfloxacin upon coadministration of chlorpromazine rather than amitriptyline could be also due, at least in part, to competition for binding at the CH3 entrance gate. Overall, our findings allow a plausible and consistent rationale to be proposed for the different inhibitory potency of chlorpromazine and amitriptyline in *S.* Typhimurium and *E. coli*.

In summary, our work corroborates experimental data on the inhibitory activity of amitriptyline and chlorpromazine against AcrB. Our in silico investigations, indeed, demonstrated that both compounds can bind the DBP, partly occupying the hydrophobic trap. We propose that chlorpromazine and amitriptyline are substrates of the AcrAB-TolC efflux pump in *E. coli* and *S.* Typhimurium, and that they are able to bind residues primary involved in substrate recognition and/or transport. These compounds, therefore, could competitively inhibit efflux of other compounds or, alternatively, affect the functional rotation mechanism of AcrB.

Chapter 4

Study on the mode of action of the inhibitor PAβN against the AcrB transporter

4.1 Introduction

The discovery of EPIs has represented an important step in the development of strategies to reinvigorate in-use antibiotics [58]. The inhibition mechanism of these compounds, however, is not fully understood at present [57]. On the basis of several studies [47, 57], two main hypothesis have been formulated. EPIs may indeed act by restraining the dynamics of efflux pumps, which is crucial for the extrusion of their substrates (see Chapter 1, Sections 1.3 and 1.4). Alternatively, they could competitively bind specific regions of the pump which are essential for substrate recognition and/or transport. The latter hypothesis was formulated for the EPIs amitriptyline and chlorpromazine, active against AcrB in *E. coli* and *S.* Typhimurium, on the basis of our MD simulations (see Chapter 3).

In order to better understand the mode of action of the EPI PAβN (see Chapter 1, Section 1.5) and its possible effects on the dynamics of AcrB, an experimental investigation has been conducted by means of hydrogen-deuterium exchange mass spectrometry[1] (HDX-MS) [128]. This technique was applied to evaluate the dynamics and hydration properties of *E. coli* AcrB under various conditions. HDX-MS was indeed performed on the wt protein in presence of (i) the antibiotic ciprofloxacin ($AcrB_{WT}$-CIP), (ii) the EPI PAβN ($AcrB_{WT}$-PAβN) and (iii) both compounds ($AcrB_{WT}$-CIP-PAβN). Apo $AcrB_{WT}$ was used as a reference. Comparison of the results obtained for such systems revealed that PAβN can considerably rigidify several regions of the protein (Fig. 4.1a), including part of the binding pockets (see Chapter 1, Section 1.4). Similar restraints in the protein dynamics have been detected in $AcrB_{WT}$-CIP-PAβN, while they are negligible in $AcrB_{WT}$-CIP (Fig. 4.1a), suggesting that PAβN may act by affecting

[1] Hydrogen-deuterium exchange mass spectrometry (HDX-MS) [126, 127] is an experimental technique based on the hydrogen exchange reaction characteristic of peptide amine groups. Upon exposure of the sample to heavy water, indeed, backbone nitrogens in solvent-accessible regions of the protein tend to exchange their bonded hydrogen with nearby deuterium atoms. The rate of the exchange reaction heavily depends on conformational dynamics. Backbone amides are indeed less exposed to HDX in structured regions. However, structural fluctuations, local unfolding events and rigid-body motions can significantly favour the exchange reaction. This technique, therefore, can be useful to investigate the protein conformational dynamics and its solvent exposure.

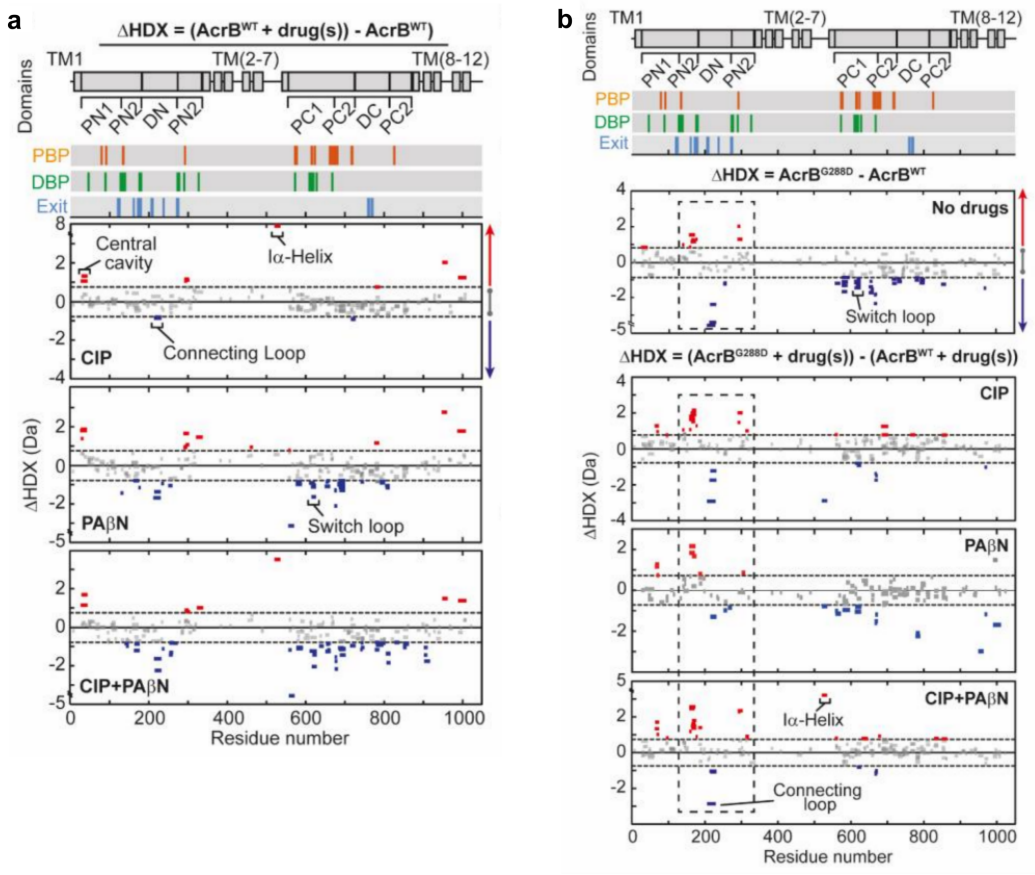

Figure 4.1: (a) HDX plots for different drug conditions (ΔHDX = ($AcrB_{WT}$+drug(s)) – $AcrB_{WT}$). Red signifies peptides with increased HDX (backbone H-bond destabilisation) in drug-bound state and blue representspeptides with decreased HDX (backbone H-bond stabilisation). 98% confidence intervals are shown as grey dotted lines and grey data are peptides with insignificant HDX. (b) HDX plots for different drug conditions (ΔHDX = ($AcrB_{G288D}$+drug(s)) – ($AcrB_{WT}$+drug(s))). Image adapted from [128].

functional dynamics in AcrB.

According to MIC assays [128], PAβN is also effective against the fluoroquinolone resistant G288D variant of AcrB ($AcrB_{G288D}$; see Chapter 1, Section 1.4). The HDX-MS procedure followed for the wt protein was then applied to $AcrB_{G288D}$ (see Fig. 4.1b) [128]. Even in this case, rigidification of several AcrB regions was detected in presence of PAβN ($AcrB_{G288D}$-PAβN), although some segments of the protein present an increased flexibility. In presence of both the inhibitor and ciprofloxacin ($AcrB_{G288D}$-CIP-PAβN), restrictions in the protein dynamics were detected to a less extent than in $AcrB_{WT}$-CIP-PAβN.

In HDX-MS assays of AcrB in presence of PAβN, the porter domain undergoes the most relevant rigidification effects, especially in the DBP and PBP. In order to investigate the correlation of the hydration and dynamics of AcrB with the binding of PAβN and of ciprofloxacin, HDX-MS assays were compared with MD simulations of systems $AcrB_{WT}$-PAβN, $AcrB_{WT}$-CIP-PAβN, $AcrB_{G288D}$-PAβN and $AcrB_{G288D}$-CIP-PAβN. Trajectories were analysed to evaluate the effect of substrate binding on the structural fluctuations and solvent-accessibility of

the protein. Moreover, simulations of $AcrB_{WT}$-CIP were performed to evaluate how the co-presence of PAβN in the binding pockets affects the binding of ciprofloxacin. In Section 4.2, we describe the computational protocol employed in this work. Section 4.3 is dedicated to the description of the results and to their comparison with HDX-MS assays. A final discussion on this investigation is reported in Section 4.5.

4.2 Methods

4.2.1 Molecular docking

A blind docking campaign was first performed using Autodock Vina [94]. As done in [129, 130], a rectangular search space of size 125Å x 125Å x 110Å enclosing the whole portion of the protein potentially exposed to ligands was adopted. The exhaustiveness parameter, related to the extent of the exploration of the search space, was set to 8192 (~1000 times the default 8) in order to improve the sampling of docking poses within the large box used (~64 times the default 30Å x 30Å x 30Å). Flexibility of both partners was considered indirectly, by employing multiple conformations in ensemble docking runs [131]. For both CIP and PAβN, 10 representative molecular conformations were obtained from 1 μs-long molecular dynamics simulations of the compounds in presence of explicit solvent (data available at *www.dsf.unica.it/translocation/db*) [90]. Namely, a cluster analysis of the trajectories of the ligands was performed as described in Malloci *et al.* [90], setting the number of cluster representatives to 10.

For the wt receptor, 10 X-ray asymmetric high-resolution structures (with PDB IDs: 2GIF [132], 2DHH [44], 2J8S [112], 3W9I [133], 4DX5, 4DX7 [45], 4U8V, 4U8Y, 4U95, 4U96 [46]) were considered, most bearing a substrate bound to the transporter. We also employed 10 structures for $AcrB_{G288D}$, namely the homology models derived on top of the $AcrB_{WT}$ X-ray structures mentioned above. Regarding the homology modelling protocol, the sequence of the G288D variant was first generated by manually modifying the FASTA file of the corresponding amino acid sequence of *E. coli* AcrB retrieved from the Uniprot database (UniProt ID: P31224) [74]. Next, 100 homology models were generated for each template with the Modeller 9.21 software [76]. The variable target function method was used to perform the optimization, and the best model (that is the one with the highest value of the MOLPDF function) was employed in docking calculations.

The ensemble docking campaign resulted in several hundred poses per ligand, most of which were located inside the distal binding pocket of the monomer in the T state (DBP_T), which is the putative binding site for the recognition of compounds with physico-chemical features similar to those of the molecules investigated in this work [134]. Because most docking poses were concentrated in this region, we performed a second docking campaign using a grid of 30Å x 30Å x 30Å and centered at DBP_T. Next, we performed a cluster analysis of the docking poses using as a metric the heavy-atoms RMSD of the substrate (setting the cutoff to 3 Å), which returned respectively 11, 9, 15 and 17 different poses for the $AcrB_{WT}$–PAβN, $AcrB_{WT}$–CIP–PAβN, $AcrB_{G288D}$–PAβN, $AcrB_{G288D}$–CIP–PAβN complexes (Tables 4.1 and 4.2; see also Fig. 4.2). Moreover, to evaluate how the presence of PAβN affects the binding of CIP in the ternary complexes, we selected three docking poses of CIP onto $AcrB_{WT}$ and $AcrB_{G288D}$. In the case of the $AcrB_{WT}$–CIP complex, to consider the largest number of putative binding modes, we purposely selected docking poses with an orientation different than that reported previously [60].

	%res	30%			40%		
	Site	N	ΔG_{max}	<G>	N	ΔG_{max}	<G>
PAβN	PBP_L	12	-11.5	-10.9 ± 0.3	3	-11.3	-10.9 ± 0.3
	PBP_T	19	-13.0	-11.6 ± 0.6	-	-	-
	$CH1_L$	1	-10.8	-10.8	1	-10.8	-10.8
	$CH1_T$	-	-	-	-	-	-
	$CH2_L$	1	-11.3	-11.3	-	-	-
	$CH2_T$	-	-	-	-	-	-
	$CH3_L$	15	-12.3	-11.3 ± 0.4	9	-12.3	-11.4 ± 0.4
	$CH3_T$	33	-12.5	-11.4 ± 0.4	22	-12.5	-11.5 ± 0.4
	DBP_T	148	-13.7	-11.6 ± 0.6	87	-13.7	-11.7 ± 0.7
CIP	PBP_L	-	-	-	-	-	-
	PBP_T	-	-	-	-	-	-
	$CH1_L$	-	-	-	-	-	-
	$CH1_T$	-	-	-	-	-	-
	$CH2_L$	-	-	-	-	-	-
	$CH2_T$	-	-	-	-	-	-
	$CH3_L$	55	-10.2	-9.4 ± 0.3	19	-9.6	-9.3 ± 0.2
	$CH3_T$	31	-10.0	-9.5 ± 0.2	4	-9.7	-9.4 ± 0.3
	DBP_T	123	-11.5	-9.7 ± 0.4	20	-10.3	-9.6 ± 0.4

Table 4.1: Number of poses (N), maximum and average (pseudo)binding free energy (ΔG_{max} and <G>, respectively) of CIP and PAβN binding to $AcrB_{WT}$. The percentages in the first row are meant to identify the poses having contacts (that is minimum ligand-residue distance below a cutoff set to 3.5 Å) at least with 30% or 40% of residues lining the corresponding site. See Fig. 4.2(a,b) for a representation of the distribution of the poses.

4.2.2 Molecular dynamics simulations

All of the 52 complexes selected from docking runs were subjected to all-atom molecular dynamics (MD) simulations (each of 1 μs in length) performed with the AMBER18 package [86].

Protomer-specific protonation states of AcrB were adopted following previous work [135]: residues E346 and D924 were protonated only in the L and T protomers, while residues D407, D408, and D566 were protonated only in the O protomer, of AcrB. The topology and the initial coordinate files were created using the LEaP module of the AMBER18 package. The proteins were embedded in a mixed bilayer patch composed of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol (POPG) in a 2/1 ratio, for a total of 660 lipid molecules symmetrically distributed in the two leaflets of the bilayer. The whole system was solvated with a 0.15 M aqueous NaCl solution. The AMBER force field protein.fb15 [137] was used to represent the protein; lipid17 (*http://ambermd.org/GetAmber.php*) parameters were used for the POPE and POPG molecules; the TIP3PFB model was employed for water [136]. The General Amber Force-Field (GAFF) parameters [138] for CIP and PAβN were taken from Malloci *et al.* [90].

Each system was first subjected to a multi-step structural relaxation via a combination of steepest descent and conjugate gradient methods using the pmemd program implemented in AMBER18, as described in previous publications [47, 60, 135]. The systems were then heated

	%res	30%			40%		
	Site	N	ΔG_{max}	<G>	N	ΔG_{max}	<G>
PAβN	PBP_L	-	-	-	-	-	-
	PBP_T	3	-11.5	-11.1 ± 0.3	-	-	-
	$CH1_L$	-	-	-	-	-	-
	$CH1_T$	2	-10.8	-10.8 ± 0.0	2	-10.8	-10.8 ± 0.0
	$CH2_L$	-	-	-	-	-	-
	$CH2_T$	2	-11.2	-11.1 ± 0.0	-	-	-
	$CH3_L$	18	-12.9	-11.5 ± 0.6	7	-12.0	-11.5 ± 0.4
	$CH3_T$	11	-13.5	-11.7 ± 0.8	4	-11.8	-11.3 ± 0.3
	DBP_T	24	-13.4	-11.6 ± 0.7	2	-11.3	-11.2 ± 0.1
CIP	PBP_L	-	-	-	-	-	-
	PBP_T	-	-	-	-	-	-
	$CH1_L$	-	-	-	-	-	-
	$CH1_T$	-	-	-	-	-	-
	$CH2_L$	-	-	-	-	-	-
	$CH2_T$	-	-	-	-	-	-
	$CH3_L$	14	-9.9	-9.2 ± 0.3	-	-	-
	$CH3_T$	9	-9.4	-9.1 ± 0.2	-	-	-
	DBP_T	5	-9.8	-9.5 ± 0.2	1	-9.3	-9.3

Table 4.2: Number of poses (N), maximum and average (pseudo)binding free energy (ΔG_{max} and <G>, respectively) of CIP and PAβN binding to $AcrB_{G288D}$. See the caption of Table 4.2 for further details; see also Fig. 4.2(c,d) for a representation of the distribution of the poses.

from 0 to 310 K in two subsequent MD simulations: i) from 0 to 100 K in 1 ns under constant-volume conditions and with harmonic restraints (k = 1 kcal·mol-1·Å-2) on the heavy atoms of both the protein and the lipids; ii) from 100 to 310 K in 5 ns under constant pressure (set to a value of 1 atm) and with restraints on the heavy atoms of the protein and on the z coordinates of the phosphorous atoms of the lipids to allow membrane rearrangement during heating. As a final equilibration step, a series of 20 equilibration steps, each of which was 500 ps in duration (total 10 ns), with restraints on the protein coordinates, were performed to equilibrate the box dimensions. These equilibration steps were carried out under isotropic pressure scaling using the Berendsen barostat, whereas a Langevin thermostat (collision frequency of 1 ps^{-1}) was used to maintain a constant temperature. Finally, production MD simulations of 1 s were performed under an isothermal-isobaric ensemble for each system. A time step of 2 fs was used for all runs before production, while the latter runs were carried out with a time step of 4 fs after hydrogen mass repartitioning [98].

During the MD simulations, the lengths of all the R–H bonds were constrained with the SHAKE algorithm. Coordinates were saved every 100 ps. The Particle mesh Ewald algorithm was used to evaluate long-range electrostatic forces with a non-bonded cut-off of 9 Å.

4.2.3 Post-processing of MD trajectories

MD trajectories were analyzed using either in-house *tcl* and *bash* scripts or the *cpptraj* tool of AMBER18. Figures were prepared using gnuplot 5.0 [120] and VMD 1.9.3 [121]. All the calculations with the exception of the cluster analysis were performed on the conformations taken

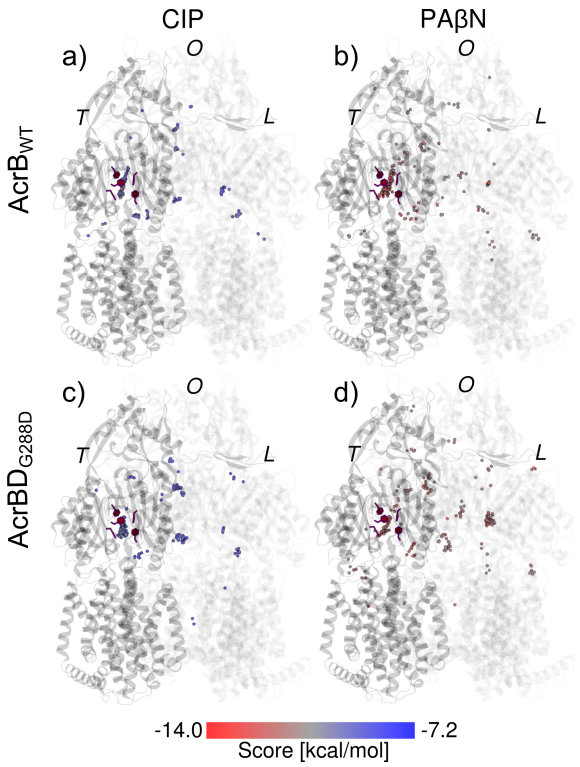

Figure 4.2: Distribution of top 200 docking poses (only the centres of mass are shown for clarity) for CIP and PAβN onto AcrB$_{WT}$ (a,b) and AcrB$_{G288D}$ (c, d). The spheres are coloured according to the value of the (pseudo)free energy of binding (docking score). The monomers L, T and O are shown as transparent ribbons (T and O darkest and lightest, respectively).

from the most populated conformational cluster (representing the most sampled conformation of the complex along the production trajectories) along the last 300 ns of the production runs.

Cluster analysis. Clustering of the ligand trajectory was carried out using the average-linkage hierarchical agglomerative clustering method implemented in cpptraj and employing an RMSD cut-off of 3 Å calculated on all the heavy atoms of the ligand.

System Stability. The RMSDs of the protein and of the substrates were calculated using cpptraj after structural alignment of each trajectory. Namely, we calculated the Cα-RMSD of the protein with respect to the initial (docking) structure after alignment of the whole trimer. The RMSDs of the substrates were calculated with respect to the corresponding structure of the selected docking pose, as well as with respect to the last frame of the MD trajectory. In particular, to evaluate the magnitude of the displacements and reorientations of the substrates during the simulations, their RMSDs were calculated upon alignment of the T monomer of the protein to the reference frame.

Interaction network. Interactions stabilizing the complexes were analysed by considering residues within 3.5 Å of each substrate in the last 300 ns of the MD trajectories. Hydrogen bonds were identified through geometrical criteria, using a cut-off of 3.2 Å for the distance between donor and acceptor atoms and a cut-off of 135° for the donor-hydrogen-acceptor angle. Such analyses were conducted through in-house tcl scripts. Occupancy levels of hydrogen bonds and water-mediated interactions (detected in the last 300 ns of each simulation) were

System	Reference
$AcrB_{WT}$-PAβN (T state)	apo$AcrB_{WT}$ (L state)
$AcrB_{WT}$-CIP–PAβN (T state)	apo$AcrB_{WT}$ (L state)
$AcrB_{G288D}$-PAβN (T state)	$AcrB_{WT}$-PAβN (T state)
$AcrB_{G288D}$-CIP-PAβN (T state)	$AcrB_{WT}$-CIP-PAβN (T state)

Table 4.3: Systems considered for the analyses of flexibility and hydration properties based on MD simulations, and respective reference structures. The state of each system was chosen in agreement with Wang *et al.* [37] (see Subsection 4.2.3).

Regions	Residues
Central cavity	25-33, 36, 37, 96, 97, 385-389, 457-466, 468, 469
Distal binding pocket (DBP)	44, 46, 48, 87, 89, 128, 130, 132, 134, 136, 139, 176-180, 273, 274, 276, 277, 327, 573, 610, 612, 615, 617, 620, 626, 628
Hydrophbic trap (HT)	136, 178, 289, 291, 573, 610, 612, 615, 626, 628
Proximal binding pocket (PBP)	79-81, 89-91, 132-134, 573, 575, 577, 617, 662-669, 672-681, 683, 717, 719, 815, 826, 828-830
Connecting-loop	206-243
Iα-Helix	520-534
Switch loop	615-620
Hosting loop	860-871
Exit gate (EG)	124, 125, 163, 164, 174, 208-221, 239, 240, 758-761, 767-770
Channel 1 (CH1)	836, 838, 840, 842, 868, 870, 872
Channel 2 (CH2)	566, 645, 649, 653, 656, 662, 676, 678, 715, 717, 719, 722, 830
Channel 3 (CH3)	33, 37, 100, 296, 298

Table 4.4: List of peptides considered in the regions of AcrB.

also computed using cpptraj. For systems $AcrB_{WT}$–PAβN, $AcrB_{WT}$-CIP–PAβN, $AcrB_{G288D}$–PAβN and $AcrB_{G288D}$–PAβN, the following analyses were also performed to evaluate their agreement with HDX-MS data.

System Flexibility. The Root Mean Square Fluctuations (RMSFs) of the protein were calculated using cpptraj after structural alignment of each trajectory as described in the previous paragraph.

Hydration properties. Residue-wise average numbers of waters within the first (second) hydration layer were calculated with cpptraj using a distance cut-off of 3.4 (5.0) Å between the nitrogen of the protein and the water oxygens.

Comparison with HDX-MS data. RMSFs and hydration properties of each system were compared with a proper reference state according to the current knowledge about the most likely conformations assumed by AcrB in the absence of ligands or complexed with substrates and inhibitors [37]. For instance, to account for conformational changes of AcrB induced by inhibitor binding, PAβN-bound and apo AcrB structures were considered in their T and L state, respectively. The T state was also considered for systems containing both PAβN and CIP ($AcrB_{WT}$-CIP–PAβN and $AcrB_{G288D}$-CIP–PAβN), hypothesizing their stability in this conformation, as evidenced by the RMSDs analyses conducted on our trajectories (Fig. 4.3, 4.5, 4.7, 4.9). The list of reference states used for each analysis are reported in Table 4.3.

4.3 Results

4.3.1 Methods for the comparison of MD data with HDX-MS assays

In this work, the effect of PAβN and ciprofloxacin binding on the dynamics and hydration properties of $AcrB_{WT}$ and $AcrB_{G288D}$ have been investigated by means of computational methods. Firstly, molecular docking has been used to create the initial configurations for systems $AcrB_{WT}$-PAβN, $AcrB_{WT}$-CIP-PAβN, $AcrB_{G288D}$-PAβN and $AcrB_{G288D}$-CIP-PAβN. MD simulations of these systems have then been performed, together with simulations of apo $AcrB_{WT}$, which were used as references in the postprocessing of the trajectories of the wt protein (see Section 4.2).

In order to evaluate the properties of each system, the RMSF and first hydration shell profiles (see Section 4.2) were compared with HDX-MS data. Overall, the average number of water molecules in the first amide NH solvation shell computed by MD simulations has been found to correlate well with HDX. A reduced hydration shell should therefore imply reduced HDX, due to the decrease in specific interactions between amide N-H bonds and the solvent. However, protein HDX is complex, with neighbouring residues having significant differences in their solvent interactions. This, combined with the stark contrast between MD simulation and HDX-MS experimental time scales (µs to ms versus seconds to hours), means that a simple quantitative comparison can often be incomplete. Nevertheless, comparisons to MD calculated hydration profiles can provide informative qualitative interpretation of protein HDX.

In the following Subsections, the results obtained for the simulated systems are reported. Moreover, their agreement with the respective HDX-MS data is discussed.

4.3.2 $AcrB_{WT}$-PAβN

According to our MD simulations, an important contribution to the stabilization of PAβN in $AcrB_{WT}$ comes from the hydrophobic trap (HT), whose residues are involved in stacking with the β-naphthylamide moiety of the inhibitor (see Pose 1 in Fig. 4.3 for a representation of the binding mode). Importantly, these interactions also involve residues of the switch loop (such as F617) or belonging to adjacent regions. These findings, in agreement with previous literature [113, 122], support the hypothesis that the stabilization of the switch loop could be key to the mode of action of PAβN. This loop is also involved in the formation of stable hydrogen bonds with the amino group of the compound (see Table 4.3.2). Additional hydrogen bonds are formed by its polar groups with polar and acid residues of the DBP, including E130, K131 (involved in interactions with the guanidino group of PAβN; see Fig. 4.3, Pose 1)) and Q176 (interacting with the carbonyl group).

Important findings on the stabilization of the switch loop come from the comparison of the hydration properties of the complex $AcrB_{WT}$-PAβN and apo $AcrB_{WT}$ (Fig. 4.4). Indeed, despite the relevant difference in the timescales of all-atom MD simulations and typical HDX kinetics, and while the switch loop itself moderately enhanced hydration, the nearby segments (residues 612 to 615 and 620 to 624) are overall dehydrated with respect to the apo form (Fig. 4.4). Considerable dehydration and rigidification are also observed for part of the PBP and the CH2 entrance (Fig. 4.4), consistently with HDX-MS data (Fig. 4.1a).

In agreement with previous studies [113, 122], all the binding modes found for $AcrB_{WT}$-PAβN (Fig. 4.3) feature the β-naphthylamide moiety of the inhibitor within the HT and interacting with the switch loop or nearby residues. Additional common interactions involve

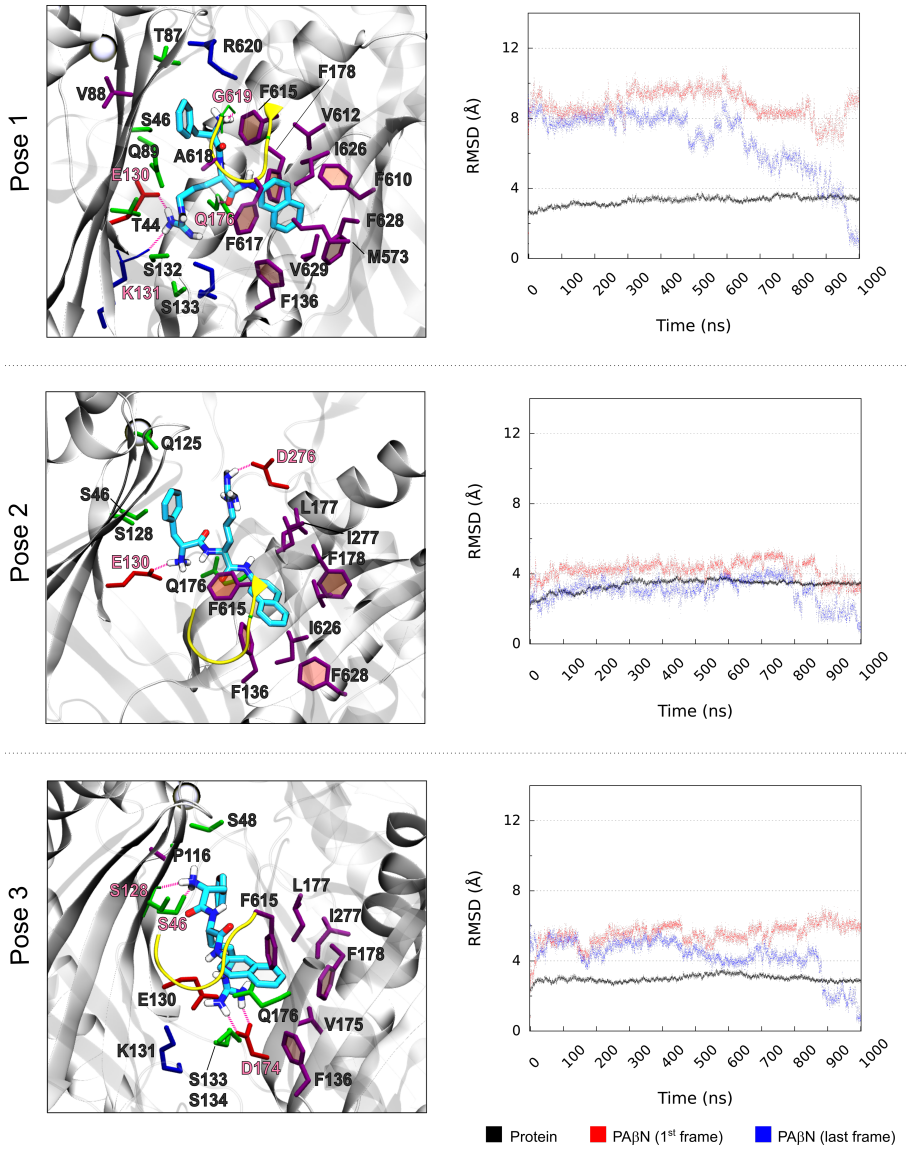

Figure 4.3: Representative binding poses and RMSDs of $AcrB_{WT}$-PAβN. In the representation of the binding poses, PAβN is coloured by atom type (C atoms in cyan, N atoms in blue and oxygen atoms in red, H atoms in white – only polar H atoms are shown). Residues within 3.5 Å are also shown, coloured by residue type (red: acidic; blue: basic; green: polar; purple: hydrophobic). Hydrogen bonds formed by PAβN are highlighted through magenta sticks, and the involved residues are labelled in pink (see Table 4.3.2 for high-occupancy hydrogen bonds involving PAβN). The switch loop is shown in yellow and the Cα atoms of the residues Q124 and Y758 belonging to the exit gate are represented as light blue spheres. Water molecules were not represented for clarity. See Table 4.9 for the RMSD of each pose with respect to reference structure PDB:4U95.

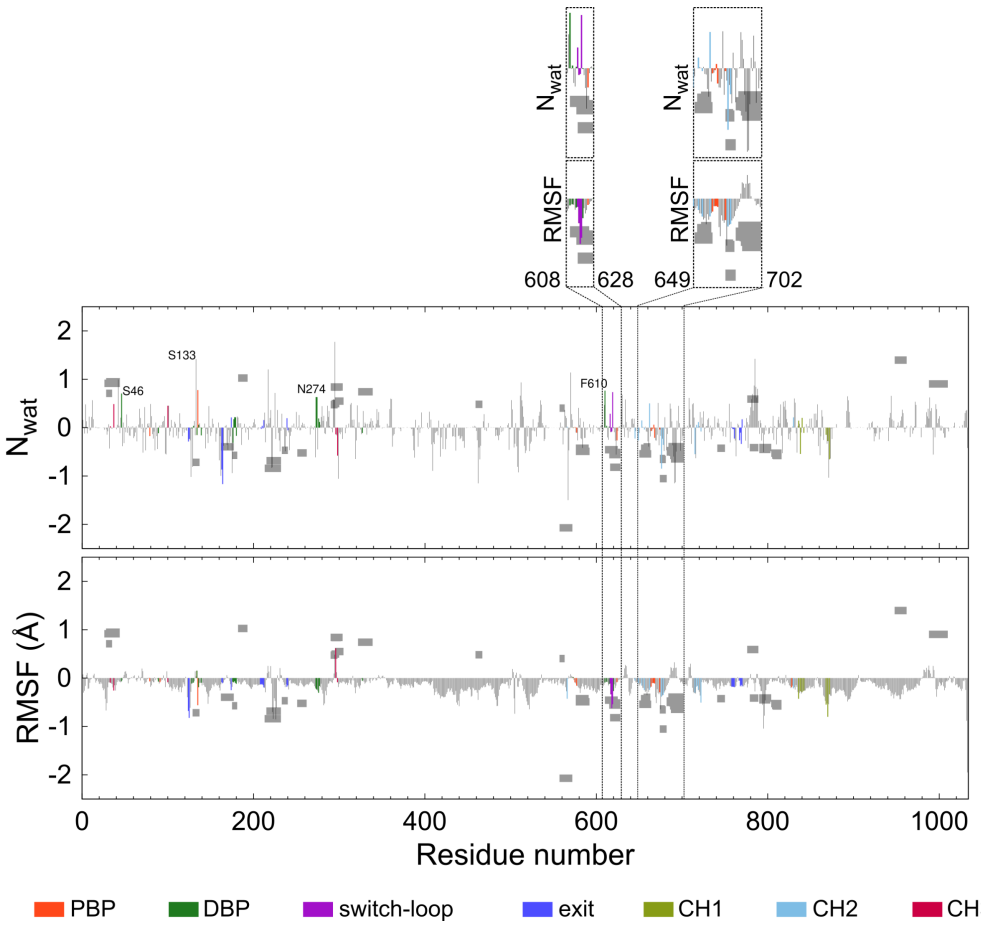

Figure 4.4: Difference in first hydration shell (N_{wat}) and RMSF between $AcrB_{WT}$-PAβN and apo $AcrB_{WT}$ (based on MD data from Pose 1 in Fig. 4.3). Differences in N_{wat} and RMSF are represented as histograms, with regions directly involved in substrate transport highlighted in different colours (see Table 4.4 for the definition of these regions). As a reference, HDX-MS data are represented as grey boxes (scale not shown). Both N_{wat} and RMSF differences have been computed between the T monomer of $AcrB_{WT}$-PAβN and the L monomer of apo $AcrB_{WT}$ (see Section 4.2). Regions of interest are highlighted in the upper part of the panel. In the N_{wat} plot, labelled residues are directly involved in interactions with PAβN and have a higher hydration in $AcrB_{WT}$-PAβN than in apo $AcrB_{WT}$.

Ligand	Pose	H-bonds		Water-mediated interactions	
		Residue	Occupancy (%)	Residue	Occupancy (%)
PAβN	1	Q176	67	E673	39
		G616	57	F617	36
		E673	40	S133	21
		G619	37		
		F617	23		
	2	D276	99	E130	67
		E130	87	D276	53
		Q176	60	L177	50
		G614	24		
	3	E273	100	D276	79
		S46	90	E273	45
		S48	39	Q176	26
		D174	26		

Table 4.5: Intermolecular hydrogen bonds (H-bonds) and water-mediated interactions involving PAβN in $AcrB_{WT}$-PAβN based on MD simulations. Analyses have been conducted on the last 300 ns of each simulation (see Section 4.2). Only interactions with occupancy higher than 20% have been reported. Representative poses are shown in Fig. 4.3.

the aromatic rings and the amino group of the inhibitor and the PBP/DBP interface, including residues of the PN1 subdomain (such as S46, S128 and E130). Residues of this region are involved in either stacking interactions with the phenyl ring of PAβN (Fig. 4.3, Pose 1, 2) or hydrogen bonds with the amino group (Pose 2; see Table 4.3.2). Furthermore, residues belonging to segment 130-134 also interact with the guanidino group of PAβN in two representative poses (Fig. 4.3, Pose 1, 3), with additional stabilization provided by Q176 and proximal residues. The other pose (Pose 2) is characterized by a different orientation of the guanidino group of the inhibitor, located in the upper part of the DBP and involved in interactions with D276 and nearby residues.

To evaluate whether the described interactions correlate with HDX-MS protection data, occupancy levels of protein-ligand hydrogen bonds and water-mediated interactions were computed (see Section 4.2). The results (Table 4.3.2) confirm that several residues belonging to protected peptides in the HDX-MS (segments 129-137, 162-181, 610- 628) form stable direct and/or water-mediated hydrogen bonds with the inhibitor. Stabilization of these regions of the DBP may thus be due to the interaction with PAβN.

4.3.3 $AcrB_{WT}$-CIP-PAβN

Interactions stabilizing CIP and PAβN in $AcrB_{WT}$ include hydrogen bonds between the two substrates (see Pose 1 in Fig. 4.5 for a representation of the binding pose; see also Table 4.10), as well as between them and the protein (*e.g.* between the guanidino group of PAβN and residues E130 and D174, or between R620 and the carboxylic and carbonyl group in CIP; see Table 4.6). Additional stabilization comes from stacking of aromatic rings, formed by PAβN with CIP and F615. Importantly, the direct interaction between the inhibitor and segments proximal to the switch loop, present in $AcrB_{WT}$-PAβN (Fig. 4.3, Pose 1; see also Table 4.3.2), is preserved also in the presence of CIP.

The comparison of the hydration properties of $AcrB_{WT}$-CIP-PAβN (Fig. 4.6) and $AcrB_{WT}$-PAβN (Fig. 4.4) reveals an analogous dehydration of the residues of the binding pockets (exception made for some residues involved in interactions with the compounds, such as E173, N174 and F615 in $AcrB_{WT}$-CIP-PAβN and similar variations in the region surrounding the switch loop (Fig. 4.6). Such region, involved in interactions with the substrates in both $AcrB_{WT}$-PAβN (Fig. 4.3, Pose 1) and $AcrB_{WT}$-CIP-PAβN (Fig. 4.5, Pose 1), is indeed considerably rigidified in both systems. This is associated with a dehydration of the segments adjacent to the loop, significantly marked in $AcrB_{WT}$-CIP-PAβN (Fig. 4.6), in agreement with HDX-MS data (Fig. 4.1a).

A common trait of the binding modes found for this ternary complex is the presence of direct interactions between the two substrates, through the formation of hydrogen bonds (involving, in all poses, the carboxylic group of CIP; see Table 4.10) as well as stacking of the aromatic rings (Fig. 4.5, Pose 1, 3, 4). While both CIP and PAβN are located inside the DBP in three representative poses (Pose 1, 3, 4), a different binding mode is predicted with CIP located within the PBP behind the switch loop (Pose 2).

Although some differences are present, comparison of the binding regions reveals several shared traits. Firstly, interactions of at least one substrate with the HT and (the region proximal to) the switch loop are preserved. Typically, such interactions involve π-stacking with the aromatic groups of PAβN, although cation-π interactions were also observed in Pose 3 (involving e.g. F178 and the amino group of PAβN). In Pose 2 and 4, additional stacking interactions are found between CIP and the switch loop and the nearby residues.

Apart from Pose 3, another conserved trait is related to the interaction with the PBP/DBP interface. Several contacts with residues of this region (such as S46, T87, S128 and adjacent residues; see Table 4.6) are formed by CIP in Poses 1 and 4, while in Pose 2 PAβN is involved in hydrogen bonds and polar interactions with E130 and nearby residues. In Pose 3, in which interactions with the PBP/DBP interface are not detected, several contacts are formed by CIP with polar and acidic residues of the PN2 portion of the DBP, including Q151 and E152. Comparison of such poses with 1-μs long MD simulations of $AcrB_{WT}$-CIP (see Section 4.2; see also Fig. 4.11 for a representative binding pose) revealed that the co-presence of PAβN determines a change in the binding region of CIP. In the absence of the EPI, indeed, CIP tends to occupy the HT, a finding consistent with previous reports [**?**]. Although starting from different orientations than those previously reported, our MD simulations confirmed that CIP establishes strong interactions with residues F136, Y327 and F628. Moreover, according to our analyses, further stabilization comes from high-occupancy hydrogen bonds involving Q176 (Fig. 4.11). Stabilizing interactions do not involve residues of the PBP/DBP interface, which instead play an important role in the ternary complex, as previously mentioned.

Analysis of the occupancy of the hydrogen bonds formed by CIP and PAβN with the protein revealed that both ligands form stable interactions in all poses (Table 4.6). Moreover, both ligands tend to form water-mediated interactions with the protein, frequently involving residues 128-133, 174-176 and 273-276. From the comparison with HDX-MS data, a good correlation between occupancy and protection data emerges for segments 138-149, 162-177, 263-274 (see Fig. 4.6 and Table 4.6). As in the case of $AcrB_{WT}$-PAβN, therefore, intermolecular interactions formed by CIP and PAβN may be a factor for the protection of significant portions of the DBP.

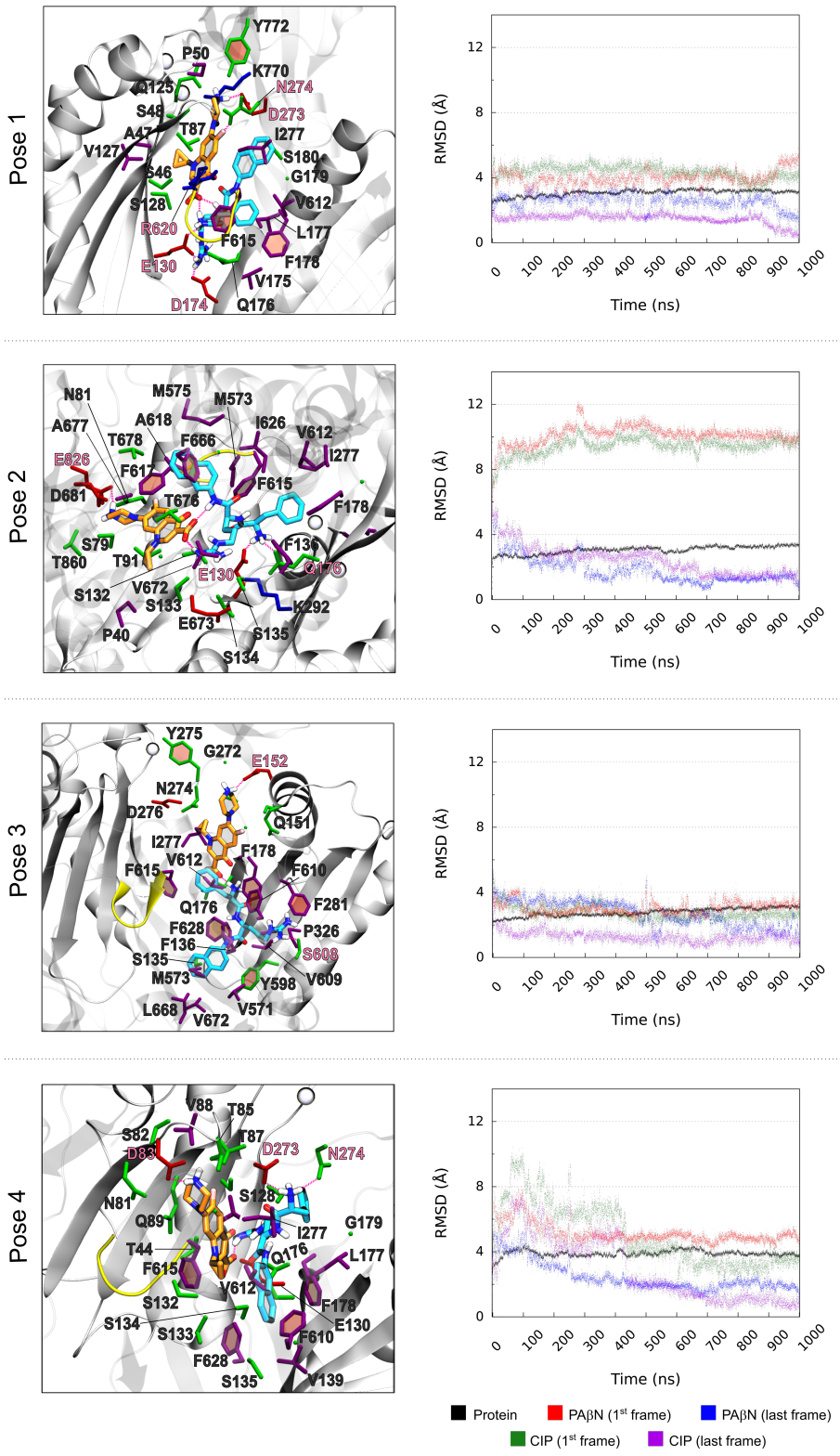

Figure 4.5: Representative binding poses and RMSDs of $AcrB_{WT}$-CIP-PAβN. To distinguish between the inhibitor and antibiotic, carbon atoms of CIP and PAβN are coloured in orange and cyan, respectively. See: Fig. 4.3 for further details; Table 4.6 for a list of direct and water-mediated hydrogen bonds established between each ligand and the protein; Table 4.10 for high-occupancy hydrogen bonds between the ligands; Table 4.9 for the RMSD of each pose with respect to reference structure PDB:4U95.

Ligand	Pose	H-bonds		Water-mediated interactions	
		Residue	Occupancy (%)	Residue	Occupancy (%)
PAβN	1	E130	100	D174	71
		D174	100	E130	57
		L177	69	Q89	42
	2	Q176	76	E130	40
		E130	66		
		S132	60		
		S133	45		
		S134	23		
		V672	20		
	3			S608	41
	4	D276	100	D276	50
		N274	69	E273	45
		S128	51	L177	43
		L177	48	S128	27
		S46	47		
CIP	1	R620	100	Q89	100
		E273	94	S128	100
		Q125	84	D681	40
	2	T676	90	E826	22
		E826	69		
		Y77	41		
	3	E152	94	E152	46
		Q176	80		
		N274	60		
	4	D83	100	L177	43
		Q176	58		
		T44	51		
		S133	51		
		T87	46		

Table 4.6: Intermolecular hydrogen bonds (H-bonds) and water-mediated interactions involving ligands in $AcrB_{WT}$-CIP-PAβN based on MD simulations. See Table 4.3.2 for details and Fig. 4.5 for representative binding poses.

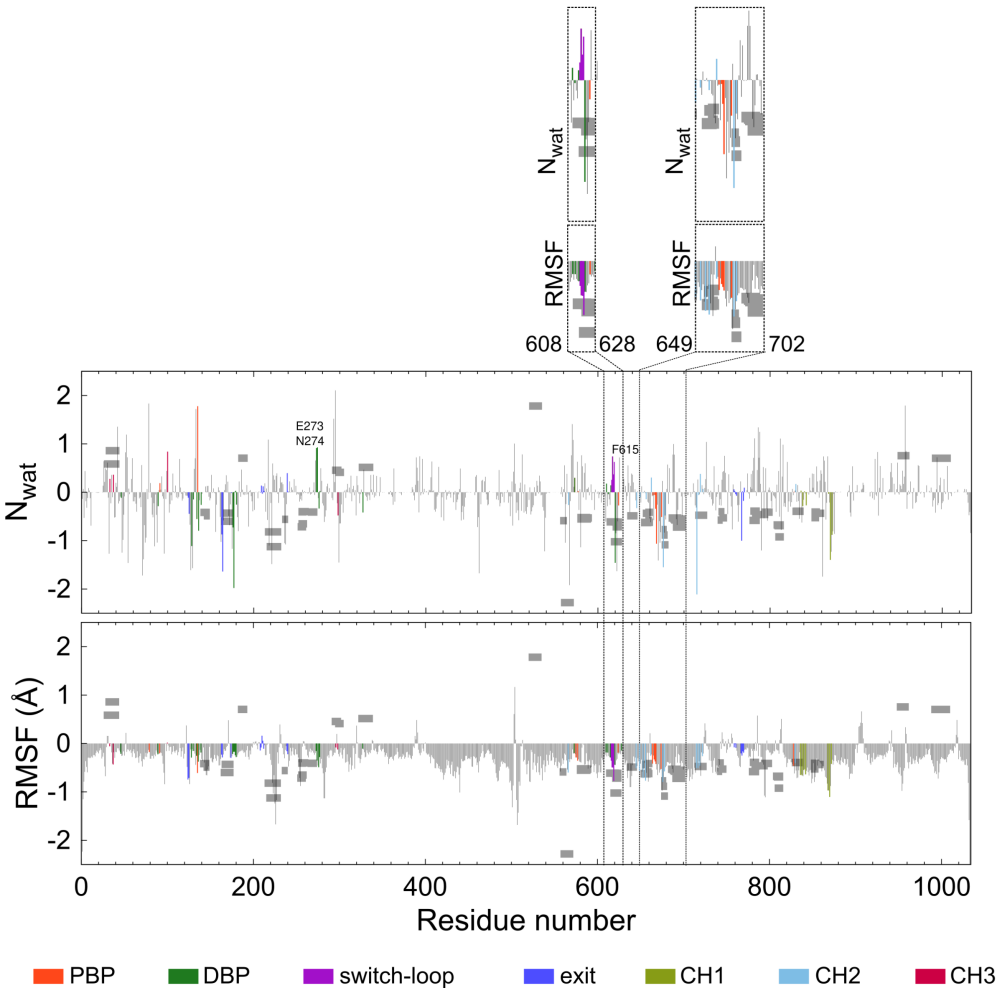

Figure 4.6: Difference in first hydration shell (N_{wat}) and RMSF between $AcrB_{WT}$-CIP-PAβN and apo $AcrB_{WT}$ (based on MD data from Pose 1 in Fig. 4.5). See Fig. 4.4 for details.

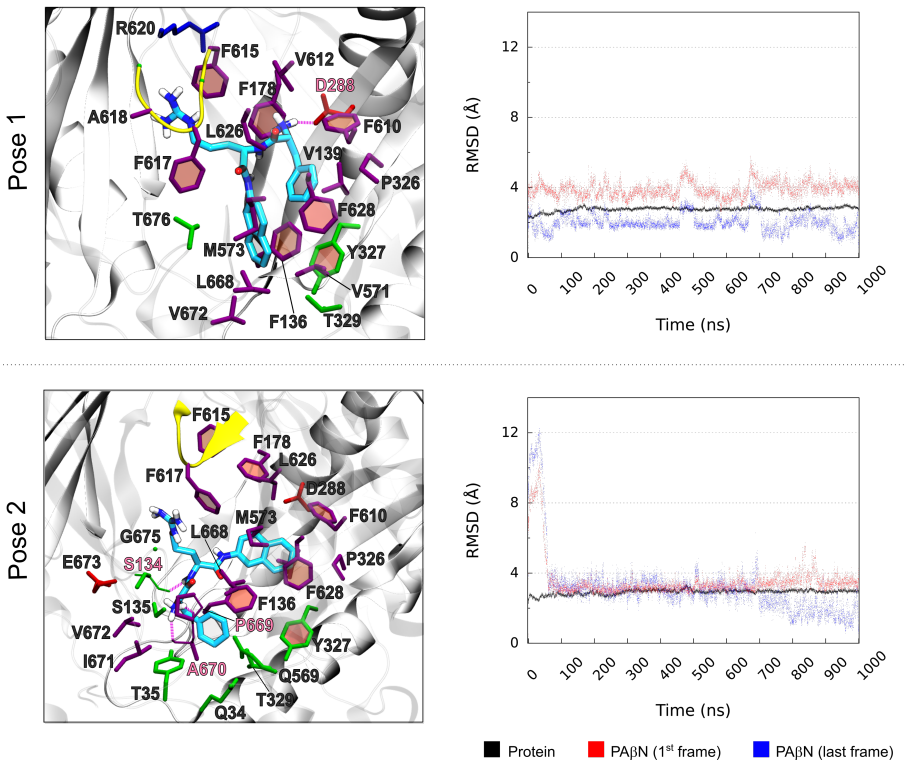

Figure 4.7: Representative binding poses and RMSDs of AcrB$_{G288D}$-PAβN. See Fig. 4.3 for details; Table 4.7 for a list of direct and water-mediated hydrogen bonds involving PAβN; Table 4.9 for the RMSD of each pose with respect to reference structure PDB:4U95.

4.3.4 AcrB$_{G288D}$-PAβN

As in the wt protein, also in the G288D mutant a significant contribution to the stabilization of PAβN comes from the residues of the HT and of the region around the switch loop. Such residues are indeed involved in stacking with the aromatic groups of the inhibitor, as well as in cation-π interactions with its guanidino and amino groups (see Pose 1 in Fig. 4.7 for a representation of the binding pose). These interactions, possibly promoted by the stable hydrogen bonds formed by the amino group of PAβN with D288 (see Table 4.7), were not detected in AcrB$_{WT}$-PAβN (Fig. 4.3, Pose 1); thus, they provide an additional contribution to the stabilization of the inhibitor specifically for the G288D mutant. Additional contacts not observed in AcrB$_{WT}$-PAβN are formed with part of the PC1/PC2 cleft (such as L668), while interactions with the PBP/DBP interface, present in AcrB$_{WT}$-PAβN, are not retained.

Although small differences in the flexibility of the binding sites emerged from the comparison of the RMSFs of AcrB$_{G288D}$-PAβN and AcrB$_{WT}$-PAβN (both considered in the T state, see Section 4.2) (Fig. 4.8), higher hydration levels were detected in AcrB$_{G288D}$-PAβN within the DBP and particularly at residues around D288, which include F178 and adjacent residues in PN2. These findings are in good agreement with HDX-MS data (Fig. 4.1b).

A comparison of the binding poses in AcrB$_{G288D}$-PAβN (Fig. 4.7) reveals a strong contribution to the stabilization of the system from stacking of the aromatic groups of the inhibitor with residues of the HT, in analogy to our findings in AcrB$_{WT}$-PAβN (Fig. 4.3) and AcrB$_{WT}$-CIP-PAβN (Fig. 4.5). Additional stabilization comes from cation-π interactions involving the

Ligand	Pose	H-bonds		Water-mediated interactions	
		Residue	Occupancy (%)	Residue	Occupancy (%)
PAβN	1	D288	100	D288	65
		G616	82	Q176	27
		F617	21		
	2	S134	95		
		I671	67		

Table 4.7: Intermolecular hydrogen bonds (H-bonds) and water-mediated interactions involving PAβN in $AcrB_{G288D}$-PAβN. See Table 4.3.2 for details and Fig. 4.7 for representative binding poses.

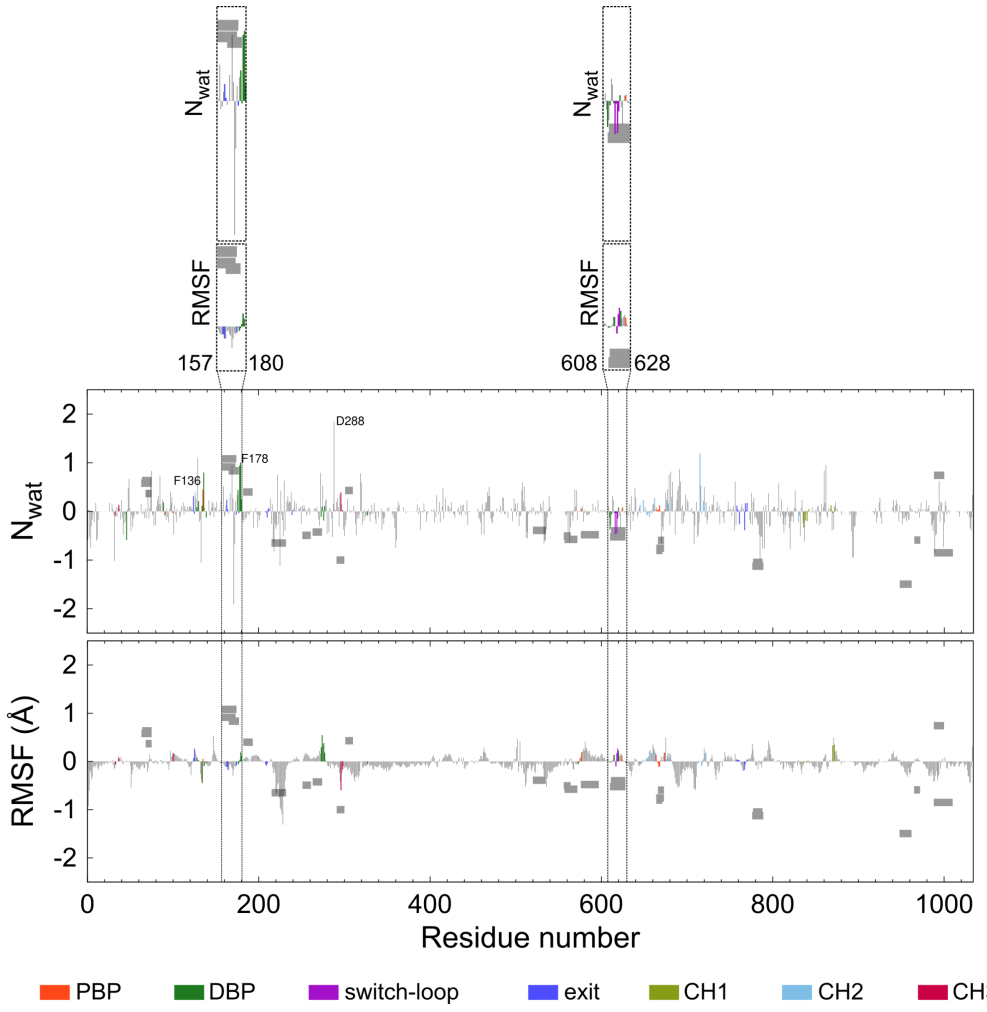

Figure 4.8: Difference in first hydration shell (N_{wat}) and RMSF between $AcrB_{G288D}$-PAβN and $AcrB_{WT}$-PAβN (based on MD data from Pose 1 in Fig. 4.7). Both $AcrB_{G288D}$-PAβN and $AcrB_{WT}$-PAβN were considered in the T state (see Section 4.2). See Fig. 4.4 for details.

guanidino group of PAβN and residues belonging or proximal to the switch loop, as well as the amino group of the inhibitor and F178 in Pose 1. Moreover, contacts are formed between PAβN and the substituted residue D288, involved *e.g.* (in Pose 1) in hydrogen bonds with the amino group of PAβN.

In both poses, stabilizing interactions further involve residues of the PC1/PC2 cleft, such as π-stacking with aromatic groups of PAβN (Pose 1, 2) or hydrogen bonds between the amino group of the inhibitor and the backbone of residues P669 and A670 (Pose 2) (see Fig. 4.7).

From the analysis of the occupancy levels, it emerged that PAβN forms stable hydrogen bonds with residues D288 and the switch loop in Pose 1 and with residues S134 and I671 in Pose 2 (Table 4.7). According to HDX-MS data, several among these residues belong to protected peptides (segments 611-629, 664-671). Moreover, peptide 291-300, adjacent to position 288, is also protected. Taken together, these data suggest that interactions with PAβN may significantly contribute to the protection of part of the DBP, as reported for the WT protein.

4.3.5 AcrB$_{\mathbf{G288D}}$-CIP-PA$\boldsymbol{\beta}$N

As in AcrB$_{\mathrm{WT}}$, CIP and PAβN are involved in direct interactions through hydrogen bonds that involve the amino group of the inhibitor and the carboxylic group of CIP (see Pose 1 in Fig. 4.9 for a representation of the binding pose; see also Table 4.10). The amino group of the inhibitor forms cation-π interactions with residue F178, while its guanidino group is oriented towards D288. Additional stabilization comes from the stacking of the aromatic groups of PAβN with the lower part of the HT (F136, Y327) and the cleft (segment 668-670). CIP is also implicated in stacking interactions with residues close to the switch- loop (such as F615), as well as in interactions with hydrophobic residues proximal to the HT (I277, V612). In analogy to AcrB$_{\mathrm{WT}}$-CIP-PAβN (4.5, Pose 1), therefore, contacts with the switch loop are retained, but the interaction with the PBP/DBP interface is weakened. Moreover, stabilizing interactions also contribute some residues of the PC1/PC2 cleft representing the entering gate towards the PBP, as well as the mutated residue D288.

From the comparison of the flexibility and hydration properties of AcrB$_{\mathrm{G288D}}$-CIP-PAβN and AcrB$_{\mathrm{WT}}$-CIP-PAβN (Fig. 4.10), it emerges that the switch loop is considerably more rigid and dehydrated in the mutant. A net increase in hydration and flexibility is also detected for part of the PN2 portion of the DP (including segment 178-182, involved in interactions with the substrates; see Fig. 4.9). These data agree with the stabilization of the switch loop and the increase in hydration of PN2 emerged from HDX-MS analyses (Fig. 4.1b).

As for AcrB$_{\mathrm{WT}}$-CIP-PAβN, direct interactions between the two substrates that involve the carboxylic and carbonyl group of CIP are present in both the binding modes detected in AcrB$_{\mathrm{G288D}}$-CIP-PAβN (Fig. 4.9 and Table 4.10). In addition, π-stacking between one or both substrates and the HT and the switch loop were detected. Further stabilization is provided by cation-π interactions established between the amino group of PAβN and F178 (Pose 1). Importantly, D288 also contributes to stabilize the complex by forming high-occupancy hydrogen bonds with PAβN (Pose 1) or CIP (Pose 2) (Table 4.8). Another common feature of both poses is the π-stacking formed with residues of the PC1/PC2 cleft. A major difference regards instead the interaction with the PBP/DBP interface, which is indeed present only in Pose 2 and involves the phenylalanine and arginine moieties of PAβN and the carboxylic group of CIP. In Pose 1 PAβN is located inside the HT and CIP interacts with regions proximal to the switch loop and to the upper part of the DBP (including, for example, residues I277 and the

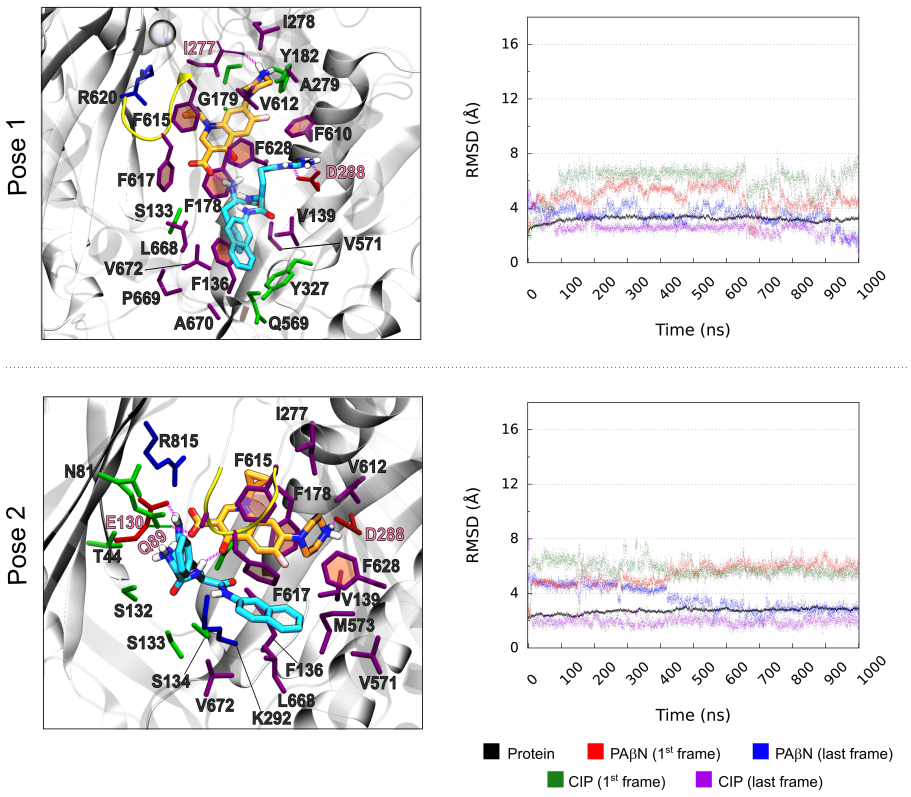

Figure 4.9: Representative binding poses and RMSDs of $AcrB_{G288D}$-CIP-PAβN. See Fig. 4.3 for details; Table 4.8 for a list of direct and water-mediated hydrogen bonds established by each ligand with AcrB; Table 4.10 for high-occupancy hydrogen bonds between the ligands; Table 4.9 for the RMSD of each pose with respect to reference structure PDB:4U95.

Ligand	Pose	H-bonds		Water-mediated interactions	
		Residue	Occupancy (%)	Residue	Occupancy (%)
PAβN	1	D288	100	D288	100
	2	E130	94	E130	68
				K131	67
				T91	64
				Q89	31
CIP	1	I277	32	E152	41
		S180	31		
		D276	23		
	2	D288	100	D288	100
		Q89	100		
		E826	69		
		Y77	41		

Table 4.8: Intermolecular hydrogen bonds (H-bonds) and water-mediated interactions involving ligands in $AcrB_{G288D}$-CIP-PAβN (MDs data). See Table 4.3.2 for details and Fig. 4.9 for representative binding poses.

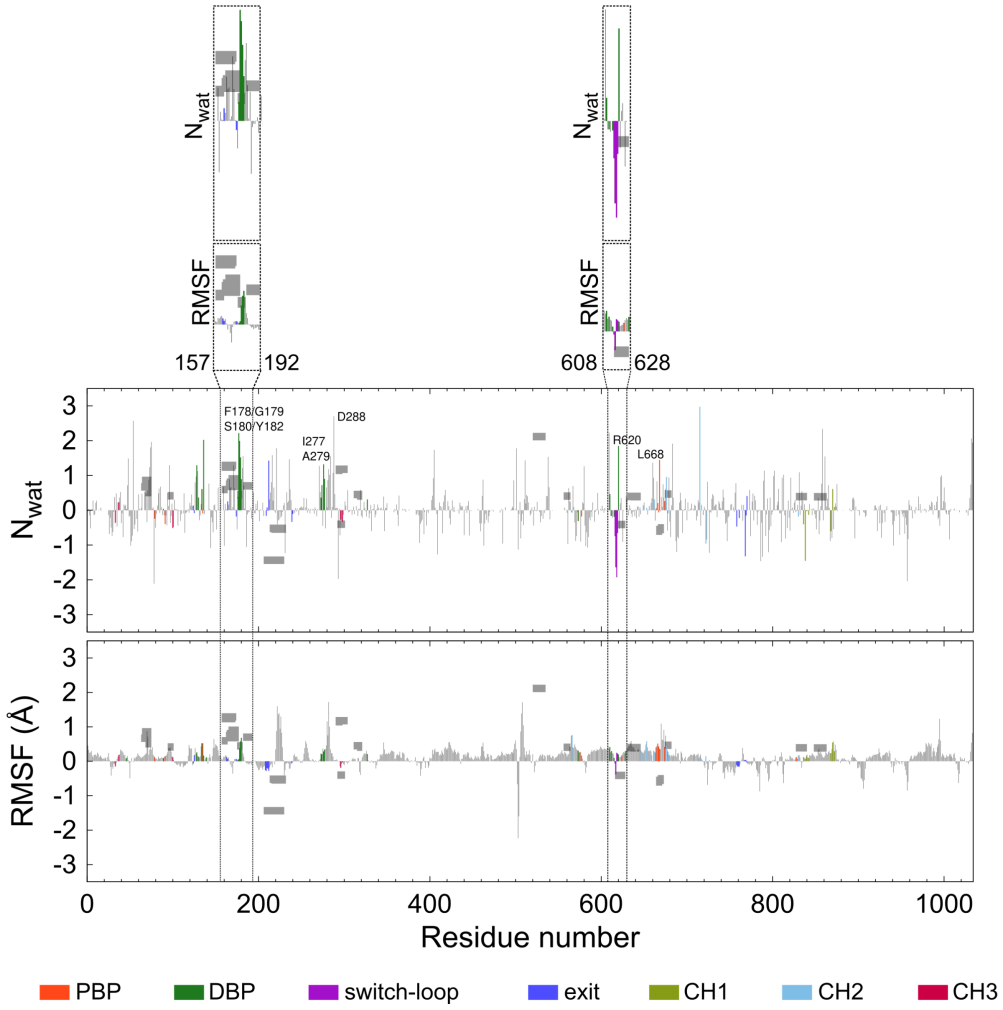

Figure 4.10: Difference in first hydration shell (N_{wat}) and RMSF between $AcrB_{G288D}$-CIP-PAβN and $AcrB_{WT}$-CIP-PAβN (based on MD data from Pose 1 in Fig. 4.9). Both $AcrB_{G288D}$-CIP-PAβN and $AcrB_{WT}$-CIP-PAβN were considered in the T state (see Section 4.2). See Fig. 4.4 for details.

System	Pose	RMSD (Å)	
		T monomer	T monomer: PDP, DBP, switch loop
$AcrB_{WT}$-PAβN	1	2.5	2.2
	2	3.2	3.3
	3	2.9	2.9
$AcrB_{WT}$-CIP-PAβN	1	2.7	2.0
	2	2.7	2.2
	3	2.6	2.3
	4	3.2	2.3
$AcrB_{G288D}$-PAβN	1	2.5	2.5
	2	2.8	2.1
$AcrB_{G288D}$-CIP-PAβN	1	2.8	2.6
	2	2.3	1.9

Table 4.9: Backbone RMSD of each pose with respect to the X-ray crystal structure 4U95 of E. coli AcrB (resolution: 2.0 Å). Calculations were performed on the T monomer of the protein (residues 1-1033) and on a sub-selection composed by the PBP, DBP and the switch loop (see Table 4.4 for a definition of these regions). For each pose, the RMSD was computed on the centre of the representative cluster of the last 300 ns of MD simulation.

segment 178-182).

In analogy with $AcrB_{WT}$-CIP, comparison between the binding poses of $AcrB_{G288D}$-CIP-PAβN and $AcrB_{G288D}$-CIP (Fig. 4.11) revealed that the co-presence of PAβN in the DBP determines a shift in the binding region of CIP. In the binary complex, indeed, CIP tends to occupy the HT, interacting with residue D288 as well as with several hydrophobic residues like F136 and F628 (Fig. 4.11). Less stabilizing interactions are formed with residues belonging to the upper part of the DBP, which instead form several contacts with CIP in the ternary complex. Similar to $AcrB_{WT}$-PAβN, analysis of hydrogen bonds occupancy revealed that both ligands establish very stable interactions with D288, both direct and water-mediated (Table 4.8). Additional interactions are mainly formed with residues E152 and S180 in Pose 1, and with residues Q89, E130 and K131 in Pose 2 (Table 4.8). Unfortunately, evaluation of the correlations between these interactions and HDX-MS results is not straightforward for this system, due to the lack of HDX-MS coverage for the segments involved in interactions with the ligands.

4.4 Discussion

From the analyses of our MD trajectories, a good correlation emerged between computational data and HDX-MS assays (see reference [128]). In $AcrB_{WT}$, HDX-MS revealed that the presence of PAβN considerably restricts the dynamics of several regions of the PBP and DBP, including the switch loop (Fig. 4.1a). These data are in good agreement with MD analyses, which revealed that binding of PAβN is accompanied by an overall rigidification of the protein that involves large patches of the DBP, PBP, switch loop, as well as the exit channel gate (EG), CH1, and CH2 channels. In particular: i) regions containing residues belonging/adjacent to the switch loop that were found to directly interact with the PAβN become more rigid in its presence, the extent of HDX protection upon PAβN binding (as revealed by the HDX-MS

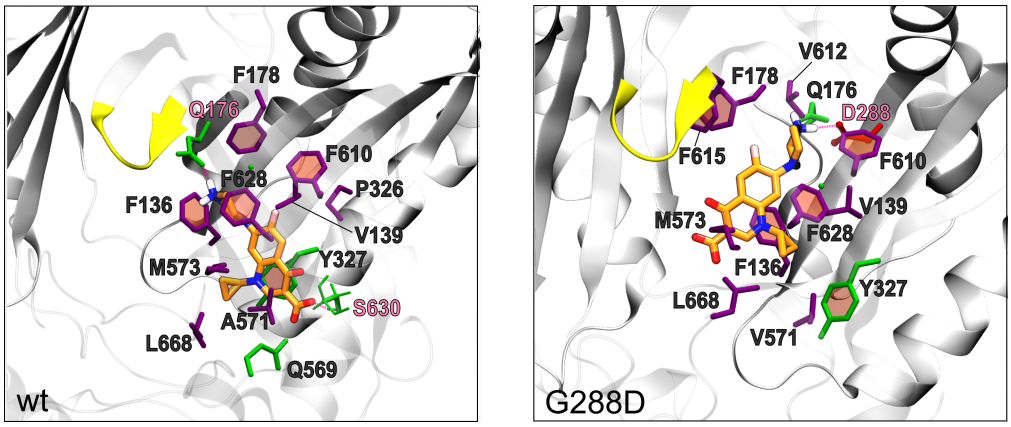

Figure 4.11: Representative binding poses of $AcrB_{WT}$-CIP and $AcrB_{G288D}$-CIP. See Fig. 4.3 for details.

System	Pose	Involved groups		Occupancy (%)
		CIP	PAβN	
$AcrB_{WT}$-CIP-PAβN	1	CO	NH_3^+ (Phe)	86
		CO	NH (Phe)	86
	2	CO_2	NH (β-napth.)	62
		CO	NH (β-napth.)	47
		CO	NH_2 (Arg)	55
	3	CO	NH (Phe)	79
	4	CO	NH (Arg)	32
$AcrB_{G288D}$-CIP-PAβN	1	CO	NH (Phe)	57
		CO	NH (Phe)	30
	2	CO_2	NH_3^+ (Phe)	93
		CO	NH (Arg)	68
		CO	NH_2 (Arg)	43

Table 4.10: Occupancies of intermolecular hydrogen bonds between the ligands in $AcrB_{WT}$-CIP-PAβN and $AcrB_{G288D}$-CIP-PAβN. For a better identification of the functional groups of PAβN, their moieties of belonging (Phe, Arg and β-napthilamide) have been indicated in parentheses. Analyses have been conducted on the last 300 ns of each simulation (see Section 4.2). Only interactions with occupancy higher than 20% have been reported.

data) correlating with the formation of hydrogen bonds between the EPI and residues of (and nearby) the DBP; ii) the switch loop itself (residues 615-620) features moderately enhanced hydration, whereas the nearby segments (residues 612-614 and 621-624) are overall dehydrated with respect to apo $AcrB_{WT}$. This supports an interaction between PAβN and the switch loop region, which could be a key factor in mediating the mode of action of this EPI. More generally, the structural stabilisation that occurs upon PAβN binding might prevent local, as well as distal, functional movements that are key to substrate efflux along the transport pathway.

The mode of action of PAβN was further investigated by considering its activity in the presence of the antibiotic ciprofloxacin. HDX-MS profiles for $AcrB_{WT}$-PAβN and $AcrB_{WT}$-CIP-PAβN revealed that the presence of ciprofloxacin did not alter the effect of the inhibitor (Fig. 4.1(a)). From the analyses of MD trajectories of $AcrB_{WT}$-CIP-PAβN, it emerged that both drugs stably bind to the DBP within the T-state monomer, with PAβN partly occupying the HT and ciprofloxacin lying in proximity of the PBP/DBP interface (Fig. 4.5, Pose 1). Several interactions contribute to stabilize this configuration, including stable intermolecular hydrogen bonds between the two ligands (Table 4.10). The simultaneous binding of CIP and PAβN has similar effects as the binding of the inhibitor only on the flexibility and hydration of the protein (Fig. 4.6), in corroboration with HDX-MS results (Fig. 4.1a). Moreover, in analogy to $AcrB_{WT}$-PAβN, HDX-MS protection data of$AcrB_{WT}$-CIP-PAβN correlate with the formation of hydrogen bonds involving the ligands and several residues of the DBP (see Subsection 4.3.3).

Overall, these data agree with a model for inhibitor action, which has been proposed to work by trapping AcrB in a conformation, possibly a T-like state, which prevents adequate functional rotation and substrate transport [37, 155]. Previous computational studies had indeed indicated that PAβN can restrict the conformational dynamics of AcrB in *K. pneumoniae* [156] and of the RND transporter AdeB in *A. baumannii* [157]. Specifically, this EPI was proven to considerably affect the dynamics of *K. pneumoniae* AcrB, preventing the completion of the extrusion process [156].

The simultaneous binding of ciprofloxacin and PAβN in the DBP is in good agreement with experimental data. Indeed, according to titration results the EPI could not effectively outcompete ciprofloxacin binding from a $AcrB_{WT}$-PAβN complex. Overall, our data support the hypothesis that PAβN does not compete or prevent antibiotic binding (competitive inhibition). Instead, we propose that it inhibits AcrB function by enforcing a more restrained state, thus, reducing the frequency and magnitude of the conformational changes within the substrate translocation path, its effectiveness being substrate dependent.

Application of the HDX-MS protocol to $AcrB_{G288D}$ revealed that the G288D mutation caused increased HDX for several peptides spanning the PN2 region of the protein, but decreased HDX within the PC1/PC2 regions and the connecting loop (Fig. 4.1b). These effects were detected in all three substrate conditions tested (CIP, PAβN, and CIP-PAβN). Upon the inspection of MD trajectories, PAβN was found to bind to the hydrophobic trap of $AcrB_{G288D}$, interacting with the mutated D288 residue through the formation of direct and water-mediated hydrogen bonds (Fig. 4.7, Pose 1). Interactions with the aromatic residues of the trap involve hydrophobic stacking as well as cation-π attraction, not observed in $AcrB_{WT}$-PAβN and possibly promoted by the direct interaction of the inhibitor with residue D288 (see Subsection 4.3.4). Similar interactions are also formed with residues of the switch loop or the surrounding region, in analogy to $AcrB_{WT}$-PAβN.

Moreover, in accordance with HDX-MS, the switch loop and the surrounding region undergo further dehydration in $AcrB_{G288D}$-PAβN (Fig. 4.8). As in the case of $AcrB_{WT}$-PAβN,

there is a good overlap between the residues of the DBP protected in the HDX-MS assay – which include peptides containing D288 - and those involved in high-occurrence interactions with the EPI (see Subsection 4.3.4). These data, together with the direct interactions detected between the inhibitor and the region of the switch loop, support the hypothesis that stabilisation of the latter has a role for the mode of action of PAβN both in $AcrB_{WT}$ and in $AcrB_{G288D}$.

Similar conclusions emerged from the comparison between $AcrB_{G288D}$-CIP-PAβN and $AcrB_{WT}$-CIP-PAβN. Indeed, MD simulations of the former complex revealed that, even upon G288D substitution, ciprofloxacin and PAβN can stably occupy the DBP at the same time (Fig. 4.9). As in $AcrB_{WT}$-CIP-PAβN, stabilizing interactions include several contacts with the protein (also involving D288) as well as intermolecular hydrogen bonds between the two compounds (Table 4.10). These data advocate that $AcrB_{G288D}$ is inhibited by PAβN in a similar manner as $AcrB_{WT}$. Our findings endorse the theory that RND-pump inhibitors act through an "altered-dynamics" mechanism, obstructing the translocation of substrates rather than preventing their binding and recognition.

Chapter 5

Computational structural analysis of the fluoroquinolone resistant AcrB variant from *Salmonella* Typhimurium

5.1 Introduction

Salmonella species include important gram-negative pathogens, especially relevant in nosocomial settings [139]. Treatment of their infections has become considerably challenging due to the occurrence of MDR [140]. In this regard, expression of multidrug efflux pumps has been identified as one of the key mechanisms [31, 141, 142].

In *S.* Typhimurium, the AcrAB-TolC efflux pump provides major contributions to the efflux of antimicrobials [142]. Importantly, the occurrence of mutations altering the specificity of this RND transporter has been detected in this species [143]. The particular mutation, which arose during the antibiotic treatment of a patient with a complex *S.* Typhimurium infection [144, 145], has resulted in a substitution, G288D, within the AcrB transporter. This mutation is responsible for altering the the occurrence of MDR. It was indeed proved to increase the MIC to ciprofloxacin over 60 folds, while incrementing sensitivity to minocycline and doxorubicin [143].

Due to the clinical relevance of *S.* Typhimurium AcrB and of its G288D variant (hereafter STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$), studies on their functioning mechanism and substrate specificity would be of considerable importance. However, such investigations have been limited by the lack of experimental structural data, due to difficulties in the crystallization of these transporters [146]. Our present knowledge of AcrB is indeed mainly based on the structural data available for the *E. coli* orthologue (see Chapter 1, Section 1.4), which presents a 94.7% sequence identity with respect to STmAcrB$_{\mathrm{WT}}$. Despite the high similarity of the two sequences, a detailed structural knowledge of STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$ is crucial to understand their substrate specificities.

Only recently, the structure of STmAcrB$_{\mathrm{G288D}}$ has been resolved through cryo-EM (resolution: 4.6 Å, see reference [146]; see also Fig. 5.1). From the analysis of the density map, it emerges that the architecture of the receptor closely resembles that of the *E. coli* orthologue. Consistenly, superposition of the three-dimensional structures of STmAcrB$_{\mathrm{G288D}}$ (obtained through homology modelling, see reference [146]) and *E. coli* wt AcrB (hereafter EcAcrB$_{\mathrm{WT}}$), both fitted in the density map, revealed considerable similarities in the ternary and quaternary

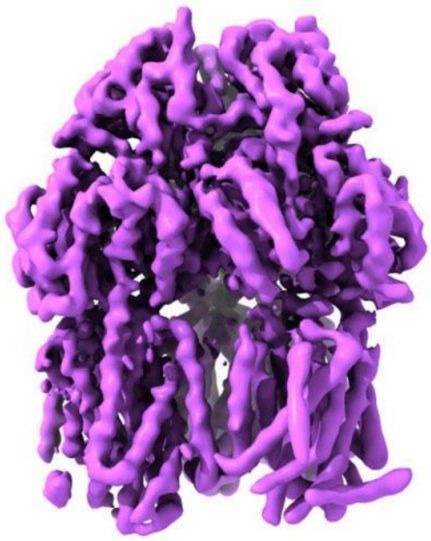

Figure 5.1: Cryo-EM map of STmAcrB$_{\mathrm{G288D}}$ (see reference [146]; resolution: 4.6 Å).

structure (Fig. 5.2). From the analysis of the STmAcrB$_{\mathrm{G288D}}$ structural model, it emerged that most residues differing in the sequence of *E. coli* and *S.* Typhimurium (see Fig. 5.3) are located on the external surface of the protein. In particular, several of them belong to regions that form crystal contacts in *E. coli* [146]. This may account for the difficulties in the crystallization of STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$, despite their high sequence identity with respect to the *E. coli* orthologue.

In agreement with previous literature, comparison of the structural models of STmAcrB$_{\mathrm{G288D}}$ and EcAcrB$_{\mathrm{WT}}$ highlighted some structural differences in the surroundings of the mutated residue 288, located in proximity of the hydrophobic trap (see Fig. 5.2G; see also Chapter 1, Section 1.4). The G288D substitution is indeed predicted to increment the hydration in the hydrophobic trap and to cause a rearrangement of its residues [143]. In order to better understand the effect of the considered mutation on the structure and dynamics of *S.* Typhimurium AcrB, we conducted a computational investigation complementary to the described cryo-EM data. Through homology modeling, we generated an ensemble of structural models of STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$ using several crystal structures of EcAcrB$_{\mathrm{WT}}$ as templates. Models of STmAcrB$_{\mathrm{G288D}}$ were further refined against the cryo-EM density map. The generated structures were then used to perform all-atom MD simulations, with the aim of analysing the structural features of the DBP and the substrate pathway and to assess the effects of the G288D substitution. For a better characterization, an ensemble of crystal structures of EcAcrB$_{\mathrm{WT}}$ was used as a reference. The computational methods used for this work are described in Section 5.2 of the present Chapter, while Section 5.3 is dedicated to the obtained results. A final discussion of our work is reported in Section 5.4.

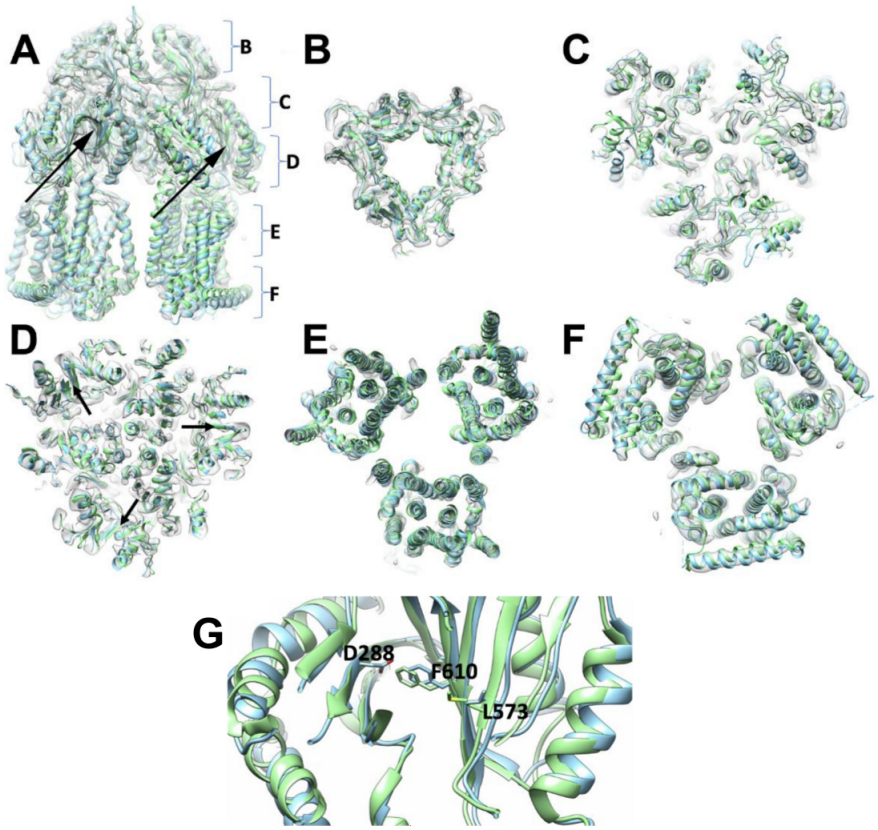

Figure 5.2: Side view (A) and horizontal slices (B-F) of STmAcrB$_{\mathrm{G288D}}$ (cyan) and EcAcrB$_{\mathrm{WT}}$ (light green). Both structures have been fitted in the STmAcrB$_{\mathrm{G288D}}$ cryo-EM map (semi-transparent grey, see Fig. 5.1). The relative position of slices (B-F) are indicated in (A). The position of residue 288 within the structure is indicated with black arrows in (A) and (D). A representation of the DBP, in proximity of residue 288, is reported in (G). Image adapted from reference [146].

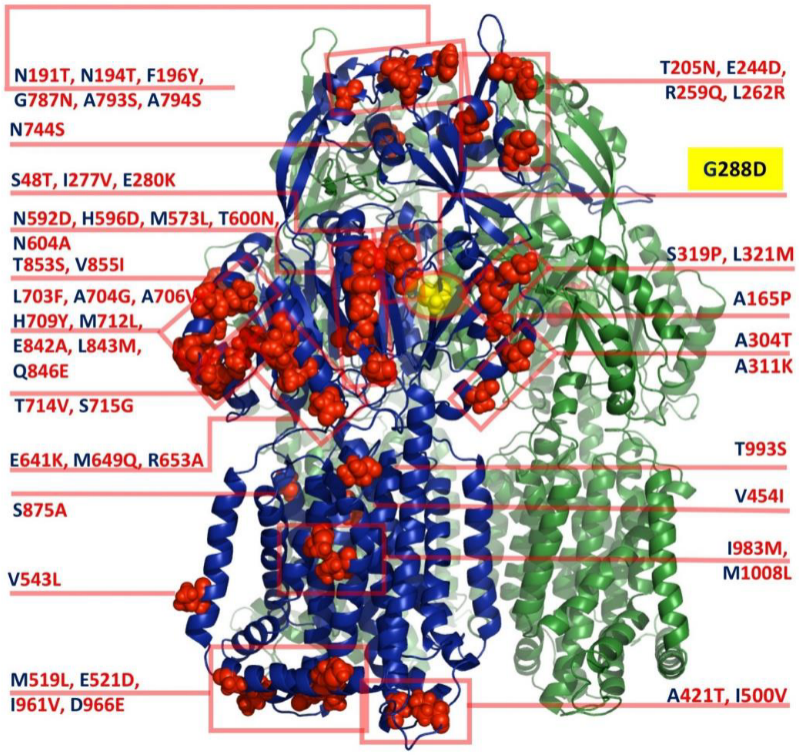

Figure 5.3: Mapping of the sequence differences between *S.* Typhimurium and *E. coli* AcrB. Differing residues are represented in red in a single protomer. The labels indicate single-letter residue code for the differing residues in *E. coli* AcrB (in blue), alongside with the equivalent positions and substituted side-chains in *S.* Typhimurium (in red). The location of mutation G288D is highlighted in yellow. Image from reference [146].

5.2 Methods

5.2.1 Homology modelling

Three homology models were built of STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$ using different X-ray crystal structures of wild type AcrB from *E. coli* (hereafter EcAcrB$_{\mathrm{WT}}$) (PDB IDs 2J8S [112], 4DX5 and 4DX7 [45]) as templates. The amino acid sequences of both EcAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{WT}}$ were obtained from the Uniprot database (Uniprot IDs: P31224 and Q8ZRA7, respectively) [74]. The absence of gaps was verified through a sequence alignment with ClustalOmega [147]. The homology models were generated using Modeller 9.3 [76], each having a MOLPDF score greater than 1.5× 105, and included the full range of residues (1–1033) in every template. The homology models of STmAcrB$_{\mathrm{G288D}}$ were further energy-minimized into the experimental C1 cryo-EM map presented in reference [146] with the program Flex-EM [148]. We performed structural optimization of the models for up to 40 iterations, and we ranked the final structures based on their cross-correlation function (hereafter CCF) (see Table 5.1).

5.2.2 Molecular dynamics simulations

The homology models of STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$ were used as starting structures to perform all-atom MD simulations. Following previous work [60, 108, 143], we simulated the truncated structure including only the porter domain and a few residues at the interface with the TM domain (namely, residue segments 32-335 and 564-870), imposing positional restraints on the C_α atoms of the residues found within 5 Å from the bottom of the structure (weight of the restraints: 1 kcal/mol). Those residues involved were in Subdomain PN1 (V32, A33, Q34, T37, I38, and A39), subdomain PN2 (A297, N298, A299, T330, P331, and F332), subdomain PC1 (L564, P565, D566, K632, D633, W634, P638, G639, E640, A670, I671, V672, T676, A677, and T678) and subdomain PC2 (P710, D711, L712, G838, E839, A840, Q865, E866, and R867). The selected portion of the protein was inserted in a truncated octahedron filled with 0.15 KCl aqueous solution, setting the minimum distance between the protein and the edge of the box to 16 Å. The topology and the initial coordinate files were created through the leap module of AMBER18 [86]. Protein and water were represented using the ff14SB force field [114] and the TIP3P model [115], respectively, while the parameters for the ions were retrieved from [149]. The system was enclosed in a truncated octahedron filled with 0.15 M KCl aqueous solution, and the minimum distance of the protein and the border of the box was set to 16 Å. The MD simulations of each system were done according to the following procedure. Firstly, we performed a multi-step structural relaxation combining steepest descent and conjugate gradient methods, using the pmemd module of AMBER18, as described in previous publications [47, 60, 135, 149]. The relaxation was followed by two MD simulations runs to heat the system from 0 to 310 K: i) from 0 to 100 K in 1 ns under constant-volume conditions and with harmonic restraints (k = 1 kcal·mol−1·Å−2) on the heavy atoms of both the protein and the lipids; ii) from 100 to 310 K in 5 ns under constant pressure (set to a value of 1 atm) and with restraints on the heavy atoms of the protein and on the z coordinates of the phosphorous atoms of the lipids to allow membrane rearrangement during heating. Next, we performed a series of 10 equilibration steps to equilibrate the box dimensions. Each step was of 100 ps in duration (total 1 ns) and was carried out under isotropic pressure scaling conditions through the Berendsen barostat. The Langevin thermostat was also used to maintain the temperature constant, with a collision frequency of 1 ps−1. Finally, for every system we

Template (PDB ID)	CCF_{init}(a.u.)	CCF_{final}(a.u.)	RMSD (Å)
2J8S	0.73	0.75	1.2
4DX5	0.73	0.75	1.1
4DX7	0.73	0.75	1.2

Table 5.1: Values of the cross-correlation function (CCF) obtained through Flex-EM [148] for the homology models of STmAcrB$_{G288D}$, before and after the optimization inside the cryo-EM map. The RMSD of the optimized models with respect to the starting ones is also reported in the last column.

performed three independent MD simulations, each with a production run of 150 ns in length. Time steps of 0.5 fs and 2 fs were used during the heating and equilibration stages, respectively. In the production run a time step of 4 fs was adopted under an isothermal-isobaric ensemble after hydrogen mass repartitioning [98]. Moreover, the lengths of all the R-H bonds were constrained with the SHAKE algorithm. Coordinates were saved every 100 ps. Long-range electrostatic forces were evaluated with the particle mesh Ewald (PME) algorithm, with a non-bonded cut-off of 9 Å.

5.2.3 Post-processing of MD trajectories

The MD trajectories of STmAcrB$_{WT}$ and STmAcrB$_{G288D}$ were firstly processed by performing a cluster analysis with the cpptraj module of AMBER18. For each trajectory, we considered only the last 140 ns of the production run, where the RMSD of the protein with respect to the first frame is fairly constant (Fig. 5.4). Every trajectory was subjected to three clustering procedures, in each of which the distance RMSD metric was applied to the DBP of a different monomer of AcrB, generating 100 clusters. In this way, we obtained 300 clusters per trajectory, divided in three equal subsets (1 subset per monomer). For each subset, the representative centroid structures of all clusters were used to perform several analyses aimed at assessing how the size and shape of the DBP are affected by the G288D mutation. To this end, we firstly estimated its volume of in the L, T, and O monomers of both STmAcrB$_{WT}$ and STmAcrB$_{WT}$.

The same analysis was then performed on 5 experimentally derived crystal structures of the EcAcrB$_{WT}$, which were chosen as reference structures to identify variations between *E. coli* and STmAcrB$_{WT}$. These structures have PDB IDs 4DX5, 4DX7 [45], 2J8S [112], 2I6W (the last being a symmetric LLL structure) [150] and 6BAJ (the structural model recently derived from cryo-EM data by Qiu *et al.* [53]). The volume calculations were performed using the POVME 2.0 software [151], adopting a grid spacing of 0.5 Å. Additional analyses were conducted to better characterize of the impact of the G288D mutation. These included the calculation of the gyration radius of the DBP, the number of (pseudo)contacts between the PC1 and PC2 subdomains and the number of waters in the first and second solvation shell of residue 288. Such analyses were conducted on every protomer of AcrB. Calculations of the radius of gyration and of the number of (pseudo)contacts were carried out using in-house tcl scripts and performed on the cluster representatives of STmAcrB$_{WT}$ and STmAcrB$_{G288D}$, as well as on the reference structures of the EcAcrB$_{WT}$ The radius of gyration was computed for three different regions of the DBP: the whole DBP (S46, Q89, S128, E130, E134, F136, V139, Q176, L177, F178, S180, E273, N274, D276, Y327, L573, F610, V612, R620, F628), the hydrophobic trap (hereafter HP trap) (F136, V139, F178, Y327, L573, F610, V612, and F628) and the upper DBP (S46, Q89, S128, E130, Q176, L177, G179, S180, E273, N274, D276 and R620). As to the number of (pseudo)contacts, it was calculated by using a distance

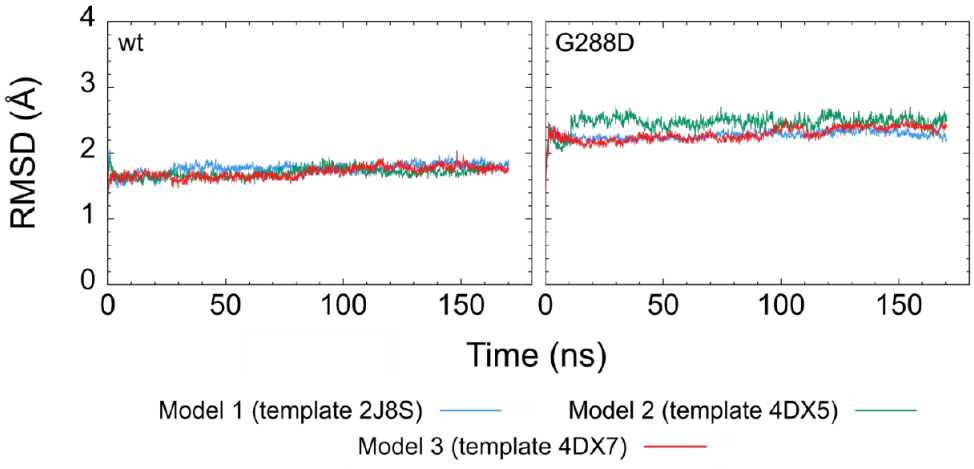

Figure 5.4: Protein RMSD calculated along the MD trajectories of STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$. Only the Cα atoms were considered.

cut-off of 10 Å among the Cα carbons of selected regions of PC1 (segment 571-667) and PC2 (segments 679-721, 822-859). Regarding the first and second water shells of residue 288, these regions were defined using distance cut-offs of 3.4 Å and 5 Å, respectively. Calculations were performed on the last 140 ns of every MD trajectory of STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$, using the cpptraj module of AMBER18. Moreover, we monitored the Loose/Tight/Open (LTO) asymmetry of STmAcrB$_{\mathrm{G288D}}$ along the MD trajectories. To perform this analysis, we used as a reference the EcAcrB$_{\mathrm{WT}}$ crystal structure with PDB ID 4DX7 [19], in which the protein is found in the LTO state. Thus, for each frame in the last 140ns of the MD production run, we calculated the RMSD of each conformer of the mutant with respect to every conformer of the *E. coli* reference structure. Only the Cα atoms were considered for this calculation.

5.3 Results

In order to evaluate the structural properties of *Salmonella* Typhimurium AcrB and the impact of the G288D mutation, several structural models of STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$ were derived through homology modelling, using three different X-ray crystal structures of EcAcrB$_{\mathrm{WT}}$ as templates (see Section 5.2). Models of STmAcrB$_{\mathrm{G288D}}$ were further minimized against the cryo-EM map presented in reference [146]. Their accuracy was evaluated through the CCF which, for each model, improved slightly with respect to the starting model (Table 5.1). Accordingly, optimization against the cryo-EM map did not cause major structural changes in each model (last column in Table 5.1).

For both STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$, the obtained structural models were used as starting structures to perform three independent all-atom MD simulations. As reported in Section 5.2, a truncated model of AcrB was used in all simulations, which only included the periplasmic portion of the protein. This protocol, also applied to other works described in this thesis (see Chapter 3), has indeed been largely validated in literature (see Section 5.2) [47, 60, 135, 149].

Despite this approximation, the asymmetric LTO conformation of AcrB reimained well preserved along every simulation, in agreement with previous studies [60, 143]. This was

Homology model	Chain	RMSD (reference structure: *Ec*AcrB$_{WT}$ 4DX7)		
		L-conformer	T-conformer	O-conformer
1 (template: 2J8S)	A	2.1 (0.1)	2.7 (0.1)	3.2 (0.1)
	B	3.0 (0.1)	2.3 (0.1)	4.1 (0.1)
	C	3.5 (0.1)	3.9 (0.1)	2.2 (0.1)
2 (template: 4DX5)	A	2.4 (0.1)	3.1 (0.1)	3.0 (0.1)
	B	3.0 (0.1)	2.2 (0.1)	3.9 (0.1)
	C	3.2 (0.1)	3.9 (0.1)	2.2 (0.1)
3 (template: 4DX7)	A	2.1 (0.1)	2.8 (0.1)	3.4 (0.1)
	B	3.1 (0.1)	2.3 (0.1)	4.1 (0.1)
	C	3.1 (0.1)	3.5 (0.1)	2.1 (0.1)

Table 5.2: Cross-RMSD of each chain of *S*TmAcrB$_{G288D}$, with respect to every conformer of the *Ec*AcrB$_{WT}$ crystal structure 4DX7 [45]. For each model, the calculation was performed on the last 140 ns of the production run (see Section 5.2); the reported values correspond to the average RMSD and its standard deviation in the Loose (L), Tight (T) and Open (O) conformer.

System	Volume of DBP (Å^3)		
	L-conformer	T-conformer	O-conformer
*Ec*AcrB$_{WT}$[1]	763 (86)	2315 (76)	1094 (76)
*S*TmAcrB$_{WT}$[2]	957 (93)	1534 (163)	991 (96)
*S*TmAcrB$_{G288D}$[2]	807 (78)	1979 (72)	1151 (64)

[1]calculated on experimental reference structures; [2]calculated on representatives of each of the 100 clusters extracted from the MD trajectories.

Table 5.3: Values of the volume of the distal binding pocket (DBP) (standard deviations in parentheses) in each conformer of AcrB, measured on the *E. coli* AcrB$_{WT}$ reference structures and on the MD trajectories of *S.* Typhimurium AcrB$_{WT}$ and AcrB$_{G288D}$ (see Section 5.2).

verified through the calculation of the cross-RMSD of each conformer of our structural models against each conformer of the reference structure (PDB ID: 4DX7; see Section 5.2). According to the obtained results, the asymmetric LTO configuration is also retained in *S*TmAcrB$_{G288D}$ (Table 5.2).

In order to assess the impact of the G288D substitution, we evaluated the volume of the DBP in all conformers of *S*TmAcrB$_{WT}$ and *S*TmAcrB$_{G288D}$. As a reference, the same analysis was conducted on an ensemble of X-ray crystal structures of *Ec*AcrB$_{WT}$ (see Section 5.2). The largest differences are seen in the volume of the distal binding pocket of monomer T (DBP$_T$), which undergoes a significant expansion with respect to the WT Salmonella protein (Table 5.3). Significantly, while relative to the *E. coli* orthologue the volume of the DBP of the *S*TmAcrB$_{WT}$ is about 800 Å^3 smaller, the effect of G288 mutant on the DBP results in an expansion of approximately 450 Å^3 in the T-conformer bringing it closer to that of the *E. coli* orthologue.

A similar trend emerges from the evaluation of the radius of gyration of the DBP, which was also performed on all conformers of the considered systems. Significant variations were indeed detected only for the DBP$_T$ (see Table 5.4). Specifically, the highest variations were observed for the hydrophobic trap of the T conformer, whose increment in *S*TmAcrB$_{G288D}$

			System		
			*Ec*AcrB$_{WT}$[1]	*S*TmAcrB$_{WT}$[2]	*S*TmAcrB$_{G288D}$[2]
Radius of gyration (Å)	L-conformer	Whole	10.2 (0.1)	10.5 (0.1)	10.5 (0.1)
		Upper	9.0 (0.1)	9.1 (0.1)	9.2 (0.1)
		HP Trap	6.4 (0.3)	6.3 (0.2)	6.4 (0.1)
	T-conformer	Whole	10.8 (0.1)	10.7 (0.2)	11.2 (0.1)
		Upper	9.4 (0.1)	8.7 (0.1)	9.2 (0.2)
		HP Trap	7.1 (0.3)	6.9 (0.2)	8.0 (0.2)
	O-conformer	Whole	10.6 (0.1)	10.9 (0.1)	11.0 (0.1)
		Upper	9.9 (0.1)	9.8 (0.1)	9.9 (0.1)
		HP Trap	6.2 (0.0)	6.3 (0.2)	6.5 (0.2)

[1]calculated on experimental reference structures; [2]calculated on representatives of each of the 100 clusters extracted from the MD trajectories.

Table 5.4: Radius of gyration of the DBP calculated for every AcrB conformer. The three regions of the DBP considered in this calculation are indicated as Whole (entire DBP), Upper (upper part of the binding site), and HP trap (hydrophobic trap) (see Section 5.2 for the definition of these regions).

amounts to almost 1 Å with respect to *Ec*AcrB$_{WT}$. Minor differences were instead detected between *S.* Typhimurium and *E. coli* AcrB$_{WT}$ (Table 5.4).

These findings well correlates with the results of the watershell analysis, which was conducted on the MD trajectories of *S*TmAcrB$_{WT}$ and *S*TmAcrB$_{G288D}$ (see Section 5.2). It emerged, indeed, that the G288D substitution is associated to a considerable increment of the number of waters within the second watershell of the considered residue (see Table 5.5; see also Fig. 5.5). The increase in the hydration of the DBP reasonably causes the detected variations in the volume and radius of gyration of the pocket, and is likely responsible for the altered specificity of the transporter, described in previous literature [143]. Indeed, mutation of a glycine into a charged and bulkier residue is expected to have the largest impact on the structure, dynamics, hydration of the surrounding (prevalently hydrophobic) region. Moreover, our findings are consistent with previous studies [143], although here we have extended the analyses to conformers other than T and we have increased confidence by using multiple and independent structural models of AcrB.

In addition to the impact of the G288D substitution on the DBP, variations in the structure of the PBP were also evaluated. Due to the peculiarities of this pocket (which opens towards the periplasm, and therefore does not present well-defined boundaries), calculation of its volume through dedicated tools (see Section 5.2) is not straightforward. To overcome this difficulty, we focused on the occurrancy of contacts between subdomains PC1 and PC2, which enclose the pocket. The number of contacts between PC1 and PC2 should indeed reflect the opening or closing of the CH2 entry (or external cleft; see Chapter 1, Section 1.4), which is the main access route to the PBP from the periplasm. These considerations are consistent with the fact that, in *Ec*AcrB$_{WT}$, the O monomer presents the highest number of contacts, followed by the T and L monomer, respectively (see Chapter 1, Section 1.4 for a description of the positioning of the subdomains in the L,T and O monomers of *E. coli*).

From the analysis of our results (Table 5.6), it emerged that the number of contacts between PC1 and PC2 was significantly lower in *S*TmAcrB$_{G288D}$ than in *S*TmAcrB$_{WT}$, especially for the L conformer. This indicates that the G288D substitution also results in a greater opening of the PBP$_L$, which in turn facilitates the entrance of substrates through the CH2 access.

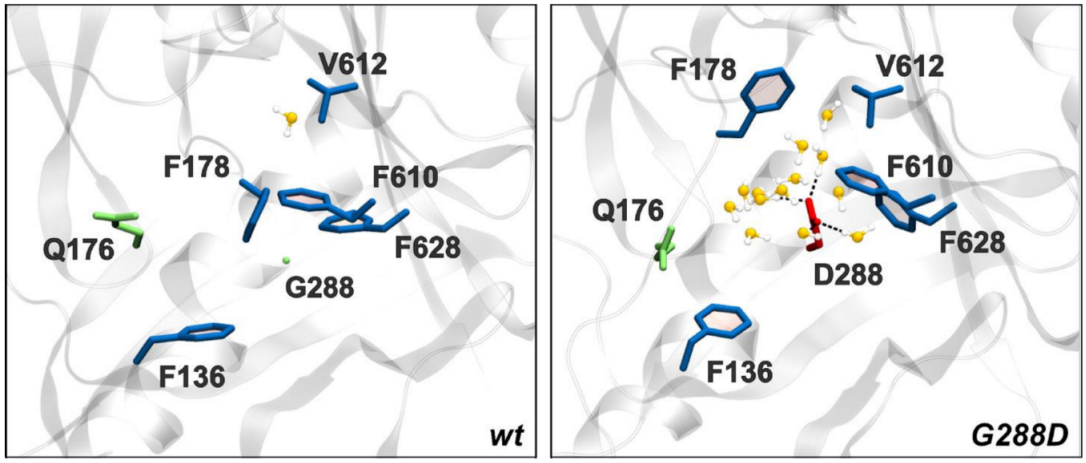

Figure 5.5: Close-up of the residue 288 and surrounding residues of the HP trap in STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$. Waters belonging to the first and second hydration shell of residue 288 (distance threshold: 5 Å, see Section 5.2) are also shown, and hydrogen bonds involving residue 288 are represented as dashed lines. This image has been created using two representative frames of MD trajectories.

System	AcrB conformer		
	L-conformer	T-conformer	O-conformer
# 1^{st} solvation shell waters			
STmAcrB$_{\mathrm{WT}}$	-	0.1 (0.3)	0.0 (0.1)
STmAcrB$_{\mathrm{G288D}}$	-	6.3 (0.6)	3.1 (0.3)
# 2^{nd} solvation shell waters			
STmAcrB$_{\mathrm{WT}}$	0.0 (0.1)	0.5 (0.4)	0.2 (0.2)
STmAcrB$_{\mathrm{G288D}}$	0.2 (0.3)	11.3 (1.2)	5.7 (0.5)

Table 5.5: Number of waters in the first and second solvation shell around residue 288 (in DBP), in STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$. The two solvation shells were defined by using a distance cut-off of 3.4 Å and 5.0 Å, respectively.

System	Number of contacts (PC1-PC2)		
	L-conformer	T-conformer	O-conformer
EcAcrB$_{\mathrm{WT}}$[1]	9 (5)	10 (2)	38 (2)
STmAcrB$_{\mathrm{WT}}$[2]	12 (4)	6 (2)	38 (2)
STmAcrB$_{\mathrm{G288D}}$[2]	1 (1)	7 (2)	31 (7)

[1]calculated on experimental reference structures; [2]calculated on representatives of each of the 100 clusters extracted from the MD trajectories.

Table 5.6: Number of contacts between the subdomains PC1 and PC2 in the three conformers of AcrB. Two residues have been considered in contact if the distance between their Cαs is below 10 Å (see Section 5.2).

5.4 Discussion

The resolution of the structure of STmAcrB$_{\mathrm{G288D}}$ through cryo-EM provided significant information on this protein, for which structural data from X-ray crystallography are currently missing due to difficulties in the crystallization process (see Section 5.1) [146]. In order to better understand its dynamical properties and the effect of the G288D substitution on the architecture of the binding pockets, we performed an *in silico* investigation based on homology modelling and MD simulations of both STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$. Specifically, the structural model of STmAcrB$_{\mathrm{G288D}}$ was refined in its cryo-EM map (see Section 5.2). Moreover, an ensamble of X-ray crystal structures of EcAcrB$_{\mathrm{WT}}$ was used as a reference to better characterize the structural properties of *S.* Typhimurium AcrB.

Interestingly, the analyses of our MD trajectories revealed that theDBP$_{\mathrm{T}}$ is markedly less voluminous in STmAcrB$_{\mathrm{WT}}$ than in the X-ray crystal structures of EcAcrB$_{\mathrm{WT}}$ (Table 5.3). In *S.* Typhimurium, therefore, a greater steric hindrance in the DBP might translate into altered substrate-processing kinetics and specificities with respect to *E. coli*. It should however be noted that a similar compression was also observed in MD simulations of EcAcrB$_{\mathrm{WT}}$in the absence of any ligands within the DBP, suggesting that unsolved ligands may have been present in the DBP$_{\mathrm{T}}$ of the experimental structures [135]. Despite that, MD analysis suggests that DBP of STmAcrB$_{\mathrm{WT}}$ exhibits markedly different dynamic properties to either the STmAcrB$_{\mathrm{G288D}}$ or EcAcrB$_{\mathrm{WT}}$, both of which show similar, larger volumes in the T-conformer (see Table 5.3; see also Table 5.4). This difference may translate into altered specificity and increased efflux of substrates that bind the DBP. Indeed, changes in the DBP characteristics have been reported to account for the discrepancies in the substrate specificity of other RND-transporters, such as AcrB and AcrD; MexB and MexY [118, 135, 152].

In STmAcrB$_{\mathrm{G288D}}$, this effect could be enhanced by the decrease in the number of contacts between subdomains PC1 and PC2 in the L-conformer (Table 5.6). This indicates, indeed, that the PBP$_{\mathrm{L}}$ is more open towards the periplasm (see Section 5.3), and is thus more accessible for potential substrates. On the basis of these considerations, it is reasonable that the structural changes detected in the DBP and PBP account for the differences in specificity between STmAcrB$_{\mathrm{WT}}$ and STmAcrB$_{\mathrm{G288D}}$, and specifically for the increased efflux of ciprofloxacin in the G288D variant. In this regard, however, it must be noticed that increased resistance to fluoroquinolones in STmAcrB$_{\mathrm{G288D}}$ is coupled to a higher sensitivity to minocycline and doxorubicin [143].

Overall, our data reinforce the relevance of residue G288 for the functioning of AcrB, which had been highlighted in previous literature [143]. Specifically, structural alterations due to the G288D substitution are not limited to the DBP, but also regard the packing of the PC1 and PC2 subdomains. Variations in the architecture of the DBP and the PBP may be communicated through different arrangements of the switch loop (see Chapter 1, Section 1.4) [154] and of the loop connecting PC2 to the funnel domain, which move closer in STmAcrB$_{\mathrm{G288D}}$ compared to STmAcrB$_{\mathrm{WT}}$. The considered substitution, therefore, seems to affect the plasticity of the drug binding pockets and, more generally, of the transport pathway.

Conclusions and future perspectives

In this thesis, several aspects of the inhibition of the AcrB efflux pump have been addressed. Specifically, the binding modes of several EPIs (the first-generation antipsychotics amitriptyline and chlorpromazine as well as the known inhibitor PAβN) have been investigated through computational methods. Our results have been integrated with experimental data as part of multisciplinary collaborations, in order to infer the inhibition mechanism of the considered compounds.

In the case of amitriptyline and chlorpromazine, experimental data from various techniques revealed that both compounds can act as EPIs against the AcrB efflux pump from *E. coli* and *S.* Typhimurium. They are indeed able to potentiate the activity of several AcrB substrates, including norfloxacin and ethidium bromide. Both amitriptyline and chlorpromazine, therefore, could be potential candidates for repurposing (*i.e.* the development of novel EPIs starting from in-use drugs with inhibitory activity). In order to understand their mode of action, we investigated the binding of both compounds to *E. coli* and *S.* Typhimurium AcrB by means of molecular docking and MD simulations. According to our results, amitriptyline and chlorpromazine preferentially bind to the DBP of the protein, forming stabilizing interactions with residues of the hydrophobic trap. Their binding modes present considerable similarities with those of norfloxacin and ethidium bromide, indicating that they might competitively bind to the DBP. Moreover, chlorpromazine well overlaps with the experimental binding pose of EPI MBX3132, whose inhibition mechanism is thought to involve competitive binding to the DBP and/or restrain of the functional dynamics of AcrB.

Computational techniques were also applied to investigate the binding of the EPI PAβN to the *E. coli* wt AcrB and its fluoroquinolone resistant variant bearing the G288D substitution. MD simulations were analysed in order to evaluate potential effects of the inhibitor binding on the flexibility and hydration of AcrB. Comparison of these results with experimental data from HDX-MS revealed that PAβN can considerably restrain the conformational dynamics of the wt AcrB and of the fluoroquinolone resistant variant. Rigidification of several regions of the binding pockets were indeed detected. Some of them are primarly involved in the substrate extrusion process, such as the switch loop or the external cleft. The inhibition mechanism of PAβN may thus involve a rigidification of key regions of the protein, some of which are located in the transport pathway. These results are in agreement with a previous model for the EPI mode of action, according to which the inhibitor may restrain the AcrB monomers in a specific conformation (possibly the T state) thus preventing substrate extrusion.

In order to evaluate the action of the inhibitor in presence of substrates, we applied our computational protocol to the ternary complex composed by AcrB, PAβN and the antibiotic ciprofloxacin. These analyses, combined with HDX-MS data, revealed that the co-presence of the antibiotic does not affect the action of the EPI. Indeed, a similar rigidification of AcrB was observed (to a less extent for the fluoroquinolone resistant variant). Moreover, our

MD simulations revealed that PAβN and ciprofloxacin can simultaneously occupy the DBP, forming direct stabilizing interactions. These data thus indicate that PAβN does not prevent or compete antibiotic binding.

Overall, our computational protocols have proven to be valuable for the evaluation of inhibition mechanism. They could thus be applied to other EPIs and repurposed drugs, in order to evaluate their interactions with AcrB and their molecular mode of action. In particular, significative results have been obtained upon the combination of our MD simulation with HDX-MS data. This approach could be applied to other EPIs and EPI/antibiotic couples, to evaluate the effect of the inhibitor on the pump conformational motions and its behaviour in presence of substrates. Moreover, our work may provide useful information for the combination of HDX-MS and MD simulation data. Indeed, integration of the two techniques is not straightforward, due to differences in their typical time scales.

In addition to the evaluation of the action of the mentioned EPIs, we applied computational methods to the structural characterization of the AcrB transporter from *S.* Typhimurium. This protein presents a high sequence identity (94.7%) with respect to its *E. coli* orthologue, whose structure has been resolved at high resolution. Despite this similarity, structural data for *S.* Typhimurium AcrB have been missing for long, due to difficulties in the crystallization process. Only recently, cryo-EM data have been obtained for its fluoroquinolone resistant G288D variant. In order to evaluate the structural differences between such variant and the wt AcrB, three dimensional structures of both proteins were realized through homology modeling. The obtained models were further relaxed through MD simulations of the truncated structure (porter and docking domains only). Structural characterization of the binding pockets revealed that the DBP is considerably compressed in the *S.* Typhimurium wt AcrB with respect to the values calculated for the G288D variant and the X-ray crystal structures of the *E. coli* orthologue. Additionally, the PBP is more open towards the periplasm with respect to the wt proteins from *S.* Typhimurium and *E. coli*. The G288D substitution, therefore, seems to have a considerable impact on the structure of the pump. These variations, in turn, may be related to the altered substrate specificity of the considered variant, which presents an increased fluoroquinolone resistance and a higher sensitivity to doxorubicin and minocycline. In our opinion, these results provide a useful characterization of the impact of the G288D substitution on the structure of AcrB. This may be of value for the design of novel EPIs and/or clinical drugs, as well as for further evaluation of substrate specificity in the considered pump.

Publications

The results of this thesis have been published as follows:

Grimsey EM, **Fais C**, Marshall RL, Ricci V, Ciusa ML, Stone JW, Ivens A, Malloci G, Ruggerone P, Vargiu AV, Piddock LJV. Chlorpromazine and amitriptyline are substrates and inhibitors of the AcrB multidrug efflux pump. *mBio.* 2020;11(3):e00465-20

Reading E, Ahdash Z, **Fais C**, Ricci V, Kan XW, Grimsey E, Stone J, Malloci G, Lau AM, Findlay H, Konjinenberg A, Booth PJ, Ruggerone P, Vargiu AV, Piddock LJV, Politis A. Perturbed structural dynamics underlie inhibition and altered efflux of the multidrug pump AcrB. *Nat Commun.* 2020; 11:5565

Johnson RM, **Fais C**, Parmar M, Cheruvara H, Marshall RM, Hesketh SJ, Feasy MC, Ruggerone P, Vargiu AV, Postis VGL, Muench SP, Bavro VN. Cryo-EM structure and molecular dynamics analysis of the fluoroquinolone resistant mutant of the AcrB transporter from *Salmonella*. *Microorganisms.* 2020;8(6):943

Bibliography

[1] Aminov RI. The role of antibiotics and antibiotic resistance iÀre. *Environ Microbiol* 2009; 11(12):2970-88 17

[2] Martinez JL. Environmental pollution by antibiotics and by antibiotic resistance determinants. *Environ Pollut* 2009; 157(11):2893-902 17

[3] Van Goethem MV, Pierneef R, Bezuidt OKI, Van De Peer Y, Cowan DA, Makhalayane TP. A reservoir of 'historical' antibiotic resistance genes in remote pristine Antarctic soils. *Microbiome* 2018; 6:40

[4] Bhullar K, Waglechner N, Pawlowski A, Koteva K, Banks ED, Johnston MD, Barton HA, Wright GD. Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS One* 2012; 7(4):e34953

[5] Holmes AH, Moore LP, Sundsfjord A, Steinbakk M, Regmi S, Karkey A, Guerin PJ, Piddock LJV. Understanding the mechanisms and drivers of antimicrobial resistance. *Lancet* 2016; 387(10014):176-87 17

[6] Waglechner N, McArthur AG, Wright GD. Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. *Nat Microbiol* 2019; 4, 1862-1871

[7] Gtari M, Ghodhbane-Gtari F, Nouioui I, Beauchemin N, Tisa LS. Phylogenetic perspectives of nitrogen-fixing actinobacteria. *Arch Microbiol* 2012; 194, 3–11

[8] Reygaert WC. An overview of the antimicrobial resistance mechanisms of bacteria. *AIMS Microbiol.* 2018; 4(3):482-501

[9] Gould K. Antibiotics: from prehistory to the present day. *J Antimicrob Chemother.* 2016; 71:572-575 17

[10] Ribeiro da Cunha B, Fonseca LP, Calado CRC. Antibiotic Discovery: Where Have We Come from, Where Do We Go?. *Antibiotics (Basel).* 2019;8(2):45 17

[11] Lewis K. New approaches to antimicrobial discovery. *Biochem Pharmacol.* 2017; 134:87-98 6, 17, 18

[12] Aslam B, Wang W, Arshad MI, et al. Antibiotic resistance: a rundown of a global crisis. *Infect Drug Resist.* 2018;11:1645–1658 17, 18

[13] Review on Antimicrobial Resistance. Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations. 2014. Available from: https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf 17

[14] Santajit S, Indrawattana N. Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens. *Biomed Res Int.* 2016; 2016:2475067

[15] Khan SN, Khan AU. Breaking the Spell: Combating Multidrug Resistant 'Superbugs'. *Front Microbiol.* 2016; 7:174 18

[16] Pogue JM, Kaye KS, Cohen DA, Marchaim D. Appropriate antimicrobial therapy in the era of multidrug-resistant humanpathogens. *Clin Microbiol Infect.* 2015; 21(4):302-12 17, 18

[17] European Centre for Disease Prevention and Control. Rapid Risk Assessment Carbapenem-resistant Enterobacteriaceae – second update. 2019. Available from: https://www.ecdc.europa.eu/sites/default/files/documents/carbapenem-resistant-enterobacteriaceae-risk-assessment-rev-2.pdf

[18] Pisani E. Review on Antimicrobial Resistance. Antimicrobial resistance: What does medicine quality have to do with it? 2015. Available from: https://amr-review.org/sites/default/files/ElizabethPisaniMedicinesQualitypaper.pdf 18

[19] Smith PA, Koehler MFT, Girgis HS, *et al.* Optimized arylomycins are a new class of gram-negative antibiotics. *Nature.* 2018; 561(7722):189–194 18

[20] Murray PR, Rosenthal KS, Pfaller MA. (2016). Medical microbiology. Elsevier Inc. 6, 19, 20

[21] Reygaert WC. An overview of the antimicrobial resistance mechanisms of bacteria. *AIMS Microbiol.* 2018; 4(3):482–501 19, 21, 49

[22] Exner M, Bhattacharya S, Christiansen B, et al. Antibiotic resistance: What is so special about multidrug-resistant gram-negative bacteria?. *GMS Hyg Infect Control.* 2017;12:Doc05 19

[23] Munita JM, Arias CA. Mechanisms of Antibiotic Resistance. *Microbiol Spectr.* 2016;4(2):10. 19, 21

[24] Nikaido H. Multiple antibiotic resistance and efflux. *Curr Opin Microbiol.* 1998;1(5):516–523 21, 49

[25] Garcia-Leon G, Salgado F, Oliveros JC, Sanchez MB, Martinez JL. Interplay between intrinsic and acquired resistance to quinolones in Stenotrophomonas maltophilia. *Environ Microbiol.* 2014; 16(5):1282-1296 21

[26] Nolivos S, Cayron J, Dedieu A, Page A, Delolme F, Lesterlin C. Role of AcrAB-TolC multidrug efflux pump in drug-resistance acquisition by plasmid transfer. *Science.* 2019; 364, 778-782 21

[27] Du D, van Veen HW, Murakami S, Pos KM, Luisi BF. Structure, mechanism and cooperation of bacterial multidrug transporters. *Curr Opin Struct Biol.* 2015;33:76-91 6, 22, 23, 25

[28] Li XZ, Nikaido H. Efflux-mediated drug resistance in bacteria: an update. *Drugs.* 2009;69(12):1555-1623 18, 21, 22

[29] Kumar S, Mukherjee MM, Varela MF. Modulation of Bacterial Multidrug Resistance Efflux Pumps of the Major Facilitator Superfamily. *Int J Bacteriol.* 2013;2013:204141 18

[30] Spengler G, Kincses A, Gajdács M, Amaral L. New Roads Leading to Old Destinations: Efflux Pumps as Targets to Reverse Multidrug Resistance in Bacteria. *Molecules.* 2017;22(3):468

[31] Du D, Wang-Kan X, Neuberger A, et al. Multidrug efflux pumps: structure, function and regulation. *Nat Rev Microbiol.* 2018;16(9):523-539 6, 21, 22, 23, 24, 25, 26, 27, 35, 86

[32] Du D, Van Veen HW, Luisi BF. Assembly and operation of bacterial tripartite multidrug efflux pumps. *Trends Microbiol.* 2015; 23(5):311-9 21, 23, 25, 26, 27

[33] Nikaido H. RND transporters in the living world. *Res Microbiol.* 2018; 169(7-8):363-371 23, 25

[34] Collu F, Cascella M. Multidrug resistance and efflux pumps: insights from molecular dynamics simulations. *Curr Top Med Chem.* 2013;13(24):3165-3183 21, 35

[35] Li XZ. (2017) Active Efflux as a Mechanism of Resistance to Antimicrobial Drugs. In: Mayers D., Sobel J., Ouellette M., Kaye K., Marchaim D. (eds) Antimicrobial Drug Resistance. Springer, Cham 18, 23, 25

[36] Du D, Wang Z, James NR, et al. Structure of the AcrAB-TolC multidrug efflux pump. *Nature.* 2014;509(7501):512-515 23, 25, 27

[37] Wang Z, Fan G, Hryc CF, et al. An allosteric transport mechanism for the AcrAB-TolC multidrug efflux pump. *Elife.* 2017;6:e24905 11, 18, 25, 27, 32, 33, 51, 68, 84

[38] Fitzpatrick AWP, Llabrés S, Neuberger A, et al. Structure of the MacAB-TolC ABC-type tripartite multidrug efflux pump. *Nat Microbiol.* 2017;2:17070 25, 26

[39] Shi X, Chen M, Yu Z, et al. In situ structure and assembly of the multidrug efflux pump AcrAB-TolC. *Nat Commun.* 2019;10(1):2635 6, 25, 26, 27, 33

[40] Ramaswamy VK, Cacciotto P, Malloci G, Ruggerone P, Vargiu AV. Multidrug Efflux Pumps and Their Inhibitors Characterized by Computational Modeling, in Efflux-Mediated Antimicrobial Resistance in Bacteria. 2016, Springer. p. 797-831 26, 27, 38

[41] Edgar R, Bibi E. MdfA, an Escherichia coli multidrug resistance protein with an extraordinarily broad spectrum of drug recognition. *J Bacteriol.* 1997;179(7):2274-228 27

[42] Ruggerone P, Murakami S, Pos KM, Vargiu AV. RND efflux pumps: structural information translated into function and inhibition mechanisms. *Curr Top Med Chem.* 2013;13(24):3079-3100 17, 27, 28, 30, 31

[43] Westfall DA, Krishnamoorthy G, Wolloscheck D, Sarkar R, Zgurskaya HI, Rybenkov VV. Bifurcation kinetics of drug uptake by gram-negative bacteria. *PLoS One.* 2017;12(9):e0184671 26

[44] Murakami S, Nakashima R, Yamashita E, Yamaguchi A. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature.* 2002;419(6907):587-593 6, 7, 27, 28, 29, 64

[45] Eicher T, Cha HJ, Seeger MA, 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 U S A.* 2012;109(15):5687-5692 7, 12, 27, 28, 30, 31, 36, 51, 64, 90, 91, 93

[46] Eicher T, Seeger MA, Anselmi C, et al. Coupling of remote alternating-access transport mechanisms for protons and substrates in the multidrug efflux pump AcrB. *Elife.* 2014;3:e03145 27, 30, 64

[47] Sjuts H, Vargiu AV, Kwasny SM, et al. Molecular basis for inhibition of AcrB multidrug efflux pump by novel and powerful pyranopyridine derivatives. *Proc Natl Acad Sci U S A.* 2016;113(13):3509-3514 11, 27, 32, 33, 52, 53, 54, 56, 58, 59, 60, 62, 65, 90, 92

[48] Du D, Neuberger A, Orr MW, et al. Interactions of a Bacterial RND Transporter with a Transmembrane Small Protein in a Lipid Environment. *Structure.* 2020;28(6):625-634.e6 27

[49] Liu M, Zhang XC. Energy-coupling mechanism of the multidrug resistance transporter AcrB: Evidence for membrane potential-driving hypothesis through mutagenic analysis. *Protein Cell.* 2017;8(8):623-627 27

[50] Fang J, Yu L, Wu M, Wei Y. Dissecting the function of a protruding loop in AcrB trimerization. *Journal of Biomolecular Structure and Dynamics.* 2013;31:4, 385-392 27

[51] Zwama M, Yamasaki S, Nakashima R, Sakurai K, Nishino K, Yamaguchi A. Multiple entry pathways within the efflux transporter AcrB contribute to multidrug recognition. *Nat Commun.* 2018;9(1):124 7, 28, 29, 30, 54, 61

[52] Tam HK, Malviya VN, Foong WE, Herrmann A, Malloci G, Ruggerone P, Vargiu AV, Pos KM. Binding and Transport of Carboxylated Drugs by the Multidrug Transporter AcrB. *J Mol Biol.* 2020;432:861-877 7, 28, 29, 30

[53] Qiu W, Fu Z, Xu GG, et al. Structure and activity of lipid bilayer within a membrane-protein transporter. *Proc Natl Acad Sci U S A.* 2018;115(51):12985-12990 28, 91

[54] Murakami S, Nakashima R, Yamashita E, Matsumoto T, Yamaguchi A. Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature.* 2006;443(7108):173-179 30

[55] Matsunaga Y, Yamane T, Terada T, et al. Energetics and conformational pathways of functional rotation in the multidrug transporter AcrB. *Elife.* 2018;7:e31715 30

[56] Cha HJ, Müller RT, Pos KM. switch loop flexibility affects transport of large drugs by the promiscuous AcrB multidrug efflux transporter. *Antimicrob Agents Chemother.* 2014;58(8):4767-4772 31, 33

[57] Opperman TJ, Nguyen ST. Recent advances toward a molecular mechanism of efflux pump inhibition. *Front Microbiol.* 2015;6:421 31, 32, 33, 62

[58] Lomovskaya O, Bostian KA. Practical applications and feasibility of efflux pump inhibitors in the clinic—A vision for applied use. *Biochem Pharmacol.* 2006;71:910-918 32, 49, 62

[59] Yu EW, Aires JR, McDermott G, Nikaido H. A periplasmic drug-binding site of the AcrB multidrug efflux pump: a crystallographic and site-directed mutagenesis study. *J Bacteriol.* 2005;187(19):6804-6815 32

[60] Vargiu AV, Nikaido H. Multidrug binding properties of the AcrB efflux pump characterized by molecular dynamics simulations. *Proc Natl Acad Sci U S A.* 2012;109(50):20637-20642 32, 33, 52, 53, 54, 64, 65, 90, 92

[61] Renau TE, Leger R, Flamme EM, Sangalang J, She MW, Yen R, et al. Inhibitors of efflux pumps in Pseudomonas aeruginosa potentiate the activity of the fluoroquinolone antibacterial levofloxacin. *J. Med. Chem.* 1999;42, 4928–4931 32

[62] Bohnert JA, Kern WV. Selected arylpiperazines are capable of reversing multidrug resistance in Escherichia coli overexpressing RND efflux pumps. *Antimicrob Agents Chemother.* 2005;49(2):849-852 32

[63] Schuster S, Kohler S, Buck A, et al. Random mutagenesis of the multidrug transporter AcrB from Escherichia coli for identification of putative target residues of efflux pump inhibitors. *Antimicrob Agents Chemother.* 2014;58(11):6870-6878 32

[64] Opperman TJ, Kwasny SM, Kim HS, et al. Characterization of a novel pyranopyridine inhibitor of the AcrAB efflux pump of Escherichia coli. *Antimicrob Agents Chemother.* 2014;58(2):722-733 33, 49

[65] Dobson CM. Biophysical Techniques in Structural Biology. *Annu Rev Biochem.* 2019;88:25-33 17, 35, 36

[66] Kandt C, Monticelli L. Membrane protein dynamics from femtoseconds to seconds. *Methods Mol Biol.* 2010. 654:423-440 17, 35

[67] Hvidsten TR, Laegreid A, Kryshtafovych A, Andersson G, Fidelis K, Komorowski J. A comprehensive analysis of the structure-function relationship in proteins based on local structure similarity. *PLoS One.* 2009;4(7):e6266 17, 36

[68] Muhammed MT, Aki-Yalcin E. Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chem Biol Drug Des.* 2019;93(1):12-20 36

[69] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Research.* 2000;28: 235-242 17, 36

[70] Deng H, Jia Y, Zhang Y. Protein structure prediction. *Int J Mod Phys B.* 2018;32(18):1840009 17, 36

[71] Krissinel E. On the relationship between sequence and structure similarities in proteomics. *Bionformatics.* 2007;23(6):717-723 36

[72] Rost B. Twilight zone of protein sequence alignments. *Protein Eng.* 1999;12(2):85-94 36

[73] Cacciotto P, Ramaswamy VK, Malloci G, Ruggerone P, Vargiu AV. Molecular Modeling of Multidrug Properties of Resistance Nodulation Division (RND) Transporters. *Methods Mol Biol.* 2018;1700:179-219 36

[74] UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 2019;47(D1):D506-D515 36, 51, 64, 90

[75] Webb B, Sali A. Protein Structure Modeling with MODELLER. *Methods Mol. Biol.* 2017; 1654, 39–54 36

[76] Webb B, Sali A. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci.* 2016;86:2.9.1-2.9.37 7, 37, 64, 90

[77] MacKerell AD, Bashford D, Bellott M, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B.* 1998;102(18):3586-3616 36, 39

[78] Topf M, Lasker K, Webb B, Wolfson H, Chiu W, Sali, A. Protein structure fitting and refinement guided by cryo-EM density. *Structure.* 2008, 16, 295–307 36

[79] Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. *Neuron.* 2018;99(6):1129-1143 38

[80] Hernández-Rodríguez M, Rosales-Hernández MC, Mendieta-Wejebe JE, Martínez-Archundia M, Basurto JC. Current tools and methods in Molecular Dynamics (MD) simulations for drug design. *Curr Med Chem.* 2016;23(34):3909-3924 38

[81] Atilgan C. Computational Methods for Efficient Sampling of Protein Landscapes and Disclosing Allosteric Regions. *Adv Protein Chem Struct Biol.* 2018;113:33-631 38

[82] Cramer JC. Essentials of computational chemistry: theories and models. 2nd ed. 2004. John Wiley and Sons Ltd 38, 41

[83] Dauter Z, Lamzin VS, Wilson KS. Proteins at atomic resolution. *Curr Opin Struct Biol.* 1995;5(6):784-790 38

[84] Martín-García F, Papaleo E, Gomez-Puertas P, Boomsma W, Lindorff-Larsen K. Comparing molecular dynamics force fields in the essential subspace. *PLoS One.* 2015;10(3):e0121114 38, 39

[85] Frenkel D, Smit B. Understanding molecular simulations: from algorithms to applications. 2nd ed. 2002. Academic press 38, 39, 41

[86] Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. *J Comput Chem.* 2005;26:1668–1688 39, 43, 46, 52, 53, 65, 90

[87] Senn HM, Thiel W. QM/MM methods for biomolecular systems. *Angew Chem Int Ed Engl.* 2009;48(7):1198-1229 7, 40

[88] Griffiths PM. Introduction to vibrational spectroscopy. In Handbook of Vibrational Spectroscopy. 2006. John Wiley & Sons Ltd 7, 40

[89] Katiyar R, Prateek J. (2018). Molecular simulations in drug delivery: Opportunities and challenges. *Wiley Interdisciplinary Reviews: Computational Molecular Science.* 2018;e1358. 10.1002/wcms.1358 7, 42

[90] Malloci G, Vargiu AV, Serra G, Bosin A, Ruggerone P, Ceccarelli M. A database of force-field parameters, dynamics, and properties of antimicrobial compounds. *Molecules.* 2015;20:13997–14021 43, 52, 64, 65

[91] Guedes IA, de Magalhães CS, Dardenne LE. Receptor-ligand molecular docking. *Biophys Rev.* 2014;6(1):75-87 44

[92] Salmaso V, Moro S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front Pharmacol.* 2018;9:923 44

[93] Chen YC. Beware of docking! *Trends Pharmacol Sci.* 2015;36(2):78-951 44

[94] Trott O, Olson AJ. Software news and update AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2009;31:455–461 44, 51, 64

[95] Durrant JD, De Oliveira CAF, McCammon JA. POVME: An algorithm for measuring binding-pocket volumes. *J. Mol. Graph. Model.* 2011; 29, 773–776 46

[96] Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov.* 2015;10(5):449-461 47, 53, 56

[97] Sjuts H, Vargiu AV, Kwasny SM, et al. Molecular basis for inhibition of AcrB multidrug efflux pump by novel and powerful pyranopyridine derivatives. *Proc Natl Acad Sci U S A.* 2016;113(13):3509-3514 48

[98] Hopkins CW, Le Grand S, Walker RC, Roitberg AE. Long-time-step molecular dynamics through hydrogen mass repartitioning. J Chem Theory Comput. 2015;11:1864–1874 44, 66, 91

[99] Bailey AM, Paulsen IT, Piddock LJ. RamA confers multidrug resistance in Salmonella enterica via increased expression of acrB, which is inhibited by chlorpromazine. *Antimicrob Agents Chemother.* 2008;52:3604–3611 49

[100] Grimsey EM, Piddock LV. Do phenothiazines possess antimicrobial and efflux inhibitory properties? *FEMS Microbiol Rev.* 2019;43:577–590 49

[101] Nehme H, Saulnier P, Ramadan AA, Cassisa V, Guillet C, Eveillard M, Umerska A. Antibacterial activity of antipsychotic agents, their association with lipid nanocapsules and its impact on the properties of the nanocarriers and on antibacterial activity. *PLoS One.* 2018;13:e0189950 49

[102] Ying YC, Yong MO, Kim LC. Synergistic interaction between phenothiazines and antimicrobial agents against *Burkholderia pseudomallei*. *Antimicrob Agents Chemother.* 2007;51:623–630 49

[103] Kristiansen JE, Hendricks O, Delvin T, Butterworth TS, Aagaard L, Christensen JB, Flores VC, Keyzer H. Reversal of resistance in microorganisms by help of non-antibiotics. *J Antimicrob Chemother.* 2007;59:1271–1279 49

[104] Coutinho HDM, Costa JGM, Lima EO, Falcão-Silva VS, Siqueira JP. Herbal therapy associated with antibiotic therapy: potentiation of the antibiotic activity against methicillin-resistant *Staphylococcus aureus* by *Turnera ulmifolia L. BMC Complement Altern Med.* 2009;9:13 49

[105] Kaatz GW, Moudgal VV, Seo SM, Kristiansen JE. Phenothiazines and thioxanthenes inhibit multidrug efflux pump activity in *Staphylococcus aureus*. *Antimicrob Agents Chemother.* 2003;47:719–726 49

[106] Martins A, MacHado L, Costa S, Cerca P, Spengler G, Viveiros M, Amaral L. Role of calcium in the efflux system of *Escherichia coli*. *Int J Antimicrob Agents.* 2011;37:410–414 49

[107] Lawler AJ, Ricci V, Busby SJW, Piddock LV. Genetic inactivation of acrAB or inhibition of efflux induces expression of ramA. *J Antimicrob Chemother.* 2013;68:1551–1557 49

[108] Grimsey EM, Fais C, Marshall RL, Ricci V, Ciusa ML, Stone JW, Ivens A, Malloci G, Ruggerone P, Vargiu AV, Piddock LJV. Chlorpromazine and amitriptyline are substrates and inhibitors of the AcrB multidrug efflux pump. *mBio.* 2020;11(3):e00465-20 7, 49, 50, 60, 90

[109] Cossu F, Cascella M. Multidrug Resistance and Efflux Pumps: Insights from Molecular Dynamics Simulations. *Current Topics in Medicinal Chemistry.* 2013;13, 000-000 50

[110] Šali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 1993;234:779–815 18, 51

[111] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 2011;7:53 51

[112] Sennhauser G, Amstutz P, Briand C, Storchenegger O, Grütter MG. Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors. *PLoS Biol.* 2007;5(1):e7 36, 51, 64, 90, 91

[113] Vargiu AV, Ruggerone P, Opperman TJ, Nguyen ST, Nikaido H. Molecular mechanism of MBX2319 inhibition of Escherichia coli AcrB multidrug efflux pump and comparison

with other inhibitors. *Antimicrob Agents Chemother.* 2014;58:6224–6234 7, 32, 52, 53, 56, 69

[114] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput.* 2015;11: 3696–3713 43, 52, 90

[115] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983;79:926–935 43, 52, 90

[116] Joung IS, Cheatham TE. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B.* 2008;112:9020–9041 52

[117] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general Amber force field. *J Comput Chem.* 2004;25: 1157–1174 43, 52

[118] Ramaswamy VK, Vargiu AV, Malloci G, Dreier J, Ruggerone P. Molecular determinants of the promiscuity of MexB and MexY multidrug transporters of *Pseudomonas aeruginosa*. *Front Microbiol.* 2018;9:1144 31, 52, 96

[119] Hopkins CW, Le Grand S, Walker RC, Roitberg AE. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J Chem Theory Comput.* 2015;11:1864–1874 52

[120] Williams T, Kelley C, Campbell J, Cunningham R, Elber G, Fearick R, Kotz D, Kubaitis E, Lang R, Lehmann A, Steger C, Tkacik T, Van Der Woude J, Woo A. 1986. GNUPLOT: an interactive plotting program. http://www-h.eng.cam.ac.uk/help/documentation/docsource/gnuplot.pdf 52, 66

[121] Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. *J Mol Graph.* 1996;14:33–38 52, 66

[122] Kinana AD, Vargiu AV, May T, Nikaido H. Aminoacyl-beta-naphthylamides as substrates and modulators of AcrB multidrug efflux pump. *Proc Natl Acad Sci U S A.* 2016;113:1405–1410 53, 69

[123] Kinana AD, Vargiu AV, Nikaido H. Some ligands enhance the efflux of other ligands by the *Escherichia coli* multidrug pump AcrB. *Biochemistry.* 2013;52:8342–8351 53

[124] Gohlke H, Kiel C, Case DA. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J Mol Biol.* 2003;330:891–913 53

[125] Matter H, Nazaré M, Güssregen S, Will DW, Schreuder H, Bauer A, Urmann M, Ritter K, Wagner M, Wehner V. Evidence for C-Cl/C-Br$\cdots\pi$ interactions as an important contribution to protein-ligand binding affinity. *Angew Chem Int Ed Engl.* 2009;48:2911–2916 60

[126] Marcsisin SR, Engen JR. Hydrogen exchange mass spectrometry: what is it and what can it tell us?. *Anal Bioanal Chem.* 2010;397(3):967-972 62

[127] Englander SW, Mayne L, Bai Y, Sosnick TR. Hydrogen exchange: the modern legacy of Linderstrøm-Lang. *Protein Sci.* 1997;6(5):1101-1109 62

[128] Reading E, Ahdash Z, Fais C, Ricci V, Kan XW, Grimsey E, Stone J, Malloci G, Lau AM, Findlay H, Konjinenberg A, Booth PJ, Ruggerone P, Vargiu AV, Piddock LJV, Politis A. Perturbed structural dynamics underlie inhibition and altered efflux of the multidrug pump AcrB. *Nat Commun.* (in press) 8, 62, 63, 82

[129] Atzori A, Malviya VN, Malloci G, et al. Identification and characterization of carbapenem binding sites within the RND-transporter AcrB. *Biochim Biophys Acta Biomembr.* 2019;1861(1):62-74 64

[130] Malvacio I, Buonfiglio R, D'Atanasio N, Serra G, Bosin A, Di Giorgio FP, Ruggerone P, Ombrato R, Vargiu AV. Molecular basis for the different interactions of congeneric substrates with the polyspecific transporter AcrB. *Biochim Biophys Acta Biomembr.* 2019 Jul 1;1861(7):1397-1408 64

[131] Amaro RE, Baudry J, Chodera J, et al. Ensemble Docking in Drug Discovery. *Biophys J.* 2018;114(10):2271-2278 64

[132] Seeger MA, Schiefner A, Eicher T, Verrey F, Diederichs K, Pos KM. Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science.* 2006;313(5791):1295-1298 64

[133] Nakashima R, Sakurai K, Yamasaki S, et al. Structural basis for the inhibition of bacterial multidrug exporters. *Nature.* 2013;500(7460):102-106 64

[134] Nakashima R, Sakurai K, Yamasaki S, Nishino K, Yamaguchi A. Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket. *Nature.* 2011;480(7378):565-569 64

[135] Ramaswamy VK, Vargiu AV, Malloci G, Dreier J, Ruggerone P. Molecular Rationale behind the Differential Substrate Specificity of Bacterial RND Multi-Drug Transporters. *Sci. Rep.* 2017; 7, 8075 31, 65, 90, 92, 96

[136] Gonzalez-Salgado D, Vega C. A new intermolecular potential for simulations of methanol: The OPLS/2016 model. *J Chem Phys.* 2016;145(3):034508 65

[137] Wang LP, McKiernan KA, Gomes J, Beauchamp KA, Head-Gordon T, Rice JE, Swope WC, Martínez TJ, Pande VS. Building a More Predictive Protein Force Field: A Systematic and Reproducible Route to AMBER-FB15. *J Phys Chem B.* 2017;121(16):4023-4039 43, 65

[138] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem.* 2004 Jul 15;25(9):1157-74 65

[139] Lee MB, Greig JD. A review of nosocomial Salmonella outbreaks: infection control interventions found effective. *Public Health.* 2013;127:199–206 86

[140] Curiao T, Marchi E, Grandgirard D, León-Sampedro R, Viti C, Leib SL, Baquero F, Oggioni MR, Martinez JL, Coque TM. Multiple adaptive routes of Salmonella enterica Typhimurium to biocide and antibiotic exposure. *BMC Genomics.* 2016;17, 491 86

[141] Misra R, Bavro VN. Assembly and transport mechanism of tripartite drug efflux systems. Biochim. *Biophys. Acta.* 2009; 1794, 817–825 86

[142] Kobylka J, Kuth MS, Müller RT, Geertsma ER, Pos KM. AcrB: A mean, keen, drug efflux machine. *Ann. N. Y. Acad. Sci.* 2019;1459, 38–68 86

[143] Blair JMA, Bavro VN, Ricci V, Modi N, Cacciotto P, Kleinekathöfer U, Ruggerone P, Vargiu AV, Baylay AJ, Smith HE, et al. AcrB drug-binding pocket substitution confers clinically relevant resistance and altered substrate specificity. *Proc. Natl. Acad. Sci. USA.* 2015; 112, 3511–3516 7, 31, 86, 87, 90, 92, 94, 96

[144] Piddock LJ, Griggs DJ, Hall MC, Jin YF. Ciprofloxacin resistance in clinical isolates of *Salmonella* Typhimurium obtained from two patients. *Antimicrob. Agents Chemother.* 1993; 37, 662–666 31, 86

[145] Piddock LJ, Whale K, Wise R. Quinolone resistance in Salmonella: Clinical experience. *Lancet.* 1990; 335, 1459 31, 86

[146] Johnson RM, Fais C, Parmar M, et al. Cryo-EM structure and molecular dynamics analysis of the fluoroquinolone resistant mutant of the AcrB transporter from Salmonella. *Microorganisms.* 2020;8(6):943 10, 86, 87, 88, 89, 90, 92, 96

[147] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 2011;7:539 90

[148] Topf M, Lasker K, Webb B, Wolfson H, Chiu W, Sali, A. Protein structure fitting and refinement guided by cryo-EM density. *Structure.* 2008, 16, 295–307 12, 90, 91

[149] Joung IS, Cheatham TE. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B.* 2008;112:9020–9041 90, 92

[150] Das D, Xu Q.S, Lee JY, Ankoudinova I, Huang C, Lou Y, De Giovanni A, Kim R, Kim SH. Crystal structure of the multidrug efflux transporter AcrB at 3.1Å resolution reveals the N-terminal region with conserved amino acids. *J. Struct. Biol.* 2007; 158, 494–502 91

[151] Durrant JD, De Oliveira CAF, McCammon JA. POVME: An algorithm for measuring binding-pocket volumes. *J. Mol. Graph. Model.* 2011; 29, 773–776 91

[152] Nakashima R, Sakurai K, Yamasaki S, Hayashi K, Nagata C, Hoshino K, Onodera Y, Nishino K, Yamaguchi A. Structural basis for the inhibition of bacterial multidrug exporters. *Nature.* 2013; 500, 102–106 96

[154] Husain F, Nikaido, H. Substrate path in the AcrB multidrug efflux pump of *Escherichia coli. Mol. Microbiol.* 2010; 78, 320–330 96

[155] Zuo Z, Weng J, Wang W. Insights into the inhibitory mechanism of D13-9001 to the multidrug transporter AcrB through molecular dynamics simulations. *J Phys Chem B.* 2016;120(9):2145-54 17, 84

[156] Jamshidi S, Sutton JM, Rahman KM. Mapping the Dynamic Functions and Structural Features of AcrB Efflux Pump Transporter Using Accelerated Molecular Dynamics Simulations. *Sci Rep.* 2018;8(1):10470 84

[157] Jamshidi, S., Sutton, J. M., & Rahman, K. M. Computational Study Reveals the Molecular Mechanism ofthe Interaction between the Efflux Inhibitor PAN and the AdeB Transporter from Acinetobacter baumannii. *ACSOmega.* 2017;2(6), 3002-3016

84

Acknowledgements

Many persons have guided and supported me through this journey. I would like to seize the opportunity to acknowledge them.

Firsty, I wish to thank my supervisors, Paolo and Attilio. Their mentoring, with all their encouragement and insightful questions, has truly incented me to deepen my studies about our research topic. Many thanks also to Giuliano, for his useful advice and the nice discussions we had during these years, and to Andrea and Giovanni, whose technical support has been fundamental during these years.

My sincere gratitude also goes to Dr. Arianna Fornili, for giving me the opportunity to join her group as visiting student and for her dedication and valuable advice. Thanks also to Dr. Alessandro Pandini for his great helpfulness.

Many thanks also to the past and present members of Paolo's group, especially Andrea, for their advice and all the interesting discussions.

A special thanks goes to Michela, Roberta and Stefania, with whom I shared my doubts and difficulties as well as a lot of fun times. Thanks also to Ljiljana for her support and to Melinda, Alessia and Sara for always being by my side, despite the distance between us.

Last but not least, a heartfelt thanks to my family. This achievement would not have been possible without their unconditional support.