Università degli Studi di Cagliari

DOTTORATO DI RICERCA

IN FISICA

Ciclo XXIX

Molecular Rationale Behind the Differential Substrate Specificity of Homologous RND Transporters in *E. coli* And *P. aeruginosa*

FIS/03

Presentata da: Venkata Krishnan Ramaswamy

Coordinatore Dottorato   Prof. Alessandro De Falco

Tutor   Prof. Paolo Ruggerone

Esame finale anno accademico 2015 – 2016
Tesi discussa nella sessione d'esame marzo – aprile 2017

Acknowledgements

R. Venkata Krishnan

Table of Contents

Chapter 1: Introduction

Chapter 2: Computational Methods

Chapter 3: Molecular Rationale Behind the Differential Substrate Specificity of AcrB and AcrD Transporters of *E. coli*

Chapter 4: Molecular Rationale Behind the Differential Substrate Specificity of MexB and MexY Transporters of *P. aeruginosa*

Chapter 5: Substrate-Transporter Interaction Analysis in AcrB and AcrD

Chapter 6: Choice of A Better Ligand-Parameterization Approach for HADDOCK

Abbreviations

ABC	ATP-Binding Cassette
AcrB	Acriflavin resistance protein B
AcrD	Acriflavin resistance protein D
AMR	Anti-Microbial resistance
AP	Access Pocket
AZT	Aztreonam
DMT	Drug/Metabolite Transporter
DOPE	Discretised Optimised Potential Energy
DP	Deep Pocket
FD	Funnel Domain
GAFF	Generalised Amber Force Field
HAE	Hydrophobic/Amphiphilic Efflux
HME	Heavy Metal Efflux
IM	Inner Membrane
L	Loose (Access)
LI	Lipophilic Index
MATE	Multi-drug And Toxic-compound Extrusion
MD	Molecular Dynamics
MDR	Multidrug Resistance
MFP	Membrane Fusion Protein
MFS	Major Facilitator Superfamily
MIN	Minocycline
MLP	Molecular Lipophilicity Potential
MM	Molecular Mechanics
MOP	Multidrug/Oligosaccharidyl-lipid/Polysaccharide
NBD	Nucleotide-Binding Domains,
OM	Outer Membrane
OMP	Outer Membrane Protein
PBC	Periodic Boundary Condition
PD	Pore (Porter) Domain
P-gp	P-glycoprotein
PME	Particle Mesh Ewald
POPE	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
O	Open (Extrusion)
QM	Quantum Mechanics
RDF	Radial Distribution Function
RMSD	Root Mean Square Deviation
RND	Resistance Nodulation and cell-Division
SMR	Small Multidrug Resistance
T	Tight (binding)
TDR	Total Drug Resistance
TM	Transmembrane
TMDs	Transmembrane Domains
TMS	Transmembrane Segment
vdW	van der Waals
XDR	Extensive Drug Resistance

Abstract

The discovery of medicinal antibiotics was a crucial breakthrough in the treatment of infectious diseases. Unfortunately, bacteria proved to be highly competitive by expeditiously developing various survival mechanisms to deal with most (if not all) antibiotics available today. One such highly efficient mechanism is the over expression of specific and general transporters that recognize a wide spectrum of substrates (including many chemically different antibiotics) and actively expel them out of the cells, thereby contributing to multidrug resistance (MDR). Resistance-Nodulation-Division (RND) transporters like AcrB and AcrD in *Escherichia coli*, and MexB and MexY in *Pseudomonas aeruginosa* are the most prominent multi-component drug efflux pumps exporting a wide range of substrates ranging from lipophilic to amphiphilic molecules. Despite a comparable overall sequence homology among these RND transporters of *E. coli* and *P. aeruginosa*, they exhibit varied substrate specificity, the underlying basis of which still remains elusive. In an attempt to provide better insights into the substrate-transporter complementarity underlying the recognition and transport events in Acr pumps of *E. coli* and Mex pumps of *P. aeruginosa*, we performed a comparative analysis of multi-copy μs-long MD simulations of the apo-forms of AcrB, AcrD, MexB and MexY transporters. To this effect, we chose a set of important descriptors like pocket volume, molecular lipophilic potential, electrostatic potential and hydration to characterize the putative binding pockets (Access and Deep Pockets) in these transporters. Owing to the absence of crystal structure for AcrD and MexY, we also modeled their 3D-structures based on the high-resolution crystal structure of their closest homologues for this study.

Our results suggest that the interactions of ligands with and their affinity to these transporters arise from an interplay between physicochemical properties, such as volume, lipophilicity, electrostatic potential, and certain specific features like changes in the loop conformations, altogether tuned by the dynamics of the systems. The thesis discusses in detail the important findings from our μs-long MD simulations of AcrB, AcrD, MexB and MexY proteins in the absence of a bound substrate, emphasizing the molecular determinants governing the partially different substrate specificity of the two couples of proteins in *E. coli* and *P. aeruginosa*. In addition, certain key interaction types needed for a substrate to bind to its transporter and/or for a transporter to recognize its substrate are also discussed for Acr transporters in *E. coli*.

Chapter 1

Introduction

1.1. Introduction

Antimicrobials (antibiotics and similar drugs) are probably one of the most successful forms of chemotherapy in the history of medicine. The use of antimicrobials dates back to 350-550 CE [1, 2] based on traces of tetracycline found in the human skeletal remains from ancient Sudanese Nubia and also from Dakhleh Oasis, Egypt [3]. It was postulated that intake of tetracyclines in these populations possibly had a protective effect against infectious diseases [3, 4]. Only since the late 19th century did scientists began to observe antibacterial chemicals in action.

The beginning of the modern "antibiotic era" is usually associated with the names of Paul Ehrlich and Alexander Fleming. The German physician, Paul Ehrlich, theorized based on selective staining of microbes with certain synthetic dyes, that chemical substances ("magic bullets") can be synthesized to selectively kill certain bacteria without harming other cells. This led to the discovery of arsphenamine in 1909, which became the first modern antibiotic although Ehrlich himself referred to it as chemotherapy (i.e. the use of chemicals to treat a disease) [5]. The term "antibiotics" was coined over 30 years later by Selman Waksman, a Ukrainian-American inventor and microbiologist. It was in 1928 that Alexander Fleming, Professor of Bacteriology at St. Mary's Hospital in London, discovered penicillin, the first true antibiotic. Subsequently, with the earnest efforts of Howard Florey, Ernst Chain and their colleagues (at the Sir William Dunn School of Pathology at Oxford University), large-scale production of penicillin became possible turning this chemical from a laboratory curiosity into a life-saving drug [6, 7]. The 1940 to 1960 period marked the birth of many valuable antibiotics like chloramphenicol, neomycin, aureomycin, erythromycin and nystatin, and became known as the Golden Age of antibiotic discovery. Since then, antibiotics have been used for millennia to treat and control life threatening infectious diseases. Between 1962 and 2000, no new structural classes of antibiotics were introduced reflecting a serious innovation gap during the genomic era [8]. Most of the antibiotics introduced since then have only been minor modifications of previously known drugs (Figure 1) [5, 9].
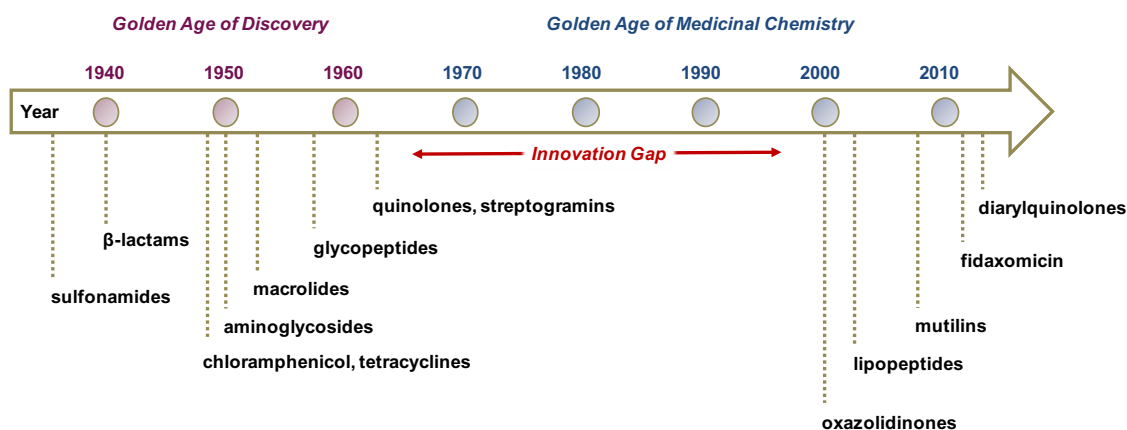
Figure 1. Timeline showing the 'Golden Age' of antibiotic discovery and the 'Golden Age' of antibiotic medicinal chemistry. Adapted from [8]

Antibiotics are vital to modern medicine with their applications reaching across a wide spectrum of areas from simple infections to major organ transplants and cancer chemotherapy. At the same time, many resistant microbes with simultaneous resistance to multiple structurally unrelated antimicrobial agents, a phenotype referred to as multidrug resistance (MDR), are emerging [10]. The development of resistance is a natural evolutionary phenomenon, and the long history of use, overuse and misuse of antibiotics are synergistically facilitating this process. An even greater concern is especially in the case of Gram-negative pathogens (such as *Acinetobacter baumannii*, *Klebsiella pneumoniae*, *Neisseria gonorrhoeae*, and *Pseudomonas aeruginosa)* as they represent most of the clinical strains endowed with multi, extreme, or total drug resistance (MDR, XDR, or TDR respectively) for which there are no promising antibiotics in the pipeline [11-13]. If unaddressed, the dramatic surge in the number of multidrug resistant bacteria (seen over the last two decades) and the dwindling number of pharmaceutical companies working on antibiotic discovery might collectively steer us into a post-antibiotic era. In fact, the most recent Global Risks Reports from the World Economic Forum have listed antibiotic resistance as one of the greatest threats to human health (Figure 2) [14, 15].
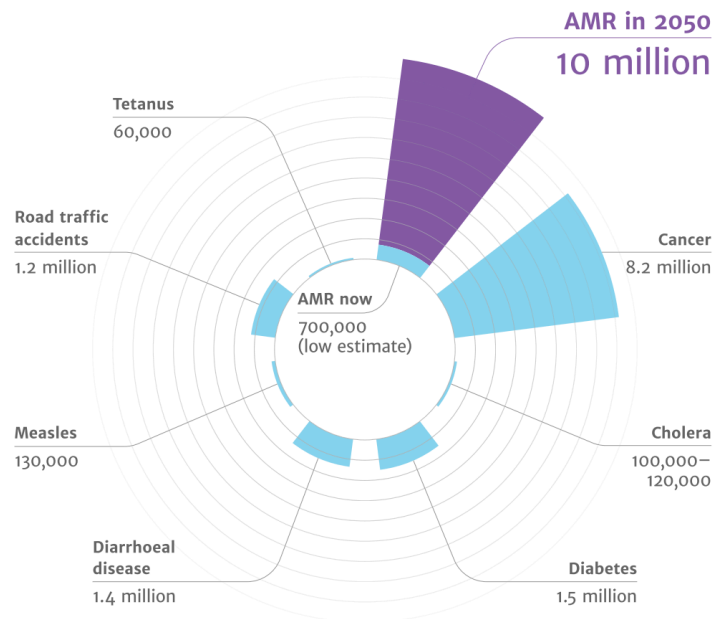
Figure 2. Deaths attributable to antimicrobial resistance (AMR) every year [14]

In general, resistance can be described as either intrinsic or acquired. Intrinsic resistance is inherent to the bacteria and is expressed by all (or almost all) strains of that particular bacterial species. It exists irrespective of the presence or absence of the antibiotic (e.g. resistance to many antibiotics due to the largely impermeable double membrane of Gram-negative bacteria like *Escherichia coli* and *P. aeruginosa*; natural resistance of anaerobes to aminoglycosides and Gram-negative bacteria against vancomycin). On the other hand, acquired resistance is obtained by mutation of chromosome or by the more common acquisition of genetic material (by any of the bacterial gene transfer methods: conjugation, transformation, transduction, transposition) that encodes for resistance function by alteration in the structural and functional features of the bacteria. Acquired resistance is limited to selected isolates of that particular species or group of microorganisms [16].

Bacteria have evolved a multitude of mechanisms that in solitude or in combination with each other function to counter the effectiveness of drugs and survive the deleterious effect of any antimicrobial agent, thus making the bacteria resistant to multiple drugs. The most common mechanisms [17-20] include (Figure 3):

(i)   the alteration of the macromolecular drug target either through chemical modification or by mutation to insensitive variants (e.g. alteration of penicillin binding protein in methicillin-resistant *Staphylococcus aureus* [21]);

(ii) the protection of the target via the production of immunity proteins, alteration of metabolic pathways (e.g. elimination of the requirement of para-aminobenzoic acid in sulfonamide-resistant bacteria for the synthesis of folic acid and nucleic acids);

(iii) the direct chemical modification or inactivation of the antibiotics (e.g. enzymatic deactivation of β-lactams by β-lactamases);

(iv) the altered transport of the compounds into the cell (e.g. reduced membrane permeability with downregulation of porins); and

(v) the increased active efflux of drugs out of the cell through efflux pumps.
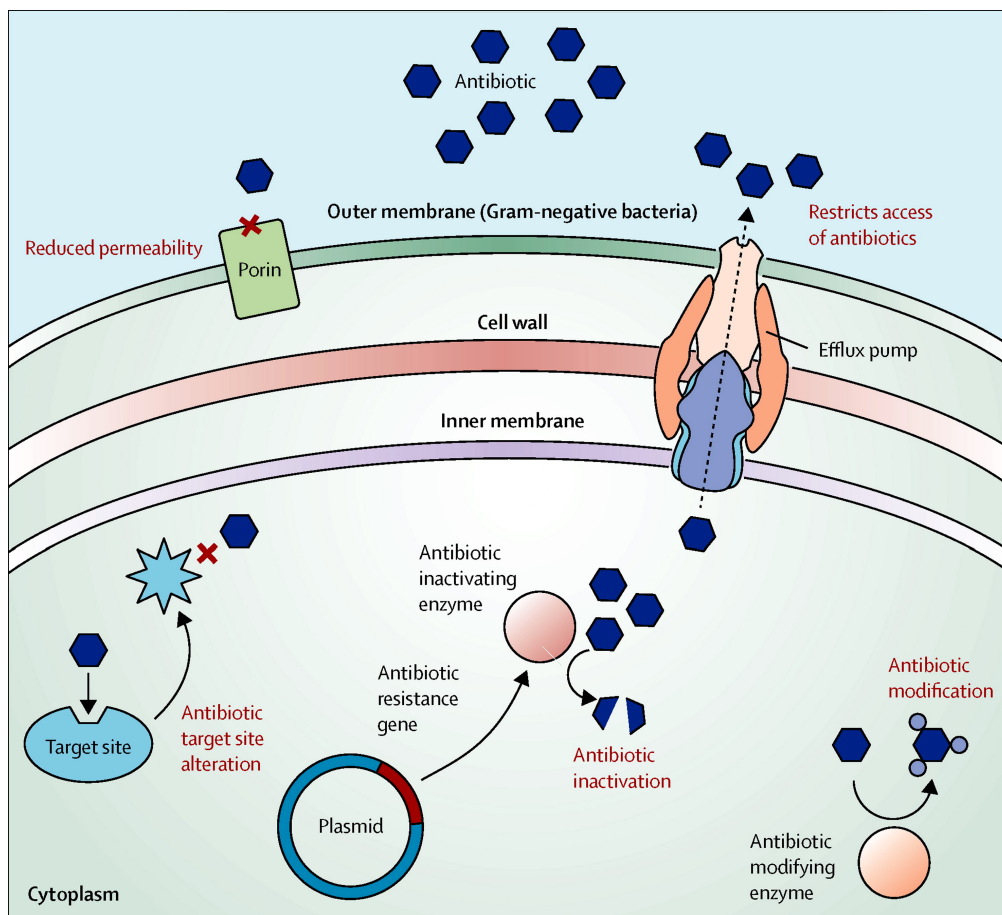
Figure 3. The most common mechanisms of antibiotic resistance in a bacteria [22]

1.2. Efflux-mediated resistance and multidrug resistance

Among the aforementioned resistance mechanisms, the efflux-mediated approach where molecular pumps actively export substrate molecules from cytoplasm to the external medium in an energy-dependent manner constitute a major mechanism for both natural and acquired resistance to a diverse range of clinically used antibiotics. Drug efflux pumps are ubiquitously expressed protein complexes residing in the bacteria membrane and were initially identified in the late 1970s for their role in drug-specific resistance. By early 1990s, they became notorious for their significant contribution to multidrug resistance expelling a wide range of structurally diverse antimicrobials

and toxins, thereby lowering their concentration inside the cell to sub-toxic levels [15, 23-25]. Extensive studies on prototypical pumps of various transporter families have greatly enhanced our understanding of their structures and transport dynamics, as well as their expression, regulation, and clinical ramifications. Efflux pumps enjoy a special status of being considered part of the primary survival kit of microorganisms as these non-specific pumps remove most of the xenobiotics from the cell interior to give the organism time to acquire resistance to that compound through more specific adaptive mechanisms [26, 27]. This way the efflux mechanisms likely contribute to a rapid emergence of resistance in the presence of antimicrobial selection pressure. Efflux mechanism also interplays with other resistance mechanisms to significantly increase the levels and profiles of resistance [28]. These characteristics pose major challenges to antimicrobial development and therapy especially in Gram-negative pathogens (e.g. *Enterobacteriaceae*, *Acinetobacter,* and *Pseudomonas*) in which the permeability barrier imposed by the outer membrane (OM) and the expression of chromosomally encoded drug efflux pumps work in unison with multitudes of specific resistance mechanisms [29]. The wide distribution and overlapping functions of MDR efflux pumps in bacteria hint at their probable role in physiological functions in addition to mediating intrinsic and acquired MDR [30]. A few of these functions include virulence, stress response, bacterial cell communication, colonization, fitness and intracellular survival, quorum sensing and transport of toxic compounds (as in the case of MacAB-TolC which is involved in exporting an extracellular peptide enterotoxin produced by enterotoxigenic *E. coli*) [31].

MDR pumps function as either a preexisting mechanism or an activated resource in response to numerous cellular stresses of antibiotics and other chemical substances such as bile salts, fatty acids, and ethanol that are often substrates of pumps relevant for drug resistance. For instance, AcrAB-TolC, the major pump belonging to Resistance Nodulation Division (RND) family of transporters in enteric bacteria, is upregulated under such stress conditions empowering the survival of these host organisms [17]. The major facilitator superfamily (MFS) pump, MdtM, also functions with AcrAB-TolC in a synergistic manner to protect *E. coli* from bile salt stress [32]. Also, NorM [33], a multidrug and toxic-compound extrusion (MATE) family transporter, and MacAB [34], a macrolide-specific ATP-binding cassette (ABC) superfamily exporter, protect the bacteria against oxidative stress. In the case of *P. aeruginosa*, several Mex pumps are upregulated in response to various stress triggers like membrane-damaging or ribosome-disrupting agents, reactive oxygen species, and/or nitrosative stress [35-37].

Apart from the previously mentioned functions, MDR pumps have also been identified to play a substantial but varying role in the formation and survival of biofilms in different species. For instance, the loss or inhibition of any of 9 MDR pumps or the TolC OM protein in *Salmonella*

impairs its biofilm forming tendency with reduced production of curli [38]. Similarly, *E. coli* mutants with a genetic deletion of one of the MDR pump genes results in reduced biofilm formation [39]. Efflux pumps are thus considered potential candidate drug targets for therapeutic interventions and open the gates for prospective combinatory products that may reinvigorate the current arsenal of decreasingly effective drugs.

1.2.1. Classification of efflux pumps

The transport proteins have been successfully classified by Milton Saier's group in over 600 families on the basis of functional and phylogenetic information (Transporter Classification Database: *http://www.tcdb.org*) [40]. The transporter genes identified in hundreds of sequenced bacterial genomes have also been documented in Ian Paulsen's database (*http://www.membranetransport.org*) [41]. Among the numerous families of transporters, the prominent ones responsible for MDR can be divided into 2 major groups based upon bioenergetical and structural features [30]:

(i) Primary transporters which hydrolyze ATP as a source of energy

    a. ATP-binding cassette (ABC) superfamily

(ii) Secondary transporters which utilize the proton (or sodium) gradient as a source of energy (the proton motive force is an electrochemical gradient in which the movement of hydrogen ions drives transport of the substrate [42]). They are further classified into four based on conserved consensus motifs and functional similarities:

    a. The major facilitator superfamily (MFS)

    b. The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily (which includes the multidrug and toxic compound extrusion [MATE] family)

    c. The resistance-nodulation-cell division (RND) superfamily

    d. The drug/metabolite transporter (DMT) superfamily (which contains the small multidrug resistance [SMR] family)

Recent studies have described novel drug transporter families including antimetabolite transporters (the AbgT family) [43] and the proteobacterial antimicrobial compound efflux transporters (the PACE family) [44]. While the major clinically relevant efflux systems in Gram-positive bacteria are usually non-RND pumps and often the singleton protein pumps belonging to the MFS, MATE, SMR or ABC, the RND efflux systems are by far the most important in Gram-negative bacteria [28].
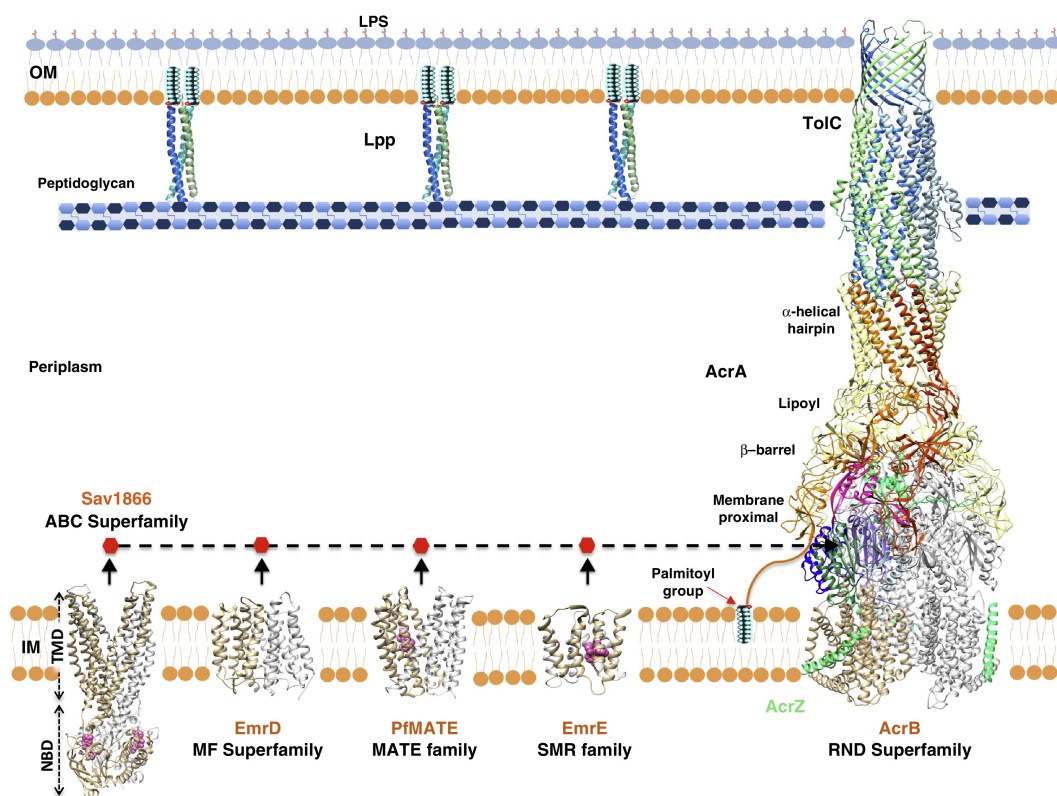
Figure 4. *Cooperative organization of MRD transporters within the cell envelope of Gram-negative bacteria (cytoplasmic facing side on the bottom). The protomers (Sav1866, EmrE and AcrB trimer) or symmetry related domains (EmrD and PfMATE) are differentially colored. The substrates (with the exception of the EmrD, for which only the apo-structure is available) are shown with pink spheres. Outer membrane (OM); Inner membrane (IM); Lipopolysaccharides (LPS); Lipoproteins (Lpp); Transmembrane Domain (TMD); Nucleotide-binding domains (NBD). The PDB codes are 3VVP, 2GFP, 3B5D, 2HYD and 3AOD for PfMATE, EmrD, EmrE, Sav1866 and AcrB, respectively. Slightly modified from [45]*

1.2.2.   Structural and functional mechanisms of efflux pumps

1.2.2.1. *ABC transporters*

ABC transporters are ubiquitous membrane systems involved in the efflux of toxins, metabolites and drugs. These transporters are typically composed of two cytoplasmic nucleotide-binding domains (NBD), and two hydrophobic transmembrane domains (TMDs) (Figure 4) [45]. The TMDs, responsible for drug recognition and transport, are fused to highly homologous NBDs, where ATP is hydrolysed. The NBDs possess the Walker A and B motifs, common to all ATP-binding proteins, and a Signature Motif, specific to ABC transporters [46]. It has been proposed that ABC efflux pumps were derived from secondary transporters by superimposition of NBD onto the transporter during evolution [46]. These transporters are found to house multiple drug binding sites, which is compatible with their broad substrate specificity and multi-drug binding capabilities. Ligand-binding and transport assays have shown that P-glycoprotein (P-gp; ABCB1; MDR1), the most extensively studied ABC member, has at least four pharmacologically distinct binding sites that are allosterically coupled [47, 48]. This family of exporters function with a mechanism termed the 'ATP switch' [49] where the nucleotide-driven interaction of the NBDs causes reorientation of

the TMDs and reduces drug affinity, thereby transporting the substrate (Figure 5A) [50, 51]. Ominous examples of ABC transporters are LmrA of *Lactococcus lactis*, MsbA conferring resistance to erythromycin in Gram-negative bacteria, and MacAB-TolC of *E. coli* able to expel macrolides.

1.2.2.2. MFS pumps

The MFS family is the largest group of secondary membrane transporters [52]. They are omnipresent systems that transport sugars, intermediate metabolites and drugs, and are the major contributors of MDR in Gram-positive bacteria. Most of these pumps are singlet transporters belonging either to 12- or 14- transmembrane segment (TMS) members of the drug/H$^+$ antiporters. In Gram-negative bacteria, they are located in the IM and transport drugs from the cytosol to the periplasm from where constitutive RND pumps, such as AcrAB-TolC and MexAB-OprM, may capture and efflux the drug molecules to the external medium thereby synergistically boosting the activity of these singlet pumps in producing resistance (Figure 4) [53, 54]. These transporters operate through an 'alternating access' mechanism (Figure 5B) in which drug-binding sites are alternately exposed to the outside or inside of the cell to uptake and release substrates. Similar to P-glycoprotein, the MFS pumps also contain several distinct (possibly overlapping) allosterically coupled binding sites [55]. There exists an indirect competition between the substrates and protons for binding to their respectively different locations, as shown in MdfA of *E. coli*, which might likely play a key role in their transport mechanism [56]. The most studied pumps of this family are NorA of *S. aureus* and its homologs Bmr and Blt of *Bacillus subtilis*, Tet pumps (12-TMS in Gram-negative bacteria and 14-TMS in Gram-positive bacteria) [57], and MdfA [58].

1.2.2.3. MATE pumps

Efflux pumps of the MATE family are mainly 12-TMS Na$^+$/drug antiporters that pump substrates from the cytoplasm to the periplasmic space (Figure 4) [59]. These transporters are widespread in bacteria and are also found in higher animals and plants. The common substrates of these pumps are cationic dyes, fluoroquinolones, and aminoglycosides. The MATE transporters have recently become a part of a new superfamily, the multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) flippase superfamily [40]. All MATE pump structures show a similar 12-TMS helix topology with an internal two-fold sequence similarity reflected in the tertiary structure [60-62] as N-terminal and C-terminal lobes. These pumps exhibit distinct binding sites for cation and drug enabling their simultaneous binding. The cation binding (with an unusual cation-pi interaction with an aromatic ring) and release promotes the interconversion between the drug-free and cation-bound configuration and drug-bound configuration as shown in the case of NorM of *Neisseria*

gonorrhoeae (Figure 5C) [61]. Transporters belonging to this family include the above mentioned NorM of *N. gonorrhoeae* and EmmdR of *Enterobacter cloacae* [36].

1.2.2.4. SMR pumps

Transporters in the SMR family [63] are the smallest drug efflux proteins known with just 100-120 amino acids composed into 4 relatively short intrahelical loops in the transmembrane (TM) region (Figure 4). They form either a homo- or hetero-dimer to exchange H^+ for pumping out either monocationic (e.g. ethidium and tetraphenylphosphonium) or dicationic (e.g. paraquat) compounds into the periplasm. The orientation of monomeric subunits in the dimer was a long debated issue with crystallographic data showing antiparallel arrangement of EmrE dimer while chemical cross-linking favoring a parallel arrangement [64]. Although the structure was withdrawn [65], this issue concluded on grounds that it can exhibit a dual topology and that the direction of insertion of the monomeric unit really does not matter for the efflux function [66, 67]. Structural plasticity and flexibility is the basis of multidrug recognition and transport in EmrE, the well studied pump of this family [68]. This transporter shows functional symmetry where conformational changes in the two monomers result in an interconversion between inward- and outward-facing states [69]. A fixed stoichiometry of two protons is exchanged per substrate molecule and this results in an electrogenic state for transport of monovalent cations but an electroneutral state for divalent cations [70]. The conserved membrane-embedded glutamate residue (Glu14) in each monomer is essential for proton and substrate binding. Hence, these transporters show an apparently simple, competitive, alternating site mechanism (Figure 5D) in which all substrates bind to the same site [71] and compete with protons for binding [64, 69]. In addition to EmrE, efflux pumps of this family include AbeS of *Acinetobacter baumannii* [72, 73] and KpnEF of *Klebsiella pneumoniae* [74].

1.2.2.5. RND transporters

RND transporters are further divided into seven families that include the hydrophobic/amphiphilic efflux (HAE) family, the heavy metal efflux (HME) family, and the SecDF protein-secretion accessory protein (SecDF) family [75]. Transporters of the HAE family, such as AcrB of *E. coli* which is considered the prototypical representative of the RND superfamily, [76] are the major clinically relevant efflux systems in Gram-negative bacteria also due to their extremely wide substrate specificity [36]. Unlike the vast majority of MDRs, which mostly pump out hydrophobic cations, the RND-type transporters from Gram-negative bacteria also interact with neutral, zwitterionic, and anionic compounds. Indeed, some of these transporters are able to recognize several tens of antimicrobials belonging to various classes (such as macrolides, quinolones, tetracyclines and β-lactams), and the different RND efflux systems in one species are altogether able to export a wide set of substrates ranging from lipophilic to amphiphilic molecules and finally

to toxic divalent cations [77-79]. Though RND transporters are characteristic of Gram-negative bacteria, certain Gram-positive bacteria like *Mycobacteria*, *Bacillus subtilis*, and *Staphylococcus aureus* also possess RND pumps [80-83]. Examples of RND pumps in Gram-negative bacteria are AcrAB-TolC and AcrAD-TolC of *E. coli*, and MexAB-OprM, MexCD-OprJ, MexEF-OprN, and MexXY-OprM of *P. aeruginosa* [36, 84]. Being representatives of these transporter systems investigated in this thesis, the next Section is devoted to a more detailed description thereof.
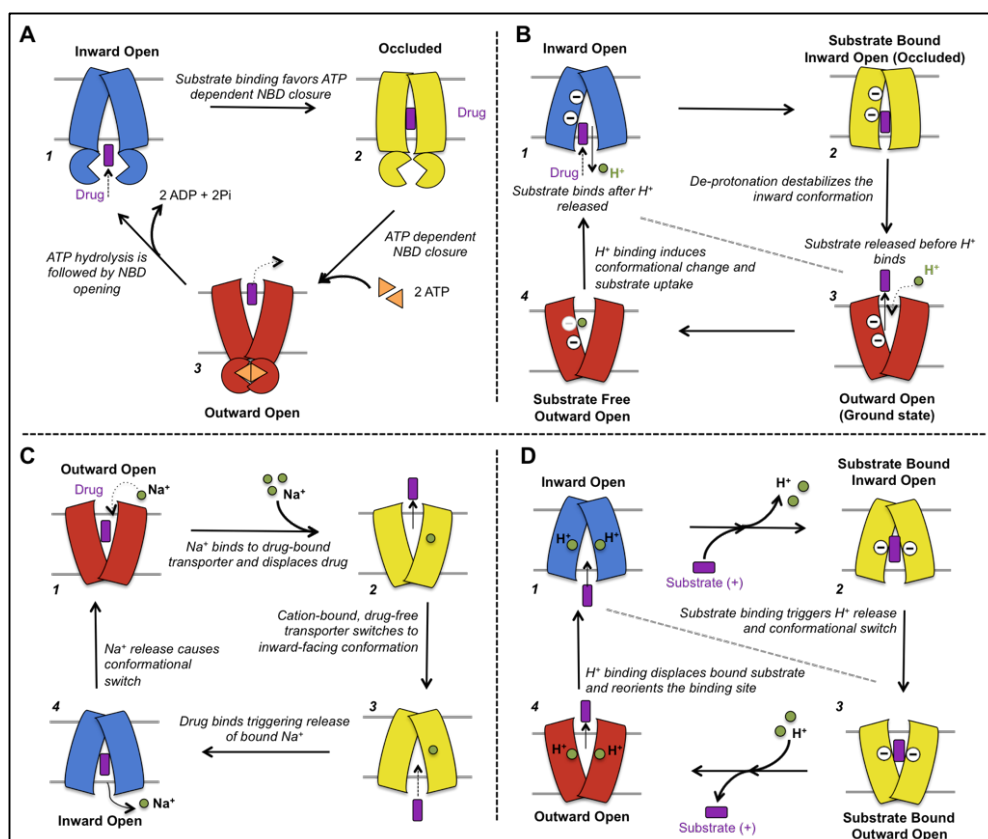
Figure 5. *Putative transport mechanisms proposed for members of the five major families of MDR efflux pumps. (A) Simplified drug transport cycle of ABC efflux pumps showing only three states: inward facing, occluded, and outward facing states. (B) Indirect competition mechanism in MFS multidrug/proton antiporters. (C) Antiport mechanism in members of the MATE family. (D) Alternating site transport mechanism in EmrE. Adapted from Refs. [45, 85, 86]*

1.2.3. RND transporters

1.2.3.1. Structural aspects

In Gram-negative bacteria, RND efflux systems span the whole periplasmic space from the IM to the OM by forming tripartite systems [87] (Figure 4) comprising an RND transporter protein (e.g. AcrB, AcrD, MexB, MexY) embedded in the inner (cytoplasmic) membrane; a periplasmic adaptor protein [a.k.a. membrane fusion protein (MFP); e.g. AcrA, MexA, MexX [88, 89]] located in the periplasmic space; and an outer-membrane protein resembling a long helical tunnel [OM protein (OMP); e.g. TolC, OprM [90-92]]. Recently, a small IM protein known as AcrZ was found to be

associated with AcrB of *E. coli* and might have potential role in enhancing the transport activity of AcrB for specific antibiotics like tetracycline, puromycin, and chloramphenicol [93]. No trace of any such protein or its homologs has been found in other RND transporters. Du *et al.* [94] presented a pseudo-atomic structure based on cryo-EM data of this entire tripartite system AcrABZ-TolC to explain the quaternary organization and key domain interactions and also proposed a cooperative process for channel assembly and opening.

The homo-trimeric RND transporter protein structurally resembles a jellyfish with each protomer comprising a total of 3 domains [95, 96]: (i) a trans-membrane domain (TMD) consisting of 12 α-helices embedded in the inner cytoplasmic membrane is the region where energy conversion via proton conduction takes place; (ii) pore (porter) domain in the periplasm is where substrate recruitment and transport mainly occur and (iii) a funnel domain (FD), also in the periplasm, couples the RND transporter to the OMP or to the hexameric assembly of MFPs in the constituted pump (Figure 6). A remarkable feature of OMP docking domain is its domain swapping within each protomer and between neighboring protomers tightening the trimeric structure.
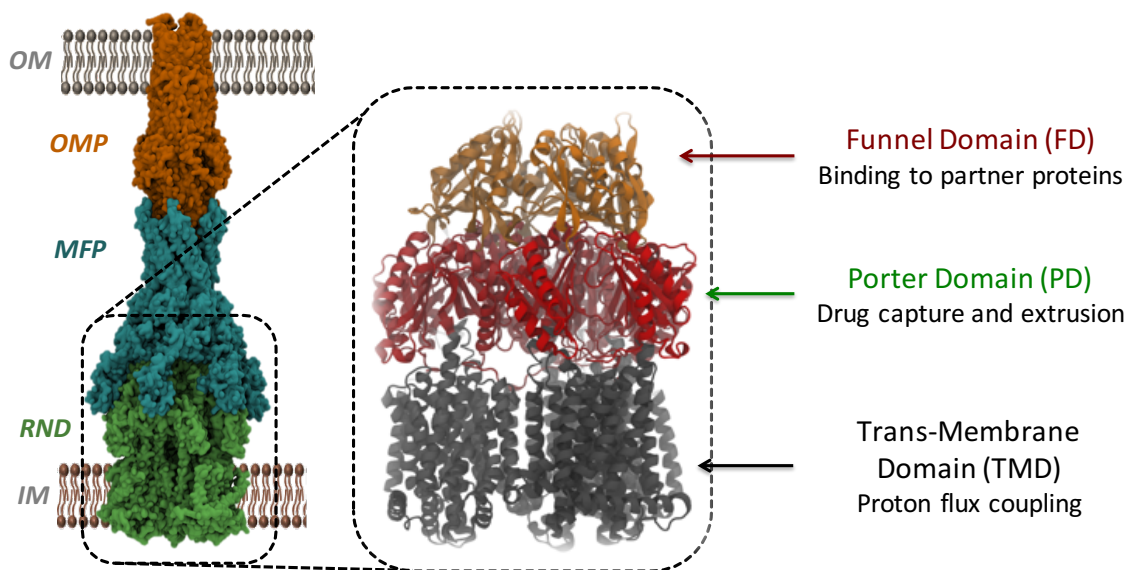
Figure 6. General structure of RND transporter showing the three main domains

Each protomer exhibits a pseudo-twofold symmetry (Figure 7) at both primary sequence as well as structural levels. It has been hypothesized that this property of RND transporters is due to evolutionary gene duplication and fusion events. A similar pseudo-twofold symmetry can also be seen in other transporter families, like the MFS, SMR, and ABC [97-99].
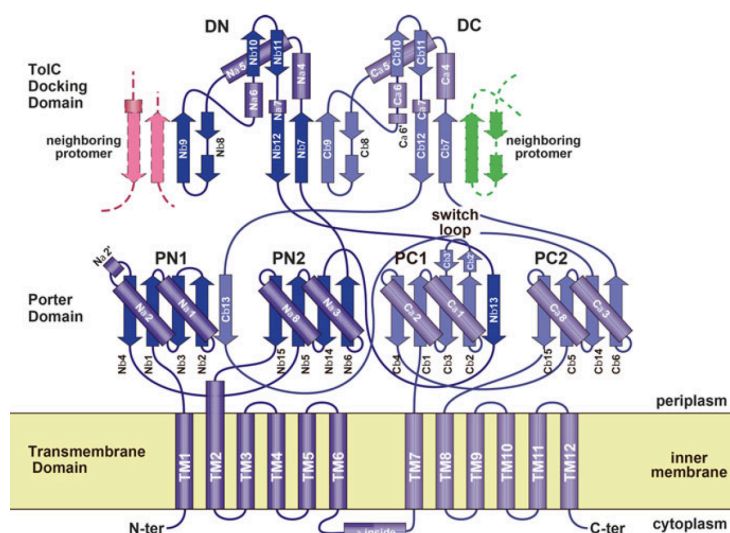
Figure 7. Topology diagram of AcrB monomer with the secondary structure elements indicated. The N- and C-terminal halves are depicted in dark and pale blue, respectively. Nα and Cα represent helices in the N- and C- terminal halves, respectively, while Nβ and Cβ represent strands in the N- and C-terminal halves of the headpiece, respectively. TM stands for transmembrane helices. The four subdomains (PN1, PN2, PC1, and PC2) of the porter domain, the two subdomains of the TolC-docking are shown in different shades. The hairpin structures for mutual insertion between TolC-docking domains of neighboring protomers are depicted in green and pink [100].

It has been postulated that the "resting state" of these transporters (i.e. the structure in the absence of substrates) corresponds to a symmetric structure in which each monomer assumes the same conformation, while the presence of substrates or inhibitors triggers conformational changes leading to an asymmetric structure [101]. The latter are characterized by three possible conformations of each monomer in the trimer, which were indeed interpreted as reaction cycle intermediates *Loose (L)* (a.k.a. *Access*) in which substrates become associated by loosely binding to an access pocket, *Tight (T)* (a.k.a. *Binding*) in which substrates bind tightly to a more deep binding pocket, and *Open (O)* (a.k.a. *Extrusion*) which corresponds to the drug-released state of a "functional rotation" mechanism [78, 102-105] (Figure 8).

1.2.3.2. Substrate binding pockets

The drug-binding sites in RND transporters like AcrB are within the periplasmic domain of the protein, in contrast to other MDR pumps discussed above [106], as evident from the drug-bound crystal structures of minocycline, doxorubicin, erythromycin, and rifampicin in the asymmetrical trimer configuration of AcrB [96, 102, 107]. There are two major substrate affinity pockets located along the substrate translocation pathway: Access (proximal) pocket and Deep (distal) pocket (Figure 8) [48, 53]. The nomenclature of these pockets refers to their proximity to the cell membrane and the entrance. The access pocket (AP) is located closer to the periplasmic bulk and is supposedly the preferred binding site for high molecular mass compounds (doxorubicin as dimer, erythromycin, and rifampicin) in the *L* conformation of the transport cycle. Recently, it has also

been identified to have β-lactam recognition sites [108]. The deep pocket (DP) is located much deeper within the substrate transport pathway, and is likely the recognition site for low molecular mass compounds (minocycline and doxorubicin) [109]. The DP is wide open in the *T* conformation but is in a collapsed state in the *L* and *O* conformations. Expansive shifting of the subdomains, accompanied with aromatic and hydrophobic residues in DP, produces a voluminous and hydrophobic binding pocket favorable for accommodating substrates by hydrophobic interaction or π-electron stacking interaction [110]. A characteristic flexible switch-loop, rich in glycine residues (therefore also named G-loop and Phe617-loop), separates the two pockets and modulates the path from the AP in the *L* conformation to the DP in the *T* conformation [111].
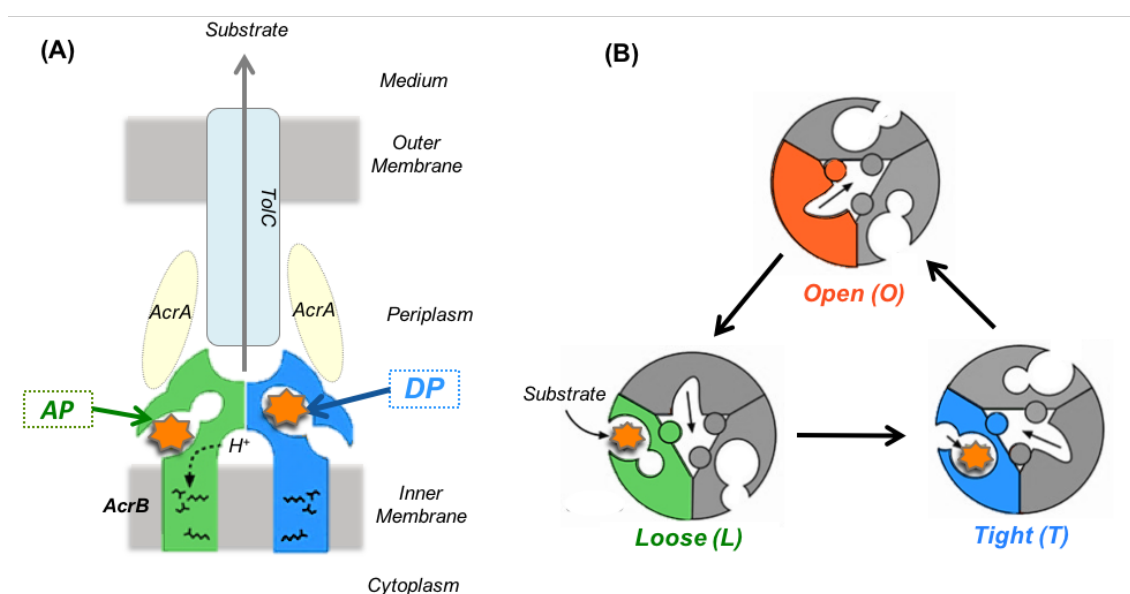
Figure 8. Schematic illustration of (A) the RND tripartite assembly highlighting the substrate transport pathway and (B) the proposed functional rotating mechanism shown by these pumps. Access pocket (AP) and deep pocket (DP) are also shown. Adapted from [112]

1.2.3.3. Substrate translocation pathways

The substrate translocation pathway is in the periplasmic domains of these transporters (Figure 9). So, cytosolic compounds which are substrates of this pump are transferred from the cytosol to the periplasm or to the upper leaflet of the cytoplasmic membrane by diffusion for the eventual uptake by their transporters. The first physiological pathway was proposed by Lomovskaya and Totrov [113] through a model calculation and also contemplating experimental observations made by Aires and Nikaido [114]. The entrances of the substrate translocation pathway in the periplasmic part of AcrB are the vestibule or cleft opening to the proximal part of periplasm and/or just above the outer leaflet of the cytoplasmic membrane [109, 115, 116]. Pathways starting from any of these entrances merge into the AP. Behind the distinct switch-loop which divides this pathway, lies the DP in the

latter part of the pathway. The DP has an exit formed by a gap between the N-terminal porter subdomains PN1 and PN2. Each exit joins the funnel-like opening in the Funnel domain. All along this pathway, a wide variety of chemical compounds (substrates) are recognized and bound at many sites. Binding characteristics for many other substrates predicted by *in silico* docking simulation analysis, protein engineering (including chimeric proteins) and mutagenesis, revealed the location of important residues around the pathway [117-123].
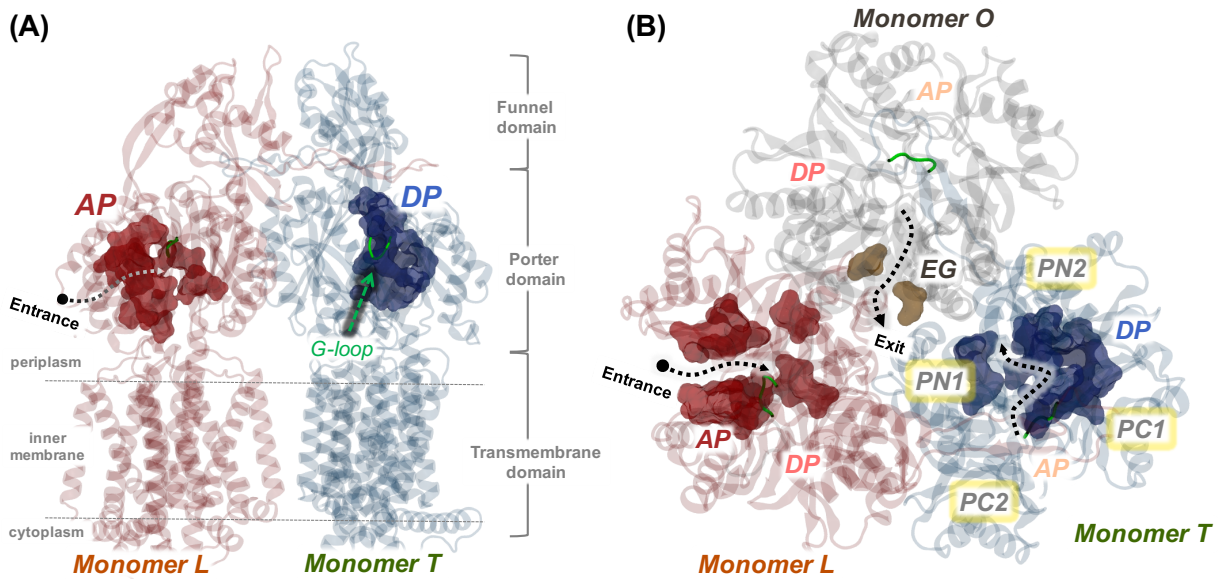
Figure 9 *Frontal and top views of an RND transporter showing their three main domains and the putative binding pockets (access pocket, AP, in monomer L and deep pocket, DP, in monomer T) along the substrate translocation pathway marked by dotted arrow. The G-loop at the interface of AP and DP is colored green. The four domains (PC1, PC2, PN1, PN2) enclosing the substrate translocation path are shown in monomer T as seen from the top view in (B). The exit gate (EG) from the porter domain to the central funnel is also shown in monomer O in the top view (B).*

1.2.3.4. Substrate transport mechanism

Substrate transport in RND transporters follows "functional rotation mechanism" [104, 124], according to which concerted but not necessarily synchronous cycling of the monomers occurs through any of the asymmetric states *L*, *T*, *O*, and back to *L* (Figure 10). During a complete functional cycle, occlusions and constrictions inside the pore domain are caused by varying interactions among the four subdomains of the porter domain (PN1, PN2, PC1, and PC2) in each protomer. This drives the binding and unidirectional transport of substrate by expanding and shrinking the putative binding pockets and thereby controlling the substrate affinity ("peristaltic pump mechanism" [103]) (Figure 10 Lower panel). In other terms, the substrate would gain access to the pore domain of the transporter via the *L* and/or *T* monomer, either from open clefts in the periplasm or through grooves between helices at the interface between pore and TMD [125, 126].

The substrate would then accommodate into a large binding pocket when the monomer assumes the *T* state and moves out toward the TolC docking domain upon a subsequent change to the *O* conformation. TolC docking domain may also change its conformation cooperatively with the MFP (i.e., AcrA) and TolC.
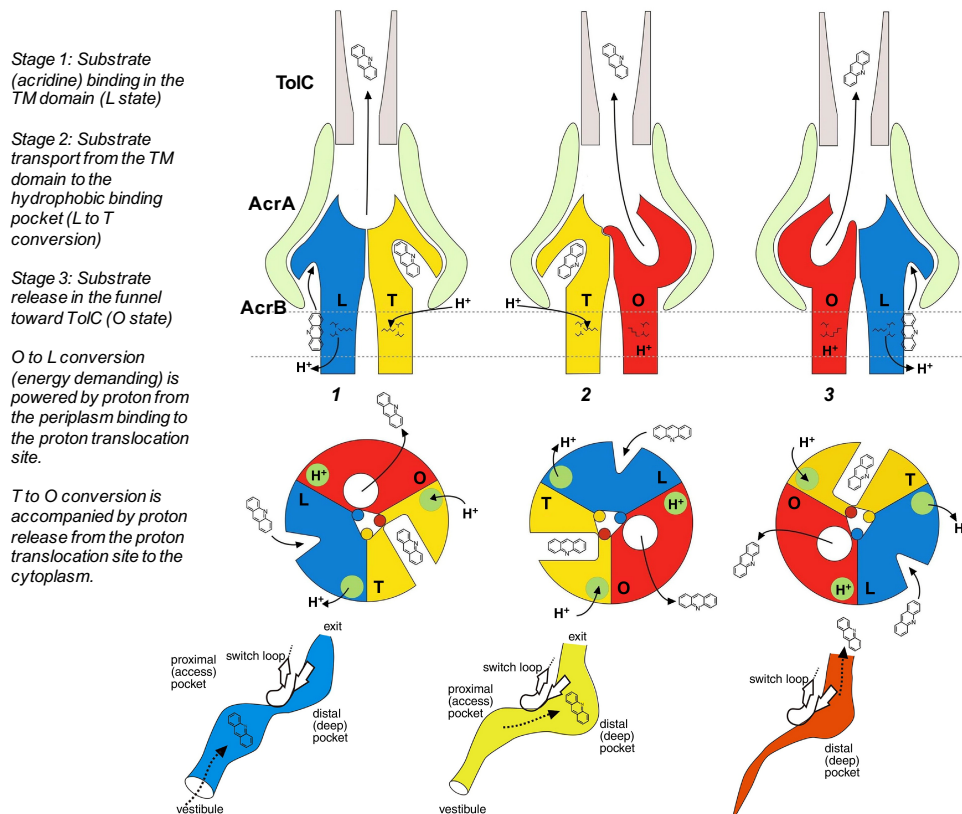
Figure 10. *Schematic representation of the functional rotating mechanism with 'peristaltic' pump-like motion in the RND transporter AcrB. The upper panel depicts a cross section (two protomers of the AcrB/TolC trimers or AcrA hexamer) through the full tripartite assembly; the middle panel is a view of the AcrB trimer, from the perspective of the porter domain of the pump; the lower panel depicts the passage of drugs between the drug-binding pockets in a protomer. Color codes correspond to the three conformational states shown in the other panels. In the upper panel, the TolC is proposed to remain in an 'open' state throughout the conformational switches in AcrB. Adapted from Refs. [92,98,104]*

These conformational changes are supposed to be triggered by protonation and deprotonation events of the functionally important titratable residues (Asp407, Asp408 and Lys970) present in the TM4 and TM10 central pair of the TMD helix bundle (Figure 11).
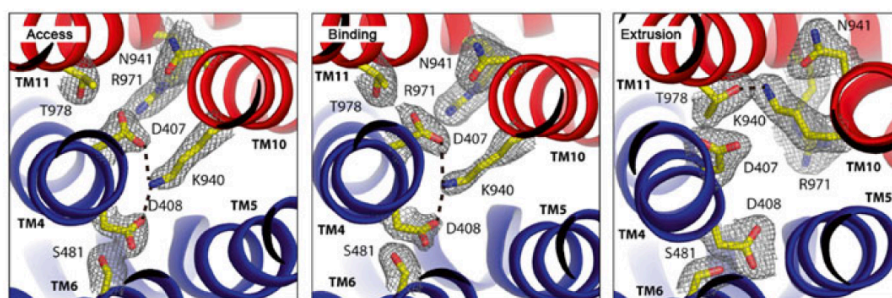
Figure 11. The switch-like conformational change shown by the key titratable residues, D407, D408, and K940, in the TM4 and TM10, involved in proton conduction. A sectional view of the transmembrane region viewed from the periplasm. Dotted lines show interaction between atoms. Slightly modified from [127]

1.3. Thesis Outline

This thesis would provide valuable insights into the substrate-transporter complementarity underlying the recognition and transport events in the principal RND transporters of *E. coli* (AcrB, AcrD) and *P. aeruginosa* (MexB, MexY). In particular, this would facilitate new drug design attempts by correlating the different substrate specificity patterns of the transporters to the physicochemical as well as topographical properties calculated on (or projected onto) the molecular surface of their multifunctional recognition sites. An extended application of this study would be in characterizing the probable substrate of a transporter from its binding pocket(s) features and vice-versa. The following initial questions form the basis of this work.

1.3.1. Initial questions

AcrB of *E. coli* and MexB of *P. aeruginosa* are the earliest reported RND transporters and the last fifteen years have seen a rush by structural biologists to provide information on key properties of both these single components of RND efflux pumps as well as on their assembly. This competition has generated a large number of structures of AcrB and MexB, thereby enabling valuable structural studies in establishing the foundation to explore similarities and differences in drug recognition and drug export mechanisms, and for future therapeutic inhibition of these transporters. However, in the case of their homologous transporters AcrD of *E. coli* and MexY of *P. aeruginosa* (Table 1), the absence of any experimental structure information has hindered such structure-based studies. In addition, despite sharing a comparable sequence homology and a high degree of global structural and functional similarity, these transporters show different substrate specificity pattern (Table 2). Moreover, concerning AcrB and MexB only few co-crystals with substrates have been obtained to allow characterization of the interaction pattern between the substrate and its transporter. In this regard, there has been no extensive comparison of the structural and physicochemical properties of

binding pockets among these transporters, which could explain their substrate specificities in terms of complementarity with the corresponding (substrate) binders.

Table 1. Sequence identity (similarity) between RND transporters of E. coli and P. aeruginosa calculated using EMBOSS Stretcher [128] (All values are in percentages)

Sequence	MexB	MexY	AcrB	AcrD
MexB	-	46.9 (65.5)	69.8 (83.2)	61.1 (76.5)
MexY	-	-	47.9 (67.0)	48.2 (66.7)
AcrB	-	-	-	65.3 (79.7)
AcrD	-	-	-	-

Table 2. Substrate selectivity of the principal homologous RND transporters of E. coli and P. aeruginosa

Substrates (Antibiotics)	*E. coli*		*P. aeruginosa*	
	AcrB [129]	AcrD [130, 131]	MexB [91, 132]	MexY [91, 131, 132]
Macrolides	Yes	No	Yes (poor)	Yes
Aminoglycosides	No	Yes	No*	Yes

Therefore, the questions to be answered were

1. What is the molecular rationale behind the different substrate specificity shown by the homologous RND pumps of *E. coli* (AcrB and AcrD)?

2. What is the molecular rationale behind the different substrate specificity shown by the homologous RND pumps of *P. aeruginosa* (MexB and MexY)?

3. How does the binding affinity of experimentally identified substrates, non-substrates and poor substrates correlate among these transporters?

1.3.2. Chapter overview

The effectively answer these questions, the flow of the thesis is organized into the following chapters.

Chapter 1: General introduction to MDR, resistance mechanism, and various efflux pumps with detailed description on RND transporters, the members of which are the prime topic of interest in this study.

Chapter 2: The theoretical background of the various computational methods adopted in the study.

Chapter 3: Our findings from the binding pocket characterization of AcrB and AcrD of *E. coli*

Chapter 4: Our findings from the binding pocket characterization of MexB and MexY of *P. aeruginosa*

Chapter 5: Molecular interaction and binding affinity analysis of substrates, non-substrates and poor substrates of Acr pumps in *E. coli*

Chapter 6: Optimization of the docking program, HADDOCK, for small-molecule docking studies in RND transporters

1.3.3. Comments

Part of this chapter has been published as a book chapter:

Ramaswamy, V. K., Cacciotto, P., Malloci, G., Ruggerone, P., & Vargiu, A. V., "*Multidrug Efflux Pumps and Their Inhibitors Characterized by Computational Modeling*", in Efflux-Mediated Antimicrobial Resistance in Bacteria. 2016, Springer. p. 797-831." [133]

Chapter 2

Computational Methods

2.1. Methodological background

Most biological effects are outcomes of actions at the molecular level, and hence gaining insights at the molecular and atomic levels is important for a detailed understanding of cellular processes. Certain experimental imaging techniques like high-resolution X-ray crystallographic analysis allow to reliably elucidate protein structure, as diffraction data is a direct output of structural averaging between the billions of molecules included in the crystals. However, the dynamic aspects of these molecules within crystals may also be averaged, resulting in just static snapshots of the protein molecules. Additionally, the temperature factor for the peptide chains within protein molecules may provide information about structural flexibility, but dynamic aspects of the molecules are predominantly lost. The observation of conformational dynamics, such as biologically relevant structural changes and the related molecular triggering mechanisms, require an extra element to be considered which is 'time'. A few experimental methods like NMR, electron paramagnetic resonance, infrared, and fluorescence spectroscopy provide information about the structural dynamics. Nevertheless, these methods are often indirect and difficult to interpret, or do not allow the observation of unstable transition states [134].

For membrane proteins like the RND systems as in our case, the access to structural and dynamic data by these experimental techniques is further limited due to various factors related to the experimental conditions and the technique. Hence, there has been only a little success particularly with a few members of the RND transporters (AcrB, MexB) in obtaining crystallographic structures [100, 102, 109, 111, 135] as the structure is dependent on the surrounding membrane. Separation of the proteins from the membrane can lead to irreversible structural changes, which inhibit ordered crystallization.

Our understanding of the structural aspects of MDR pumps from the available crystallographic structures has been significant but not sufficient to fruitfully assist structure-based drug design. Biomolecular simulations and other molecular modeling techniques [136] have addressed these limitations of experimental techniques and are valuable tools for complementing

such experimental procedures. These techniques are increasingly being used both for rationalizing existing data as well as for various predictions, for instance, about the drug recognition and binding, translocation mechanism and structural relations with the surrounding environment using three-dimensional structures.

Among the transporters discussed in this thesis, the experimental structures for only two (AcrB, MexB) of the four RND transporters (AcrB, AcrD, MexB, MexY) are available. Hence, we used homology modeling to build a hypothetical model for AcrD and MexY comparable in their structural (sterochemical) quality to the best available crystallographic structures of AcrB and MexB. Subsequently, Molecular dynamics simulations (MD) for all the four transporters were performed under near-physiological conditions in the presence of membrane, water and ions to obtain insights into the molecular basis of their biological activities. In this work, MD was also used to stabilize the homology model structures. Further, molecular docking studies were used to investigate the similarities and difference in the types of interaction shown by ligands with the putative substrate binding pockets in the Acr transporters of E. coli.

In this section, we briefly discuss the theoretical background of the computational methods used to study the RND transporters.

2.2. Theoretical framework

2.2.1. Homology modeling

To efficiently use MD simulations in the study of bimolecular systems, a proper initial structure of the system under study is necessary. Since protein functionality is strongly related to their three-dimensional structure [137-139], a bad initial model description may lead to largely different results [140]. Therefore, several experimental techniques have been developed to determine the structure of proteins, such as X-ray crystallography [141, 142], Nuclear Magnetic Resonance (NMR) spectroscopy [143-145] or Electron Microscopy (EM) [146-149]. However, an experimental description of the three-dimensional structure for all the proteins in nature is very far to be reached, since only a small fraction of them has been structurally determined to date [150, 151]. To fill this gap in the proteins structure knowledge, several computational methods have been developed in recent years. One of the main computations approaches is the so-called homology (or comparative) modeling that predicts three-dimensional structure of proteins [152-154].

The homology modeling procedure consists of three main steps [155]:

2.2.1.1.   Template search and sequence alignment

The modeling procedure requires at least one reference structure that has been previously obtained from experiments [152-154, 156]. This protein of known 3D structure (template) should have a significant amino acid sequence similarity with the protein of unknown structure (target) in order to be able to predict a realistic structure, [157]. The term significant refers to the definitions [158, 159] of the (i) easily detected relationship (>30% sequence identity), (ii) the twilight zone [160] with statistically significant sequence similarity (10%-30% sequence similarity) and (iii) the "midnight zone" [160] of statistically insignificant sequence similarity. Independent of the method used to align sequences, searching in the midnight zones of the sequence-structure relationship often results in false negatives, false positives, or alignments that contain an increasingly large number of gaps and alignment errors [159]. The amino acid sequence of the template and target proteins are then aligned optimally to with least gaps and maximum coverage. The success of this model building method relies on the sequence alignment between target and suitable templates. In this work, sequence searches from PDB of known structures using BLAST [161] was performed to identify related sequences. The sequences were aligned using ClustalOmega [162].

2.2.1.2.   Model building

Among the different programs available for building a 3D model, the MODELLER package [157-159] was used to build the proteins structures of AcrD and MexY studied in this thesis. The first step of model building by MODELLER is to compute distance and dihedral angle restraints for the target sequence, derived from the template three-dimensional structure using a database of alignments [163]. This program builds a structure with an extended strand for target and fold by satisfaction of spatial restraints [152] from the alignment of the target and its templates. The hydrogen bonding features and main chain dihedral angles are preserved from the template structure. Relationships such as the correlations between two equivalent $C\alpha$-$C\alpha$ distances from two related proteins are evaluated and expressed through probability density functions (pdf) and are used as spatial restraints. Then, stereochemical restraints like bond lengths, bond angles, dihedral angles and non-bonded atom-atom contacts are extracted from the Charmm22 force field [164] or from the so-called homology-derived restraints analysis, in which restraints are calculated from template protein structures (see Figure 1 for a schematic description of the modeling process). Models are then generated by minimizing the violations of all the restraints.

1. Sequence alignment
QITTELPGRTSAYRIAEVRPQVSGIILKRNFKEGSDIEAGVSLYQIDPAT Template
.·:·||||||||:|:|||||||||:·|||||·||||||::||··|||||||
TLNTELPGRTNAFRIAEVRPQVNSIILKRLFKEGSDVKAGQQLYQIDPAT Target

2. Spatial restraints

3. Three-dimensional structure
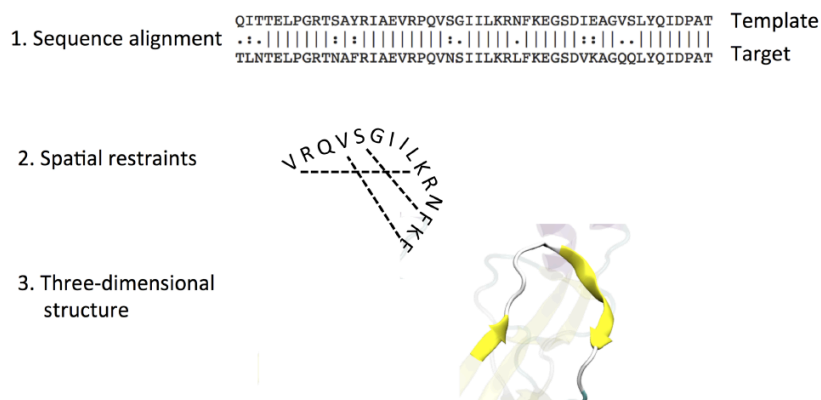
Figure 1. Main steps in the modeling process with MODELLER.

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

2.2.1.3. Model evaluation

It is important to check the quality of the constructed models using the quality assessment tools. The quality check should be able to verify the reliability of the model. It should be able to distinguish between properly versus improperly folded models, and evaluate steric and geometric properties of the models. Most of the methods have been developed using empirical data from globular proteins of known structure. In our work, the models quality was assessed using ProSa, Verify3D, WHAT_CHECK Packing, and PROCHECK. Most of these are available as web services at MolProbity (http://molprobity.biochem.duke.edu) and SAVES (http://services.mbi.ucla.edu/SAVES/).

ProSa [165, 166] is a program to check the potential error in 3D models of protein structures. It uses statistical potential of distance and surface-dependent statistical for Cα atoms to obtain the model. Verify3D [167] is the method that analyzes the 3D atomic model with its own amino acid sequence 1D. WHAT_CHECK Packing [168] checks all possible atom types in all possible positions around the fixed fragments. PROCHECK [169] addresses the stereochemical parameter of a protein. The structure is classified into highly populated to forbidden regions by the Ramachandran plot. It shows the Φ/ψ torsion angles for all residues in the structure. The entire process of homology modeling may have to be repeated until a satisfactory model is obtained. In

this work, molecular dynamic simulations of the satisfactory model in lipid bilayer environment is also used to evaluate the quality and to stabilize the model.

All the steps involved in the homology modeling run are represented as a schematically in Figure 2.
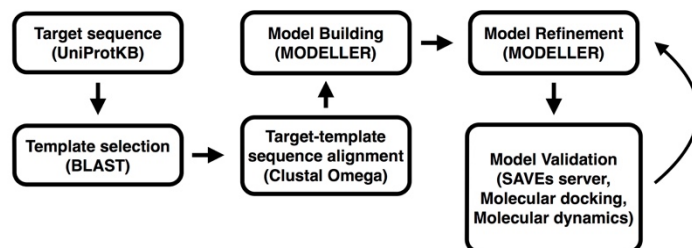
Figure 2. *Steps within a homology modeling run: from target and template sequence selection to alignment, homology modeling and finally models refinement.*

In the Methods section of chapters 3 and 4, we describe the specific details related to the homology modeling of AcrD and MexY, respectively.

2.2.2. MD simulations

MD is a powerful technique that can provide atomic level descriptions of molecular systems with high temporal resolution. They are also often employed to validate the stability of homology models. Atomic level simulations in the scale of several hundred nanoseconds are routinely performed for obtaining detailed insight into conformational changes and free energies of interactions, alongside identifying drug-binding locations, translocation processes, and interactions with the surrounding lipid bilayer [170]. This could lead to the identification of the movements intimately related with the translocation process, aiming for a better understanding of the first steps of the efflux mechanism [171]. Recently, the use of 'coarse-grained' simulations, where typically four atoms are combined into one particle, and biased MD simulations have increased dramatically allowing sampling of large conformational changes that would normally be inaccessible because of the large free energy barriers between such conformations and the consequent limitations due to lack of computational time [172-175].

The theoretical basis of MD simulations to describe biological world lies in statistical mechanics, which provides mathematical expressions that relate microscopic properties such as atomic positions and velocities, computed in the simulations, to macroscopic observables such as pressure, energy, etc. The thermodynamic state of a system is defined by the set of its macroscopic parameters, like temperature T, pressure P and number of particles N, while the microscopic state is defined by the positions, r, and momenta, p of all the particles of the system. The collection of these

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

$$\langle A \rangle_{ensemble} = \iint dp^N dr^N \, A\,(p^N, r^N)\, \rho\,(p^N, r^N)$$

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

$$\rho\,(p^N, r^N) = \frac{1}{Q}\, e^{-H(p^N, r^N)/k_B T}$$

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

$$Q = \iint dp^N dr^N \, e^{-H(p^N, r^N)/k_B T}$$

As the integral over the phase space is extremely difficult to solve, MD simulations calculate each point of the phase space step by step. Therefore, to calculate the ensemble average properly the simulations must explore all the possible states accessible by the system. This means that, in principle, the trajectories need to be long enough to explore the whole phase space. However, such condition is not always guaranteed, even with multiple MD simulation [176].

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

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

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

$$\langle A \rangle_{ensemble} = \langle A \rangle_{time}$$

The idea that support the ergodic hypothesis is that, if the system is free to evolve indefinitely in time, it will pass through all the accessible states. To preserve the validity of such hypothesis, it is thus necessary that MD simulations generate enough conformations that might be representatives of all the states of the system. As macroscopic experiments might be performed in different conditions so do MD simulations, and the choice of the ensemble does make a difference when computing the observables averages [177]. However, techniques exist to relate fluctuations in different ensembles [178] allowing a comparison of the results of MD simulations performed in different conditions. Several algorithms have been developed to mimic the different ensembles properties, such as the Andersen, Nosé-Hoover, Berendsen and Langevin dynamics algorithms [179-183]. With these methods, it is thus possible to perform MD simulations in:

   i.   *NVE* (micro-canonical) ensemble: The number of particles N, the volume V and the total energy E are constant;

  ii.   *NVT*: The number of particles N, volume V and temperature T; are conserved. This is one of the main ensembles used in MD simulations

 iii.   *NPT*: The constant number of N particles, pressure P and temperature T are constant

2.2.2.1.   Classical Mechanics (classical MD)

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

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

$$\sum_j F_{ij}(t) = m_i \frac{\partial^2 r_i(t)}{\partial t^2}$$

However, if the classical approximation considerably simplifies the physical description of the system, it makes theoretically impossible to find an analytical solution for the above equation. In fact, such solution for a system of N interacting particles is possible only for $N = 2$. The first step to move around this limit is to assume that the force F_i (t) that acts on particle i at the time t is constant for a Δt time interval, so that the new particle position $r_i(t+\Delta t)$ and velocity $v_i(t+\Delta t)$ can be computed. Several algorithms based on Taylor series have been developed to integrate the equation of motion [184-188]. An example is reported in the equation that describes the *Verlet Leapfrog* algorithm, that evaluates the positions at time $t + \frac{1}{2}\Delta t$ by using the velocities at time $t - \frac{1}{2}\Delta t$.

$$F_i = -\nabla U_i, i = 1 \dots N$$

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

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

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

$$U_i = \sum_j U_{i,j}^{bonds} + \sum_j U_{i,j}^{angles} + \sum_j U_{i,j}^{dihedrals} + \sum_j U_{i,j}^{electrostatic} + \sum_j U_{i,j}^{vdW}$$

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

Bonded Interactions: This class includes the contributions of the *bond, angles* and *dihedral* interactions and each term has the following functional form [189]:

$$U^{bonded} = U^{bonds} + U^{angles} + U^{dihedrals}$$

$$U^{bonds} = \sum_{bonds} k_r \left(r - r_{eq} \right)^2$$

$$U^{angles} = \sum_{angles} k_\theta \left(\theta - \theta_{eq} \right)^2$$

$$U^{dihedrals} = \sum_{dihedrals} \frac{V}{2} \left(1 + cos(n\emptyset - \gamma) \right)$$

While a more realistic description for the bonds interaction is provided by the Morse potential, for small deviations from the equilibrium distance r_{eq}, it can be approximated to the harmonic interaction in equation for bonds above, as described in Figure 3 and Figure 4a. In this approximation, the interacting particles are described as rigid bodies, connected by a spring with an empirical k_r constant. Analogous to bonds, angle θ between two particles also fluctuates around a reference value as shown in Figure 4b.
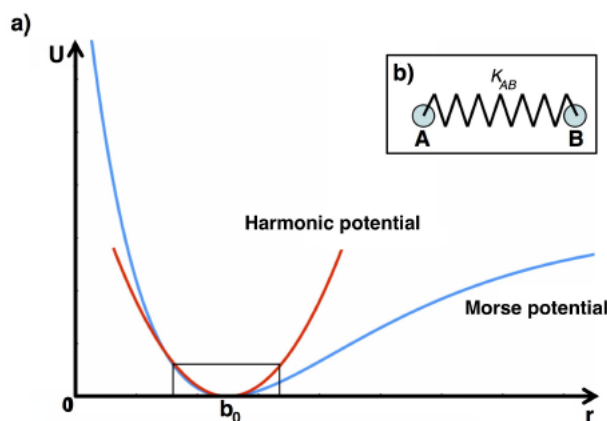
Figure 3. a) *The approximation of the Morse potential [Blue] with the harmonic potential [Red] for small deviations from the average distance b_0;* ***b)*** *The harmonic interaction between particle A and particle B with a spring with a k_{AB} constant*

The potential energy associated with this motion is a harmonic function, with k_θ as spring constant. The torsional motion along a covalent bond is taken into account in the dihedral potential function, where ϕ (see Figure 4c) is the angle between the *ijk* and *jkl* planes, *n* is the number of the minima and γ is a phase angle. In addition to the last term of the bonded interactions, the equation

below describes the improper dihedral contribution, where the angle ξ is chosen to keep planarity in a molecular structure.

$$U^{Improper\ dihedrals} = \sum_{dihedrals} \frac{1}{2} k_\xi \left(\xi - \xi_{eq} \right)^2$$
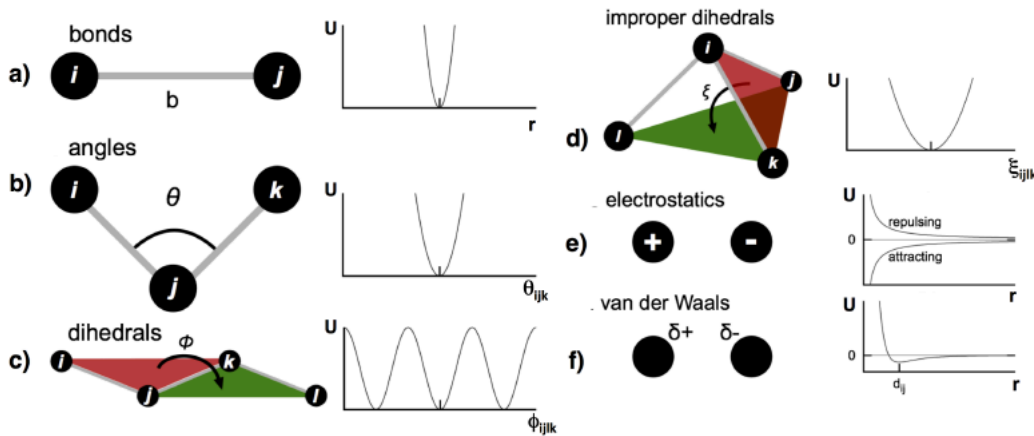
Figure 4. a-c) *Bonded interactions for the bonds, angle and proper dihedral terms.* ***d)*** *Bonded interaction for improper dihedral contributions.* ***e-f)*** *Non-bonded interaction terms described by the Coulomb and the Lennard-Jones potentials*

Non-bonded Interactions: Non-bonded interactions are described by the electrostatics and van der Waals (vdW) interactions. The Coulomb potential describes the attractive and repulsive interactions between charged atoms at distance r_{ij} (see Figure 4e) as described in equation below, where the ϵ_0 is the vacuum dielectric constant. Electrostatic interactions are, however, possible also between uncharged atoms, whose asymmetric charge distribution may form transient dipoles as shown in Figure 4f.

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

The dipoles interact through the so-called van der Waals potential, parametrized in the Lennard-Jones potential in the equation below, where the $\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}$ is the repulsive long-range term and $\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}$ is the attractive short-range term, ϵ is the depth of the potential well and σ_{ij} is zero-potential distance between particle i and j.

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

Force Fields: Several force fields have been designed to efficiently compute the potential energy terms described above such as the force fields in the AMBER package [190-192], used for the simulations in this thesis, the force fields in the CHARMM package [193, 194], GROMOS [195-197] and MARTINI, developed as coarse-grained force field [198].

2.2.2.2. Periodic Boundary Conditions (PBC)

Even in the best laboratory conditions, real experiments are always influenced by the surrounding boundaries. Such influences can, in most cases, be neglected or taken into account as systematic errors. In the computational world, the simulated experiments are limited by the simulation box that contains the system. Since the box is not infinite, the edge effect will influence the evolution of the simulation. To avoid this limitation, the so-called Period Boundary Conditions (PBC) have been developed. The idea is to simulate an infinite system, composed by the copies of the finite initial box. The result is a three-dimensional system where identical images are arranged in a periodic array. In this extended box, the atoms can leave the original simulation box on one side and enter the simulation box as periodic image on the opposite side as described in Figure 5. As consequence, the total number of particles in the system is conserved. One drawback of the PBC method is the possible interaction, e.g. electrostatic, between the system under study and its periodic images [199]. In order to avoid such interactions a proper choice of the simulation box is needed.
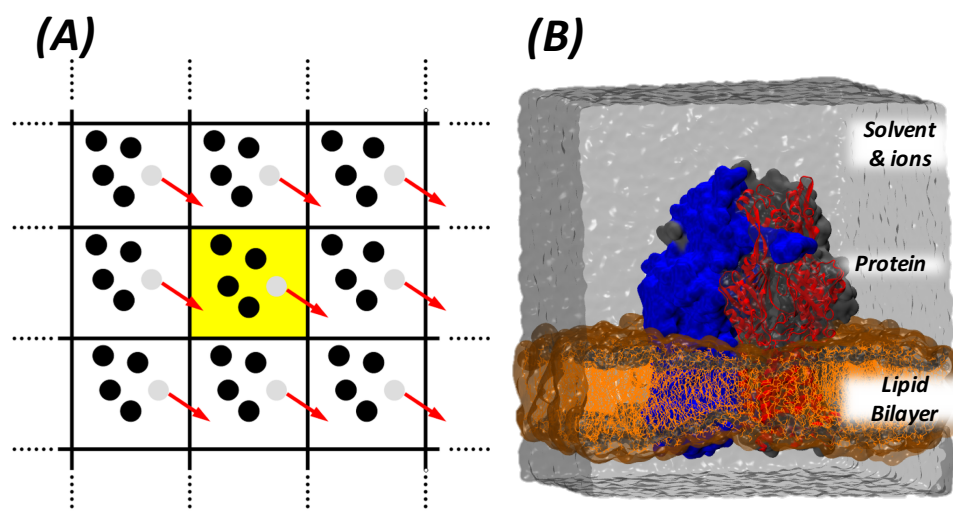
Figure 5. *(A) Schematic representation of periodic boundary conditions [200]. Atoms that leave a simulation system a side enter the original system on the opposite side. (B) Simulation box of RND transporter (AcrB identified by PDB code 4DX5 [111]) protein embedded in POPE lipid bilayer membrane. The protein chains are colored differently and shown in surface representation except for one chain shown as cartoon representation.*

However, if a very large box would be potentially the best choice it is necessary to consider that the higher the number of the atoms the slower the simulation is. Another point that should be taken into account in the PBC context is how long-range interactions can be computed during the simulation of an infinite three-dimensional periodic system.

As example, the Coulomb potential might include interactions between images that should be avoided. Since a simple truncation of the potential is not allowed [177, 201] it became essential to find a way to efficiently compute long-range interactions in the PBC context. A way to take into account the long-range interaction is through Ewald summation and the Particle Mesh Ewald (PME) methods [202-206]. In a system of N charged particles located in a cube of volume L^3, under PBC, the Coulomb potential in equation can be written as,

$$U^{electrostatic} = \frac{1}{2} \sum_n \sum_{ij}^{N} \frac{q_i q_j}{4\pi\epsilon_0 |r_{ij} + nL|}$$

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

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

where *f(r)* is a generic function of the distance r between the particles. The choice of the function *f(r)* has been proved to be crucial for the efficiency of the method, although different choices yield to comparable results [207]. The identity (equation above) is introduced to separately consider the two main complications of the Coulomb potential, i.e. the large variation at small r and the slow decay for large r.

Therefore, the first term of the identity (equation above) should be negligible beyond a cutoff r_{max} in order to avoid the divergence when r $\rightarrow$ 0. At the same time, the second term should vary slowly for all r so that its Fourier transform could be represented by limited number of reciprocal vectors. An example of one traditional function [205] that fits these requirements is the function $2\pi^{-1/2} \int_r^\infty dt \exp(-t^2)$

The methodology used in this thesis: MD simulations of the crystal structure of AcrB (PDB ID: 4DX5), MexB (PDB ID: 3W9I) and the homology models of AcrD and MexY were carried out using AMBER14 molecular modeling software [191]. The sequence of steps involved are summarized in the Figure 6.
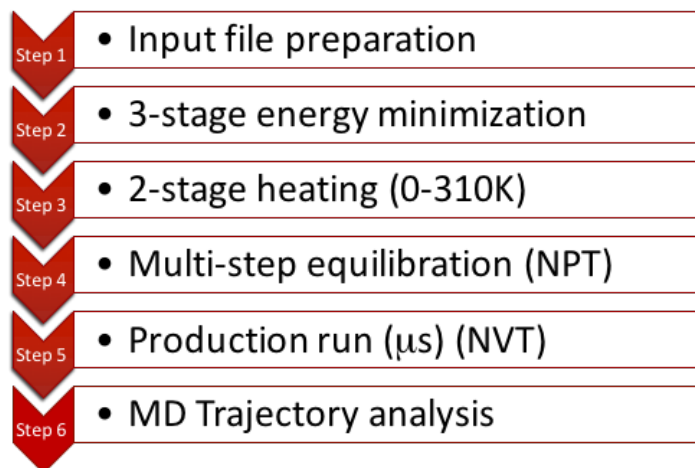
Figure 6. MD methodology showing sequence of steps starting from input file preparation to trajectory analysis.

1. <u>Input file preparation:</u> The topology and the initial coordinate files for these apo-protein structures were created using the *LEaP* module of AmberTools14 after suitably embedding them in 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) bilayer patches, solvating with explicit TIP3P water model, and neutralizing the residual charge of the system with required number of randomly placed K^+/Cl^- ions [208-211]. The ions count was suitably adjusted to account for an osmolarity of 0.15M KCl. Embedding of the protein into a pre-equilibrated POPE bilayer patch was done using the PPM server [212] and subsequently the CharmmGUI tool [213]. The lipid residue nomenclature was converted from the CHARMM to AMBER format using the *charmmlipid2amber.py* python script provided with AmberTools. The central pore lipids were added after calculating the number of lipids to be added to each leaflet by dividing the approximate area of the central pore by the standard area per lipid of POPE molecules [192]. The ff14SB [214] version of the all-atom Amber force field was used to represent the protein systems while lipid14 [192] parameters were used for the POPE bilayer. Periodic boundary conditions were used and the distance between the protein and the edge of the box was set to be at least 15 Å in each direction.

2. <u>Energy minimization:</u> Multi-step energy minimization with a combination of steepest descent and conjugate gradient methods was carried out using the *pmemd* program

implemented in AMBER14 to relax internal constrains of the systems by gradually releasing positional restraints.

3. <u>Heating</u>: Following this, the systems were heated from 0 to 310 K by a 1 ns heating (0-100 K) under constant volume (NVT) followed by 5 ns of constant pressure heating (NPT) (100-310 K) with the phosphorous heads of lipids restrained along the Z-axis to allow membrane merging and to bring the atmospheric pressure of the system to 1 bar. A Langevin thermostat [215] (collision frequency of 1 ps^{-1}) was used to maintain a constant temperature.

4 & 5. <u>Equilibration and Production runs</u>: Multiple short equilibration steps of 500 ps under anisotropic pressure scaling (Berendsen barostat) in NPT conditions were performed to equilibrate the box dimensions. A time step of 2 fs was used during all these runs, while post-equilibrium MD simulations were performed with a time step of 4 fs under constant volume conditions after hydrogen mass repartitioning [216]. The particle-mesh Ewald (PME) algorithm [203] was used to evaluate long-range electrostatic forces with a non-bonded cutoff of 9 Å. During the MD simulations, the length of all R-H bonds was constrained with SHAKE algorithm [186]. Coordinates were saved every 100 ps. Multi-copy μs-long MD simulations for each system was performed for better statistical reliability in the obtained data.

6. <u>Trajectory Analysis</u>: Trajectory analysis was done using *cpptraj* module [217] of AmberTools14 and VMD1.9.1. Several analysis tools were used to calculate the pocket descriptors all of which are explained below. Graphs were generated using the xmgrace [218] plotting tool.

2.2.3. Molecular docking

Molecular docking is a widely used computational method that aims at predicting the bound conformation and the binding affinity of a complex from the interactions between a small molecule (ligand) and a larger macromolecule (receptor). The original concept of molecular docking is that of lock (the receptor protein) and key (the ligand). However, experimental observations of ligand-receptor binding reveal, in most cases, conformational changes of the native unbound structures. The historical lock-and-key model has therefore evolved to the so-called induced-fit model and the available docking software try to predict protein-ligand association taking into account flexibility (compilation of available docking software and their main features in Ref. [219]). In particular, a full-flexible docking in which both ligand and protein are flexible should be the best option but the high computational cost prevents its massive application. The preferred approach is therefore a semi-flexible docking in which the protein is kept fixed in space (or only some torsional angles in the active site are set as rotatable [220]) and different conformations of the ligand are generated on

the fly during the docking process. An alternative and reliable approach is to rigidly dock ensemble of conformations of a ligand on ensemble of receptor conformations [221, 222]. In this approach the conformational analyses of ligand and receptor are made externally, for example they are generated from MD simulations. To compute relative affinities of putative poses of ligands docking programs use the so-called "scoring function", which approximates the free energy of binding by means of a "master equation" in which the several terms contributing to the binding are summed up. Molecular docking has proven to be a powerful tool to investigate the mechanism of recognition of different compounds by the MDR transporters AcrB and MexB [223-226].

In our study, we used the ensemble docking approach with AutoDock VINA [220], which is among the most used programs for molecular docking [219, 227]. AutoDock VINA uses an empirical scoring function to evaluate the binding affinity between the molecules and employs the iterated local search global optimizer for global optimization, achieving a significantly improved speed and better accuracy of the binding mode prediction. It is based on the definition of a rectangular grid of points where the molecular potential of the receptor is evaluated and sensed by the ligand during the docking process, and allows to set rotatable bonds for ligand and receptor (although flexible backbone or induced-fit are not taken into account explicitly).

2.3. Analysis methods

MD simulations produce trajectories of atomic positions (and optionally velocities and energies) as a function of time. Several methods are available to extract meaningful information from the MD trajectories. In this section, we describe the main analyses performed for the study described in this thesis.

2.3.1. RMSD Analysis

The evaluation of the Root Mean Square Deviation (RMSD) is in most cases the first analysis that gives an overview of the global behavior of the system during the MD simulation. Although it cannot be used to estimate the convergence it is useful to check whether the simulation has not reached the convergence [228]. To evaluate the RMSD between two structures, the first structure is superimposed to the reference and the RMSD value is computed as described in equation below. Generally, Cα atoms, backbone atoms or heavy atoms are applied as reference atoms in this analysis.

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

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

2.3.2. Cluster analysis

A useful way to study the MD simulation trajectory is to group molecular configurations into subsets based on the similarity of their conformations [229, 230]. Clustering algorithms collect data elements (points) in sets called clusters using a function that measures the distance between pairs of points. The points among in each cluster are more similar to each other than to points of other clusters and thus, the size of each cluster are likely to be different as higher energy states will be less populated than lower energy states. Several algorithms have been developed to perform cluster analysis such as hierarchical, centripetal, complete-linkage, centroid-linkage, average-linkage, means, Bayesian and COBWEB. Among these, we used the average-linkage algorithm in the *cpptraj* module of AMBER [217] in this thesis, in which the distance between two clusters is defined as the average distance between each point in one cluster and every point in the other clusters. The average-linkage algorithm was selected among the others as it was proved to be one of the most useful algorithms for such analysis [230].

We performed clustering of the MD trajectories using this method to sample structures still retaining the large conformational space sampled during the MD runs. The cluster representatives were for used for the analysis of various pocket descriptors (volumes, the molecular lipophilicity indexes and the electrostatic potential) described later as well as for use as receptor structures in molecular docking. In this approach, we clustered the trajectory independently based on RMSD (cutoff set to 3 Å) of the AP in *L* monomer and the DP in *T* monomer for both the proteins. All non-protein molecules were stripped from the trajectory during post-processing to reduce additional memory usage and to speed up file processing.

2.3.3. Pocket descriptors

The various pocket descriptors used to characterize the binding site in this study were calculated using specific programs after validating their applicability to RND systems by assessing results against available crystal structures and experimental data.

A. Cavity Volume

Evolution of size and shape of the AP and DP during the MD simulations was examined using the two-probe sphere method of *rbcavity* program bundled in rDock suite [231]. This allows obtaining detailed information on the pocket volume and plasticity of the site. In this method, the potential binding site volume was identified by a fast grid-based cavity detection algorithm [232] within a sphere of radius 14 Å, centered over the pockets, using large and small probe radii of 6.0 Å and 1.5 Å, respectively.

The cavity detection algorithm includes the following steps (Figure 7):

i. A grid of resolution GRIDSTEP (default 0.5 Å) is placed over the cavity mapping region, encompassing a sphere of radius=RADIUS (14 Å in our study), centre=CENTER. Cavity mapping is restricted to this sphere. All cavities located will be wholly within this sphere. Any cavity that would otherwise protrude beyond the cavity mapping sphere will be truncated at the periphery of the sphere.

ii. Grid points within the volume occupied by the receptor are excluded (colored red). The vdW radii of the receptor atoms are increased by VOL_INCR (default 0.3 Å) in this step.

iii. Probes of radii LARGE_SPHERE (6Å in our study) are placed on each remaining unallocated grid point and checked for clashes with receptor excluded volume. To eliminate edge effects, the grid is extended beyond the cavity mapping region by the diameter of the large sphere (for this step only). This allows the large probe to be placed on grid points outside of the cavity mapping region, yet partially protrude into the cavity mapping region.

iv. All grid points within the cavity mapping region that are accessible to the large probe are excluded (colored green).

v. Probes of radii SMALL_SPHERE (1.5 Å in our study) are placed on each remaining grid point and checked for clashes with receptor excluded volume (red) or large probe excluded volume (green).

vi. All grid points that are accessible to the small probe are selected (yellow).

vii. The final selection of cavity grid points is divided into distinct cavities (contiguous regions). In this example shown here (Figure 7) only one distinct cavity is found.

This approach is a faster grid-based implementation of previously published methods [233]. The effect of Step 3 (large sphere) is to eliminate all large cavities and smooth convex regions of the target surface. The radius of the small sphere used in step 5 represents a small ligand fragment or solvent molecule. Taken together, steps 3 to 6 identify regions of the grid corresponding to 'deep' cavities, i.e. those that are accessible to small spheres but not to larger ones.
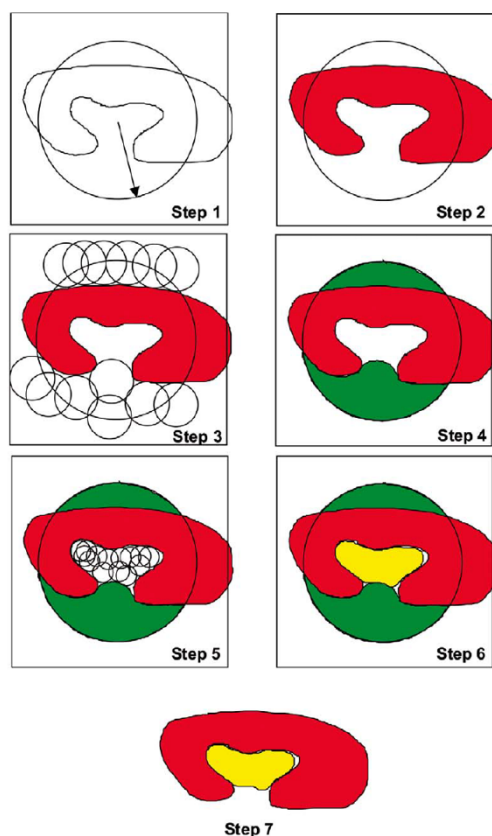
Figure 7. *Cavity detection algorithm by rbcavity of rdock [232].*

B. Molecular Lipophilicity Potential

The interaction energy of a ligand in the receptor binding site is a function of several physicochemical parameters [234]. Lipophilicity is one such guide of the various intermolecular interactions (viz. steric, hydrophobic and electrostatic) that the molecule can elicit in a receptor-bound state. The three-dimensional distribution of lipophilicity in space or on a molecular surface can be described using Molecular Lipophilicity Potential (MLP), which represents the influence of all lipophilic fragmental contributions of a molecule on its environment. The MLP value of a point in space (k) is generated as the result of intermolecular interactions between all fragments in the molecules and the solvent system, at that given point. Thus, MLP can be calculated from the fragmental system of logP [235] and a distance function as in the following equation [236]:

$$MLP_k = \sum_{i=1}^{N} F_i \cdot f(d_{ik})$$

where, N is the number of fragments, F_i is the lipophilic contribution of fragment i of the molecule and $f(d_{ik})$ is a distance function based on the distance of the measured point in space k to fragment i.

In this way, summing up all positive and all negative MLP values associated to each point on the binding pocket yields the lipophilic index (LI) as:

$$LI = \frac{\Sigma MLP^+}{\Sigma MLP^+ + |\Sigma MLP^-|} \cdot 100$$

The lipophilicity of AP in the L monomer and of DP in the T monomer were qualitatively and quantitatively estimated likewise using the MLP Tools [237] plugin available for PyMOL.

C. Electrostatic Potential

Understanding electrostatic complementarity is essential for the study of biomolecular recognition and binding processes. In particular, the long range of these electrostatic interactions makes it a vital component of molecular energetics within and between molecules [238]. The electrostatic potential maps of AP and DP allow to highlight the relative differences in the electrostatic surface exposed to the substrate in the corresponding binding pockets of Acr and Mex transporters, which can be correlated to their difference in recognition of charged substrate molecules (e.g. zwitterionic or negatively charged β-lactams; positively charged aminoglycosides).

The electrostatic potential surface maps were computed by APBS [239], after preprocessing the protein structures to assign charges and radii using PDB2PQR server [240]. All electrostatic potential calculations were performed at 0.15 M physiologic salt concentration, with a solvent probe of radius 1.4 Å, a solvent dielectric constant of 78.5, a biomolecular dielectric constant of 2.0, a temperature of 310 K, a minimum grid spacing of 0.5 Å while keeping the other Poisson-Boltzmann parameters at default.

D. Hydration Analysis

The local stereochemistry and distribution of functional groups in a region govern both the ordering of water molecules and their biologically important interactions in that region. The structure and the dynamics of the first water hydration shells around a putative binding pocket is of primary importance, given the relevance of water displacement for the free energy balance of the recognition event [241]. The Radial Distribution Function (RDF) profiles indicate the probability of finding water molecules at a certain distance from a region or residue of interest and is commonly used to analyze the solution structure revealed from either experimental or computer simulations data.

In this work, the RDF analysis of water oxygen atoms was performed using the *cpptraj* module of AMBER14, in which the RDF is computed from the histogram of the number of solvent particles found as a function of the distance R from the selection of interest, normalized by the expected number of solvent particles at that distance. The normalization is thus estimated as:

$$d \cdot \left(\left[\frac{4\pi}{3}(R + dR)^3 \right] - \left[\frac{4\pi}{3} dR^3 \right] \right)$$

where dR is equal to the bin spacing and d is set to 0.033456 water molecules per Å^3, which corresponds to a density of water of about 1.0 g mL^{-1}. A bin spacing of 0.1 and a maximum bin value of 4.0 were used in this case to calculate the RDF of all water oxygen atoms to each atom of AP in L monomer and of DP in T monomer over the entire length of the simulation.

Though RDF clearly shows a difference in the water distribution around the desired regions, it lacks the ability to present the information about the spatial positions of these differences. Hence, the Spatial Distribution Function (SDF) [242] of waters around the whole protein was calculated on the trajectory of the most populated cluster using the Gromacs utility *g_spatial* [243, 244] to determine the three-dimensional density distribution of aqueous solution around the binding pockets of these transporters. RDF and SDF together highlight the hydration level of the binding pockets in these proteins in a dynamic manner.

2.3.4. Free energy of binding: The MM/GBSA method

The binding free energy of a compound or ligand [B] to the protein or receptor [A] is a thermodynamic quantity describing the affinity between A and B, thus related to the ratio between the concentration [AB] of the complex and the product of those of the partners [A] and [B][245, 246] as shown in equation:

$$K_b = {[AB]} \Big/ {[A][B]}$$

Free energies of binding can be calculated with a plethora of methods [245, 247-250], including some computationally cheap ones characterized by an implicit treatment of solvent, namely the Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) and the MM-Poisson-Boltzmann Surface Area (MM-PBSA) methods [251-254], which have been used extensively by us in order to estimate the affinity of substrates and inhibitors to RND transporters AcrB and MexB [223, 225].
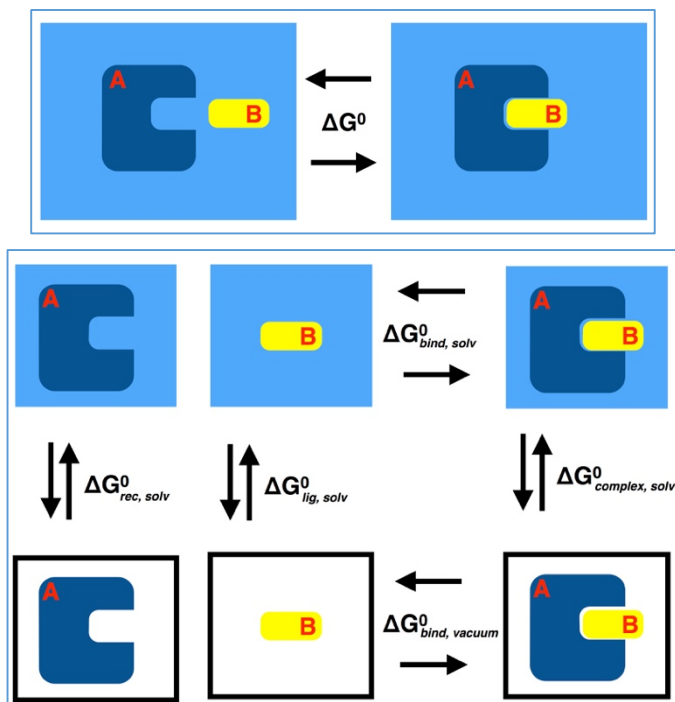
Figure 8. *(Top panel) Scheme of the ideal mechanism to calculate the binding free energy for a receptor (A) - ligand (B) and (Bottom panel) scheme of the extra cycle used in MM-G(P)BSA in which the vacuum contributions are included.*

The idea of such methods is to calculate the free energy difference between two states, which we can assume as representatives of the bound ([AB]) and unbound ([A] + [B]) states of the system, by decomposing the total free energy of binding into gas-phase energies, solvation free energies and entropic contributions [254] (Figure 8).

In the MM-G(P)BSA framework, the free energy of binding for each compound is evaluated as:

$$\Delta G_b = G_{com} - (G_{rec} + G_{lig})$$

where G_{com}, G_{rec} and G_{lig} are the absolute free energies of complex, receptor and ligand, respectively, averaged over the trajectory at equilibrium of the complex (single trajectory approach). Each term in equation above can be decomposed as:

$$\Delta G = \Delta E_{MM} + \Delta G_{solv} - T\Delta S_{conf}$$

where the free energy is the sum of the molecular mechanics energy difference ΔE_{MM}, the solvation-free energy difference ΔG_{solv} and the solute conformational entropy difference ΔS_{conf}. The term ΔE_{MM} contains the molecular mechanics energy contribution of bonded (E_{bond}, E_{angle}, $E_{torsion}$) and

non-bonded terms (E_{vdw}, E_{elec}, with no cutoff) estimated from the force field and, thus, can be calculated as:

$$\Delta E_{MM} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{torsion} + \Delta E_{vdw} + \Delta E_{elec}$$

The second term (ΔG_{solv}) which is the solvation-free energy is modeled as sum of an electrostatic [evaluated with MM-G(P)BSA approach] and a non-polar contribution (proportional to the difference in the solvent-exposed surface area).

$$\Delta G_{solv} = \Delta G_{solv,p} + \Delta G_{solv,np}$$

The last term in $T\Delta S_{conf}$ is the solute entropy contribution, which is composed by two terms: the rototranslational contribution, calculated through classical statistical mechanics, and the vibrational term, which can be estimated through normal-mode analysis [255].

The MM-GBSA approach also provides an alternative to the alanine scanning approach [256] by means of a per-residue decomposition of the contributions to the binding free energy. In this case, the energy contribution of a single residue is calculated by summing its interactions over all residues in the system.

2.4. Comments

Part of this chapter has been published as a book chapter:

Cacciotto, P., **Ramaswamy, V. K.**, Malloci, G., Ruggerone, P., & Vargiu, A. V., "*Molecular modeling of multi-drug properties of RND transporters*." Bacterial Multidrug Exporters of the series Methods in Molecular Biology, Ed. Yamaguchi A. (Springer Verlag, Heidelberg, New York, 2017) *(In Press)*

Chapter 3

Molecular Rationale Behind the Differential Substrate Specificity of RND Transporters AcrB And AcrD

3.1 Introduction

AcrB and AcrD are the two major RND transporters in *Escherichia coli*, and in spite of sharing an overall sequence identity of nearly 66% (similarity of nearly 80%), they feature distinct substrate specificity patterns. As listed in Table 1, for example, macrolides and compounds of tetracycline family are transported by AcrB but not by AcrD while AcrD manages to export aminoglycosides. They also share certain common substrates viz. most of the beta-lactam class of antibiotics. These substrates when categorized based on their physicochemical properties (Table 2), simplifies the specificity pattern highlighting the fact that substrates of AcrB and AcrD are essentially hydrophobic and hydrophilic, respectively, while both transporters might shuttle out amphiphilic compounds.

Table 1. Substrate specificities of homologous RND transporters AcrB and AcrD in E. coli [108, 112, 114, 118, 130, 257-263]. Examples of compounds are mentioned within parentheses.

Transporter(s)	AcrB	AcrD	AcrB and AcrD
Substrates	Macrolides (erythromycin), chloramphenicol, ciprofloxacin, Tetracyclines (tetracycline, tigecycline, minocycline), doxorubicin, acriflavine	Aminoglycosides (amikacin, gentamicin, kanamycin, neomycin)	Most beta-lactams (aztreonam, sulbenicillin), Aminocoumarins (novobiocin)

Table 2. General physicochemical properties of substrates of AcrB and AcrD

Substrate Type (Physicochemical property)	*E. coli*	
	AcrB	AcrD
Hydrophobic	+	-
Hydrophilic	-	+
Amphiphilic	+	+

One of the most likely explanations for this difference in substrate specificities between the homologous transporters would be to link them with possible dissimilarities in the structural and chemical features within their putative substrate-binding pockets. However, very little is known about the nature of the binding sites in these transporters. Two previous studies attempted to identify substrate recognition site(s) in these Acr pumps by using chimeric analysis [108, 118]. The importance of periplasmic loop regions was pointed out by Elkins and Nikaido [118], and successively deepened by Kobayashi *et al.* [108]. The latter authors, in particular, identified a few residues in the AP as potential determinants of specificity towards negatively charged β-lactams (aztreonam, carbenicillin, and sulbenicillin) in these Acr transporters. The three residue replacements in AcrB by those in AcrD (Q569R, I626R, and E673G) in their study, were all accompanied with a gain in positive charge and loss of negative charge thereby favoring recognition of anionic β-lactams in AcrB. However, these findings concerning the overall location of the sites responsible for substrate recognition were restricted to a subclass of compounds, and no comprehensive molecular-level rationale for the different specificities of AcrB and AcrD has been proposed yet. An extensive comparison of the structural and physicochemical properties of binding pockets of the two transporters could explain, in a more general and broader sense, their substrate specificities in terms of complementarity with the corresponding (substrate) binders.

Under this aim, we report here the first exhaustive study comparing in a systematic way the physicochemical nature of the main putative substrate binding sites (AP and DP) (Structural aspects of RND transporters have been discussed in Chapter 1) between AcrB and AcrD. The lack of any experimental information on the structure of AcrD has been a major hindrance for accomplishing such comparative studies. Nevertheless, given the high sequence identity and similarity between these Acr pumps of *E. coli*, computational modeling of AcrD and related structure-based studies are possible. An important point of our study regards the inclusion of the dynamical aspects of the two transporters in the analysis, extending the realm of structure-function relation to account for subtle interplay between behavior of the solvent, exposure and/or burial of charge distribution, structural changes associated with the time evolution of the system under physiological-like conditions. Thus, we characterize and compare the details of their molecular properties (pocket descriptors) like accessible binding volume, lipophilic index, electrostatic potential, and hydration not only on crystal structures or homology models but considering also the impact of the dynamics on all the above properties on the basis of extensive molecular dynamics simulations on both transporters.

Our study would be informative to new drug design attempts to correlate the different specificity patterns of these two transporters to the physicochemical as well as topographical

properties calculated on (or projected onto) the molecular surface of their multifunctional recognition sites [264].

3.2. Methods

3.2.1. Homology modeling of AcrD

A reliable structure of the system of interest is the starting and main ingredient of any structure-based computational study. Since the structure of AcrD has not yet been resolved experimentally, we built it by template-based homology modeling. The amino acid sequence of full length AcrD transporter protein from *E. coli* was retrieved from the UniProt database [265] (UNIPROT ID: P24177), and subsequently searched for the best available template structures bearing homologous relationship to the query sequence using the NCBI-BLAST tool [161] against the Protein Data Bank (PDB) (www.rcsb.org) [266]. AcrB sequence showed the highest identity (~66%) (and similarity of ~80%) with least gaps over a maximum sequence coverage; therefore its high resolution crystal structure, 1.9 Å (PDB ID: 4DX5 [111]) was chosen as template for modeling AcrD. The two protein sequences were optimally aligned using ClustalOmega [162] and the results were visually inspected to ensure the absence of gaps in important secondary structure regions. Modeller 9.13 [152] was used to generate a total of 100 asymmetric models of AcrD based on AcrB template using an optimization method combining slow MD with very thorough variable target function method through 300 iterations, and this whole cycle was repeated twice unless the objective function MOLPDF was greater than 10^6. The resulting models were ranked using DOPE [267] score values, and the top 5 models (with the lowest DOPE score) were selected for individual structure quality checks. Each model was further subjected to loop refinement using Modeller, and to overall structure relaxation by energy minimizations using AMBER14 [191]. The most reliable model was then selected based on various geometric and stereochemical quality [268] factors calculated for backbone angles, side chains flips, rotamers, steric clashes etc. using PROCHECK [169], ERRAT [269], ProSA [165], Verify3D [167] programs available in MolProbity [270] and Structure Analysis and Verification Server tools (SAVES) (http://services.mbi.ucla.edu/SAVES/).

We also performed comparative structural studies by superimposition of the modeled AcrD structure over the experimentally determined X-ray crystal structure of AcrB used as the template. All the above methods were validated by their application to the crystal structure of AcrB. Visual inspections were performed with VMD1.9.1 [271] and PyMOL [272].

3.2.2. Molecular dynamics simulations of AcrB and AcrD

MD simulations of the crystal structure of AcrB (PDB ID: 4DX5) and of the homology model of AcrD were carried out using the AMBER14 molecular modeling software [191]. Monomer specific protonation states [127] were adopted with E346 (E346) and D924 (D922) protonated in both L and T monomers while deprotonated in O monomer of AcrB (AcrD). The residues D407 (D407), D408 (D408), D566 (L565) were protonated only in O monomer of AcrB (AcrD). The topology and the initial coordinate files for these apo-protein structures were created using the *LEaP* module of AmberTools14. The proteins were successively embedded in POPE bilayer patches, solvated with explicit TIP3P water model, and neutralized with the required number of randomly placed K^+ ions [208-211]. The ions count was suitably adjusted to account for an osmolarity of 0.15M KCl. Embedding of the protein into a pre-equilibrated POPE bilayer patch was performed using the PPM server [212] and subsequently the CharmmGUI tool [213]. The lipid residue nomenclature was converted from the CHARMM to AMBER format using the *charmmlipid2amber.py* python script provided with AmberTools. The central pore lipids were then added after calculating the number of lipids to be added to each leaflet by dividing the approximate area of the central pore by the standard area per lipid of POPE molecules [192]. Periodic boundary conditions were used and the distance between the protein and the edge of the box was set to be at least 15 Å in each direction.

Multi-step energy minimization with a combination of steepest descent and conjugate gradient methods was carried out using the *pmemd* program implemented in AMBER14 to relax internal constrains of the systems by gradually releasing positional restraints. Following this, the systems were heated from 0 to 310 K by a 1 ns heating (0-100 K) under constant volume (NVT) followed by 5 ns of constant pressure heating (NPT) (100-310 K) with the phosphorous heads of lipids restrained along the Z-axis to allow membrane merging and to bring the atmospheric pressure of the system to 1 bar. A Langevin thermostat [215] (collision frequency of 1 ps^{-1}) was used to maintain a constant temperature, and multiple short equilibration steps of 500 ps under anisotropic pressure scaling (Berendsen barostat) in NPT conditions were performed to equilibrate the box dimensions. A time step of 2 fs was used during all these runs, while post-equilibrium MD simulations were carried out with a time step of 4 fs under constant volume conditions after hydrogen mass repartitioning [216]. The particle-mesh Ewald (PME) algorithm [203] was used to evaluate long-range electrostatic forces with a non-bonded cutoff of 9 Å. During the MD simulations, the length of all R–H bonds was constrained with SHAKE algorithm [186]. Coordinates were saved every 100 ps. The ff14SB [214] version of the all-atom Amber force field was used to represent the protein systems while lipid14 [192] parameters were used for the POPE bilayer. Multi-copy µs-long MD simulations for each system, namely, two 1.5 µs-long simulations

for each transporter for a total simulation time of 6 μs, were performed for better statistical reliability in the obtained data. Graphs were generated using the xmgrace [218] plotting tool. Trajectory analysis was done using *cpptraj* module [217] of AmberTools14 and VMD1.9.1.

3.2.3. Clustering of MD trajectories

A cluster analysis of the MD trajectories was performed using the average-linkage hierarchical agglomerative clustering method [230] implemented in *cpptraj* module of AMBER. Such clustering helps to reduce the number of structures for analysis yet retaining the large conformational space sampled during the MD runs. In this approach, we clustered the trajectory independently based on RMSD (cutoff set to 3 Å) of the AP in *L* monomer and the DP in *T* monomer for both the proteins. The representative structure from each of the 10 top clusters generated in each of the four cases considered (AP in *L*, DP in *T* for both proteins) was taken for the successive quantitative analyses in order to account for the dynamical behavior of the proteins. All non-protein molecules were stripped from the trajectory during post-processing to reduce additional memory usage and to speed up file processing.

3.2.4. Pocket descriptors

The list of the pocket descriptors identified for the present study contains:

1. Cavity volume;
2. Molecular lipophilicity potential;
3. Electrostatic potential;
4. Hydration.

The various pocket descriptors used to characterize the binding site [273] were calculated using specific programs after validating their applicability to RND systems by assessing results against available crystal structures and experimental data.

3.2.4.1 Cavity Volume

Evolution of size and shape of the AP and DP during the MD simulations was examined using the two-probe sphere method of *rbcavity* program bundled in the rDock suite [231]. This allows obtaining detailed information on the pocket volume and plasticity of the site. In this method, the potential binding site volume was identified by a fast grid-based cavity detection algorithm [232] within a sphere of radius 14 Å, centered over the pockets, using large and small probe radii of 6.0 Å and 1.5 Å, respectively. These radii were identified as optimal after testing several combinations so as to correctly estimate the pocket volume by keeping the possible inclusion of non-pocket region at its least.

3.2.4.2. Molecular Lipophilicity Potential

Lipophilicity is one of the indicators of the various intermolecular interactions [234] (viz. steric, hydrophobic and electrostatic) that the molecule can elicit in a receptor-bound state. The three-dimensional distribution of lipophilicity in space or on a molecular surface can be described using Molecular Lipophilicity Potential (MLP), which represents the influence of all lipophilic fragmental contributions of a molecule on its environment. The lipophilicity of AP in *L* monomer and DP in *T* monomer were qualitatively and quantitatively estimated in this way using MLP Tools [237] plugin available for PyMOL.

3.2.4.3. Electrostatic Potential

Understanding electrostatic complementarity is essential for the study of biomolecular recognition and binding processes. In particular, the long range of these electrostatic interactions makes it a vital component of molecular energetics within and between molecules [238].

The electrostatic potential surface maps were computed by APBS [239], after preprocessing structures of AcrB and AcrD to assign charges and radii using PDB2PQR server [240]. All electrostatic potential calculations were performed at 0.15 M physiologic salt concentration, with a solvent probe of radius 1.4 Å, a solvent dielectric constant of 78.5, a biomolecular dielectric constant of 2.0, a temperature of 310 K, a minimum grid spacing of 0.5 Å and keeping the other Poisson-Boltzmann parameters at default.

3.2.4.4. Hydration analysis

The radial distribution function (RDF) profiles indicate the probability of finding water molecules at a certain distance from a region or residue of interest and is commonly used to analyze the solution structure revealed from either experimental or computer simulations data. The RDF analysis of water oxygen atoms was performed using *cpptraj* module of AMBER14, in which the RDF of all water oxygen atoms to each atom of AP in *L* monomer and of DP in *T* monomer were computed over the entire length of the simulation.

Though RDF clearly shows a difference in the water distribution around the desired regions, it lacks the ability to present the information about the spatial positions of these differences. Hence, spatial distribution function (SDF) [242] of waters around the whole protein was calculated on the configurations extracted from MD simulations and forming the most populated cluster using the Gromacs utility *g_spatial* [274] to determine the three-dimensional density distribution of aqueous solution around the binding pockets of these transporters. RDF and SDF together highlight the hydration around the binding pockets of these proteins, which can be effectively used to understand the molecular mechanism of interaction of water molecules penetrating the pocket in a dynamic manner.

3.3. Results

3.3.1. Homology modeling of AcrD
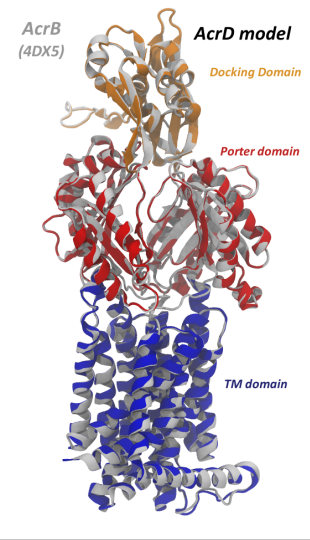
RND transporters share a common overall fold but subtle differences exist in the orientation of loop regions and in certain regions with non-identical residues, especially in the AP and DP, which might contribute towards the differential substrate specificities of these proteins. Comparing AcrB to AcrD (sequence alignment given as Figure 1 in Appendix), we noticed that the AP was better conserved than the DP with nearly 60% identical residues in contrast to around 40% in the DP. Most of the conserved residues were identified to play an important role in forming polar, van der Waals and ring stacking interactions with substrates and inhibitors based on mutation and interaction studies [108, 109, 111, 116, 118, 211, 275, 276].

Visual inspection of the top 5 models generated by MODELLER revealed an identical overall 3-dimensional structural fold with minor difference in the loop regions. The final AcrD model featured 97.3% and 99.7% of residues in the favored and allowed regions of the Ramachandran plot, respectively. The evaluation results were fully within the permissible limits of a good model and none of the outlier residues were in the region of interest. The overall quality factor for non-bonded atomic interactions assessed with ERRAT plot further confirmed the high quality of our model with a score of 97.21%, which is in the range of high-resolution experimental structures. The ProSA evaluation showed that the overall model quality (Z-score) of the homology model (-12.18) and the template crystal structure (-12.62) were quite similar and within the range of scores typically found for native proteins of similar size. The local quality of the model based on interaction energies for each residue as evaluated with ProSA served as an additional evidence of model quality with most of the residues falling in the negative energy scale and only a minor set on the positive side. All the observed values were within ranges observed for experimentally determined structures, thereby increasing the confidence of acceptability of the built homology model.

The evaluation results are summarized in Table 3, which also shows the superposed structures of the AcrD model with its AcrB template, highlighting the similarity in their general structural fold.

Table 3. *Evaluation results of the AcrB-based AcrD model. The figure on the left shows the structural superposition of AcrD (colored representation) onto AcrB (gray representation). Only a single monomer is shown for the sake of clarity. The table on the right shows the model evaluation results.*

Evaluation Criteria	AcrB-based AcrD model
RMSD (backbone)	0.12 Å
TM-score	0.99
Ramachandran favored	97.33%
Errat	97.21%
Verify-3D	Pass

3.3.2. Molecular dynamics simulations of AcrB and AcrD

The stability of the AcrD model as well as its suitability for subsequent analysis was validated in two independent μs-long MD simulations. According to monomer-wise RMSD analyses of the protein backbone of different domains and of the whole protein with reference to the initial structure, we considered 0.5 μs a suitable equilibration time. The cluster representatives from the equilibrated trajectories of AcrB and AcrD were used to characterize the distribution of accessible binding volume, molecular lipophilicity and electrostatic potential, while the MD trajectories were considered for the analysis of hydration patterns within the AP and DP, so as to identify possible differentiating factors between them. In the following, we present the results from these characterizations on the two aforementioned sites.

3.3.2.1. Access Pocket

The access pocket was recently identified to be a probable "substrate selection filter" for anionic β-lactams based on replacement of selected residues in AcrB with their AcrD counterparts [108]. In order to identify if any physicochemical property of the pocket could differentiate the AP of these proteins and determine the probable nature of their substrates, we calculated the following descriptors.

A. Pocket Volume

The extent and the shape of the three dimensional space that a ligand is allowed to explore to find its optimal binding pose in any pocket is governed by multiple factors, the primary of which is the accessible binding volume (to allow the ligand to enter) prior to and/or in addition to factors like

shape and electrostatic complementarity [231]. Especially with promiscuous proteins like the RND transporters, one would expect a large binding site with reasonable degree of plasticity to facilitate binding of molecules of a wide range of sizes.
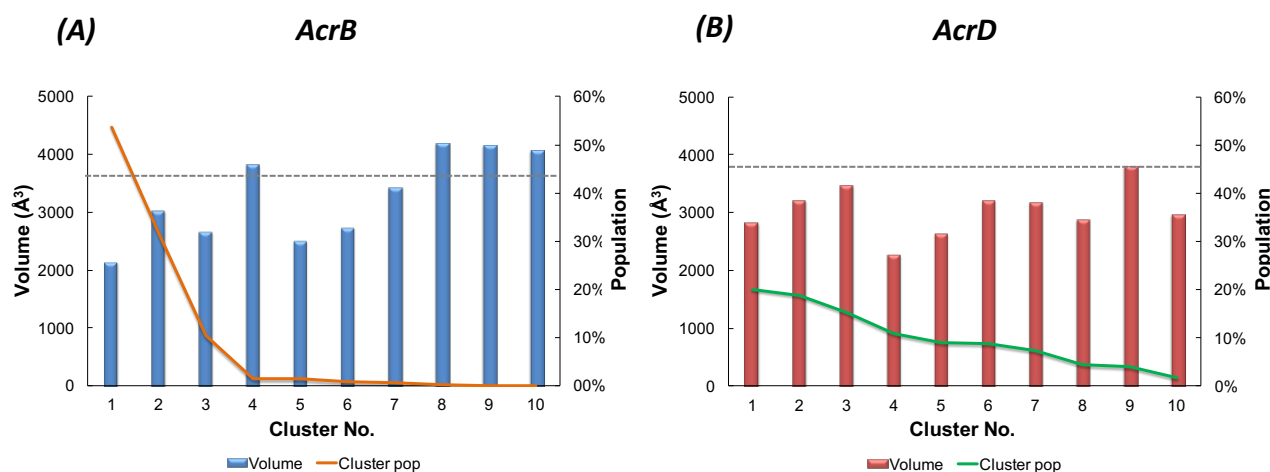
Figure 1. *Volume distribution of AP in the L monomer of (A) AcrB and (B) AcrD over the simulation timescale. Histograms refer to the volume distributions, lines to the population of the clusters. The corresponding volume in their Pre-MD structures are shown as dashed line.*

The distribution of the volumes of AP in *L* monomer with respect to their cluster population mirrored the relative size of the substrates transported by these pumps as well as their different dynamicity. The AP of AcrB featured an average volume of 2515 Å^3 ($\pm$438 Å^3) with the three most populated clusters (54%, 31%, 10% of the population, respectively) peaked around 2140, 3020, and 2660 Å^3, respectively (Figure 1A). Note that the pocket in the AcrB structure used as initial structure for MD simulations, i.e., the crystal structure identified by PDB code 4DX5 [111] (hereafter called Pre-MD) is 3620 Å^3 while it is 3820$\pm$180 Å^3 when a set of 8 crystal structures was considered. In the case of AcrD, the average value was 3015 Å^3 ($\pm$385 Å^3). The major clusters sampled volumes of 2820 Å^3 (20% of simulation), 3200 Å^3 (19%), 3460 Å^3 (15%) and 2260 Å^3 (10%) (Figure 1B). In the Pre-MD structure (for AcrD, pre-MD structure refers to the final optimized model used as starting configuration for MD simulations), the volume was 3760 Å^3. Also, note that for AcrD the cluster distribution was more extended than for AcrB: the fourth cluster in AcrB contained only 1.5% of the population while clusters 5, 6, 7 of AcrD still covered 9%, 8.7%, and 7.3%.

B. Molecular Lipophilicity Potential

The results for the lipophilic index of AP of both transporters proved the higher lipophilic nature of AP in AcrB when compared to that in AcrD (Table 4), which is required for AcrB to provide a favorable environment for its hydrophobic substrates to bind. The maximum LI for AcrB, 9.5, was associated with the fifth cluster covering 1.4% of the population whereas in AcrD a cluster with 7.3% of the population had the highest LI, 2.8. However, the specific chemical environment of the access pocket is neither entirely hydrophobic nor entirely polar in both the proteins (Figure 2A). Interestingly, the dynamics affected only slightly the lipophilic index for AcrB (7.2 vs. 7.0 for pre-MD and average value, respectively) and induced a small increase in magnitude for AcrD (1.2 vs. 1.6). According to the molecular lipophilic potential in the representatives of the most populated clusters extracted from MD, regions of particular lipophilicity were located for AcrB in the side part of the pocket and at the border with the vestibule (Figure 2B, top panel) while no predominant spots were recognizable for AcrD (Figure 2B, bottom panel).

Table 4. *Lipophilic index of AP in the L monomer of AcrB and AcrD. For AcrB, the Pre-MD structure corresponds to the crystal structure identified by PDB code 4DX5 [111] while for AcrD it is the final optimized model used as starting configuration for MD simulations.*

System	Lipophilic index	
	Pre-MD	From MD
AcrB	7.2	7.0 ± 1.0
AcrD	1.2	1.6 ± 0.6

(A)

(B)
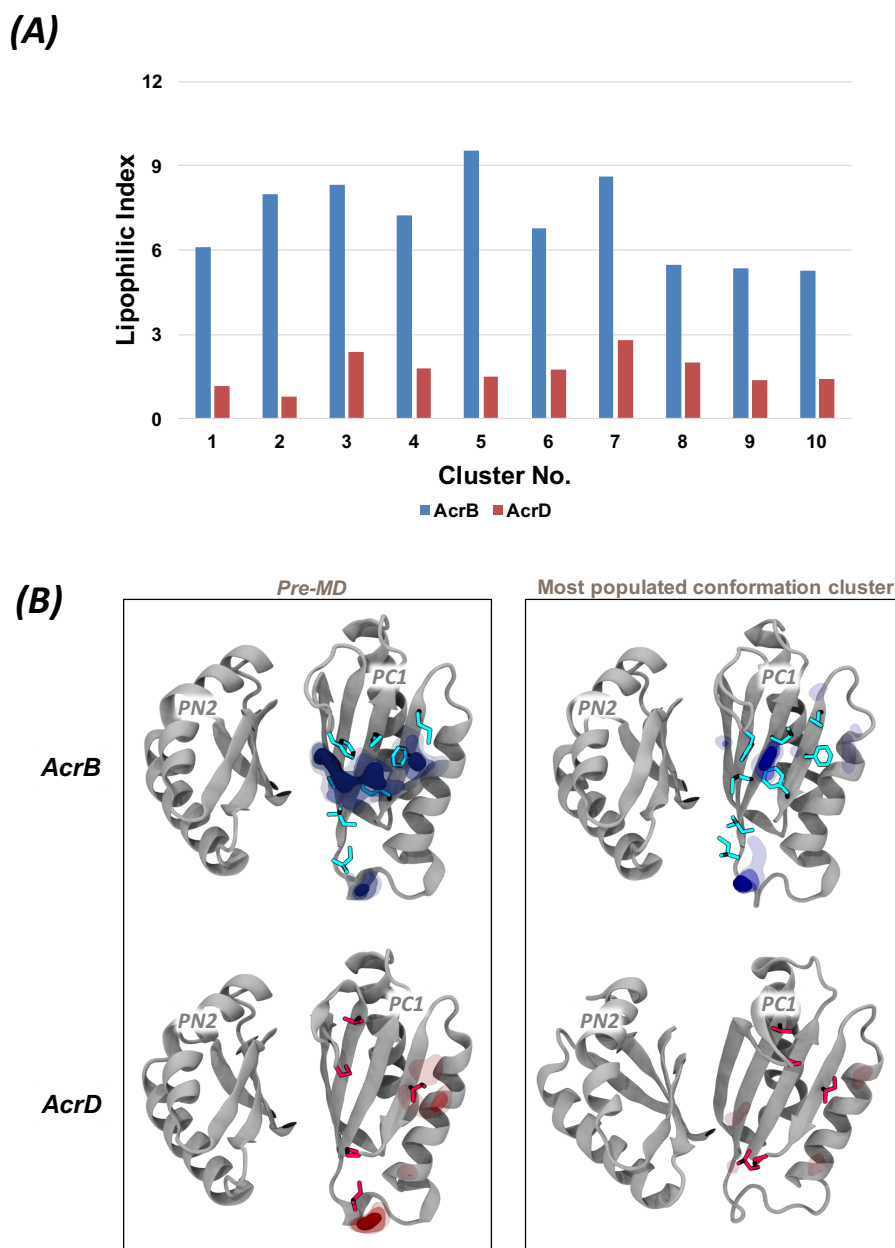
Figure 2. *(A) Distribution of the lipophilic index for AP in the L monomer of AcrB and AcrD over the clusters. (B) Molecular lipophilic potential surface observed within 4 Å of AP in the L monomer of AcrB (blue) and AcrD (red) in Pre-MD (left panel) and the representatives of the most populated cluster (right panel) as seen from PC2-PN1 side. The hydrophobic/aromatic residues of the pocket are shown as sticks in the respective structures. Isosurfaces at 0.75 (solid), 0.5 and 0.25 (transparent) are shown.*

C. Electrostatic Potential

The electrostatic potential of the pocket in AcrB and AcrD are shown in Figure 3. Panel A collects the results for the Pre-MD structures, and Panel B those of the most populated clusters. Positively charged patches (blue in Figure 3B) were predominant in AcrD and essentially compacted by the dynamics to form a large region of positive potential. A similar effect of the dynamics appeared in AcrB but the distribution of positive and negative regions appeared unaffected by the dynamics.
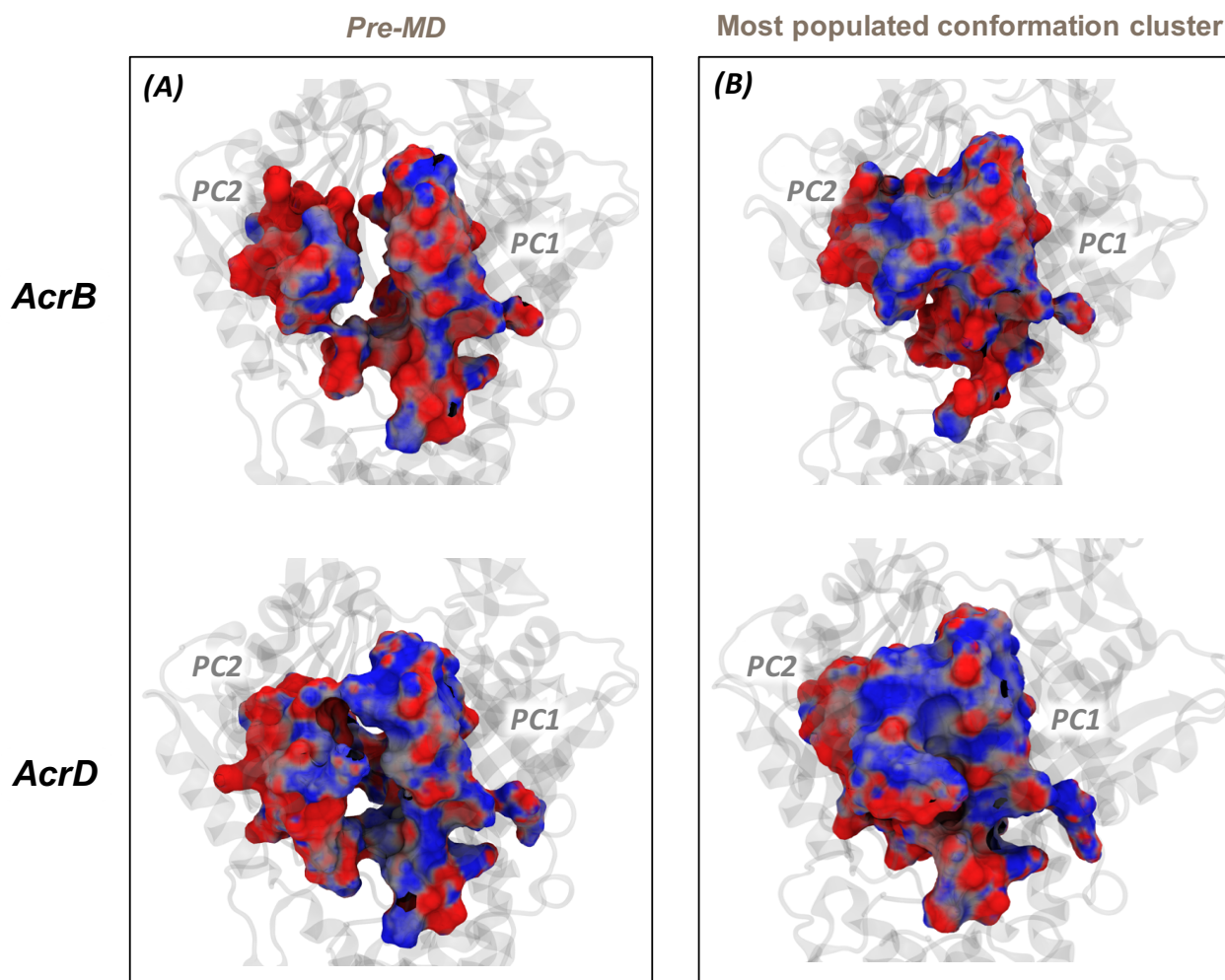
Figure 3. *Molecular surface representation of AP in the L monomer of AcrB and AcrD colored according to the electrostatic potential [red to blue from negative (-10 k_bT/e_c) to positive (+10 k_bT/e_c) potential]. The upper and lower images refer to AcrB and to AcrD, respectively. The left panel displays the potential of the regions before MD simulations, the right panel contains the potential of the most populated clusters extracted from the MD trajectories.*

D. Hydration Analysis

The RDF profile around the access pocket residues of AcrB and AcrD calculated over the entire trajectory were rather similar with only a minor difference in the intensity of hydration (Figure 4A). The first solvation shell was observed around 1.9 Å in both the proteins with a slightly reduced probability in AcrB. The SDF calculated on the configurations extracted from MD simulations and forming the most populated cluster, however, featured no water density spots near the hydrophobic residues in AP of AcrB but showed many dense regions in AcrD at identical density isovalues (Figure 4B).
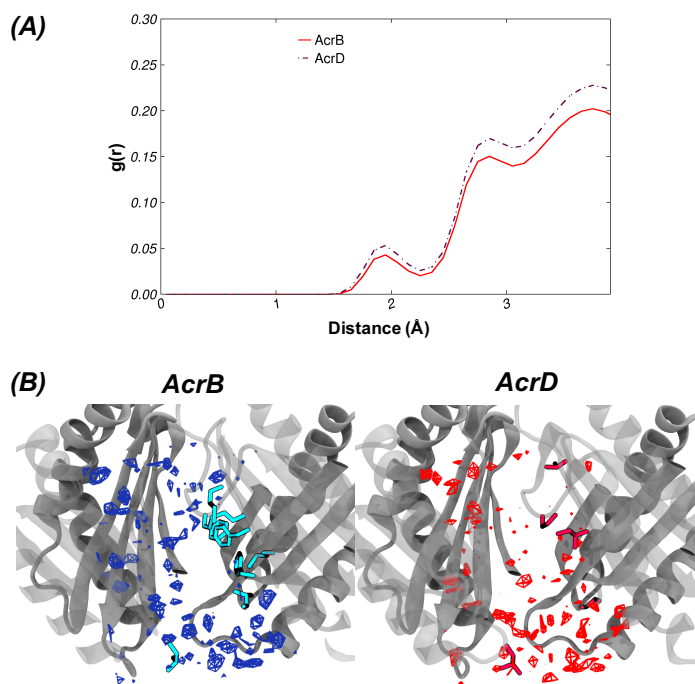
Figure 4. *(A) Comparison of RDF profiles of water oxygen atoms around AP in the L monomer of AcrB (red solid line) and AcrD (brown dash-dotted line) extracted from the whole trajectories. (B) Comparison of SDF for waters in the AP of L monomer calculated on the configurations belonging to the most populated cluster of (left panel) AcrB and (right panel) AcrD illustrating the variation in the immediate environment of the hydrophobic residues. The isosurfaces are shown at density isovalue of 6.*

3.3.2.2. Deep Pocket

The DP is a more deeply located cavity within the substrate transport pathway in RND pumps, and is likely the recognition site for low molecular mass compounds [109]. According to the crystal structures, this pocket exists in a collapsed state in the *L* and *O* monomers but is wide open in the *T* monomer; therefore, all the analyses concerning this site were performed on the *T* monomers of AcrB and AcrD.

We noticed most of the hydrophobic and bulky residues in the DP of AcrB replaced by polar/charged and less voluminous amino acids in AcrD. Such substitutions involving a large number of polar charged residues would definitely result in a major variation of the chemical environment of the pocket space.

A. Pocket Volume

In AcrB, this pocket featured volumes between 2650 and 2500 Å^3 and around 2900 Å^3 for about 44%, 23%, and 13% of the simulation time (Figure 5A). The average volume considering the ten clusters identified in the analysis was 2610 Å^3 ($\pm$250 Å^3) in comparison with the average value of 4040 Å^3 ($\pm$70 Å^3) evaluated over a set of 10 crystal structures and of 3708 Å^3 calculated for the pre-

53

MD structure. In the case of AcrD, the DP sampled volumes around 2500 and 3100 Å^3 for about 30% and 20% of the total simulation time, respectively (Figure 5B). The mean value of the volume was 2770 Å^3 (±306 Å^3). The DP in the Pre-MD structure had a volume of 3855 Å^3. As in the case of AP, the cluster distribution was more extended for AcrD than for AcrB.

(A)   AcrB   **(B)   AcrD**
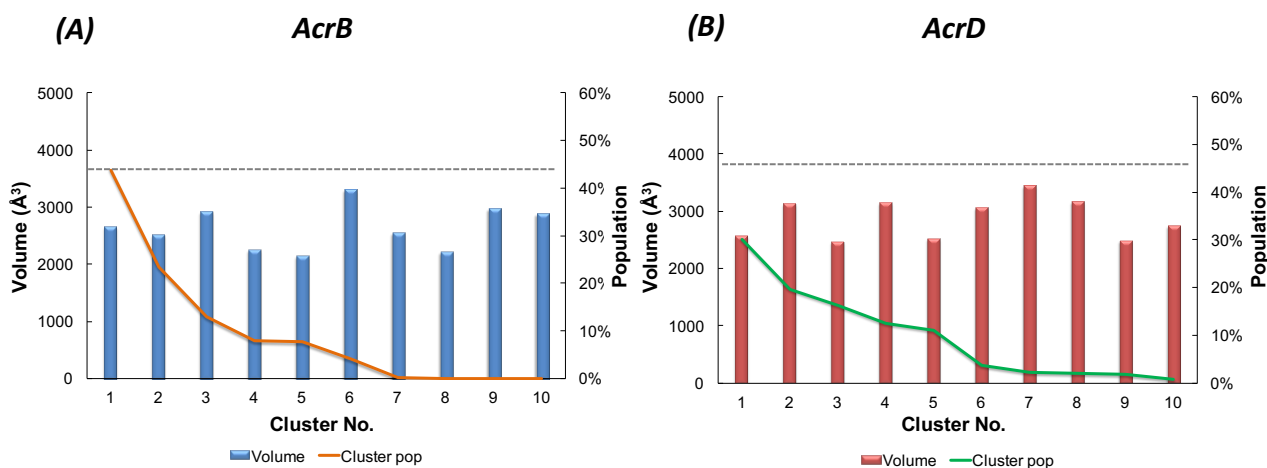
Figure 5. Volume distribution of DP in the T monomer of (A) AcrB and (B) AcrD over the simulation timescale. Histograms refer to the volume distributions, lines to the population of the clusters. The corresponding volume in their Pre-MD structures are shown as dashed line.

B.   Molecular Lipophilicity Potential

We observed a remarkably stronger reduction of the LI in AcrD than in AcrB, as can be seen from the values reported in Table 5. The dynamics reduced the LIs by 60% in AcrD and 20% in AcrB from their Pre-MD values. The maximum LI for AcrB, 17.2, was associated with cluster 8 covering 0.1% of the population but high lipophilic indices characterized clusters 1, 3, and 5 (LI of 14.4, 17.1, 17.0, respectively) (Figure 6A). Together, the three clusters embraced 64% of the simulation time. In AcrD, a cluster with 1.8 % of the population had the highest LI, 5.6 while the most populated clusters had LIs ranging from 0.5 to 2.8 (Figure 6A). The lipophilic potential surfaces of the Pre-MD and the most populated clusters are reported in Figure 6B. Pronounced lipophilic regions appeared for AcrB while for AcrD three less extended spots were present.

It is apparent from the MLP calculations discussed above that the difference in the LI between AcrB and AcrD is greater in the lesser-conserved DP than that in the AP.

Table 5. Lipophilic index of DP in the T monomer of AcrB and AcrD. For AcrB, the Pre-MD structure corresponds to the crystal structure identified by PDB code 4DX5 [111] while for AcrD it is the the final optimized model used as starting configuration for MD simulations.

System	Lipophilic index	
	Pre-MD	From MD
AcrB	16.3	13.1 ± 3.3
AcrD	4.9	1.9 ± 1.0

(A)
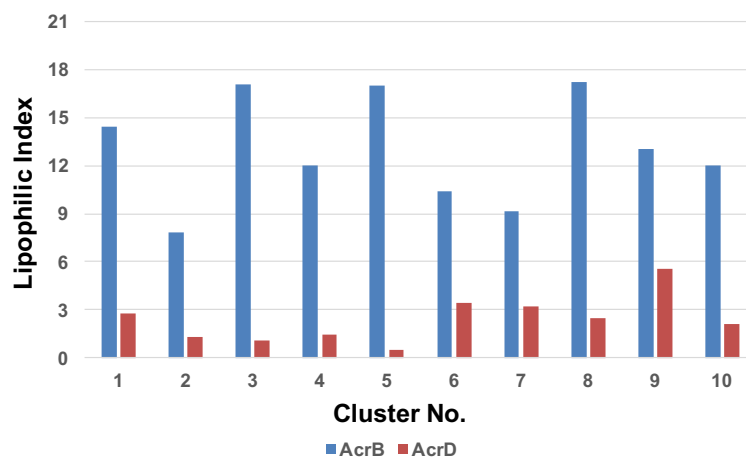
(B)
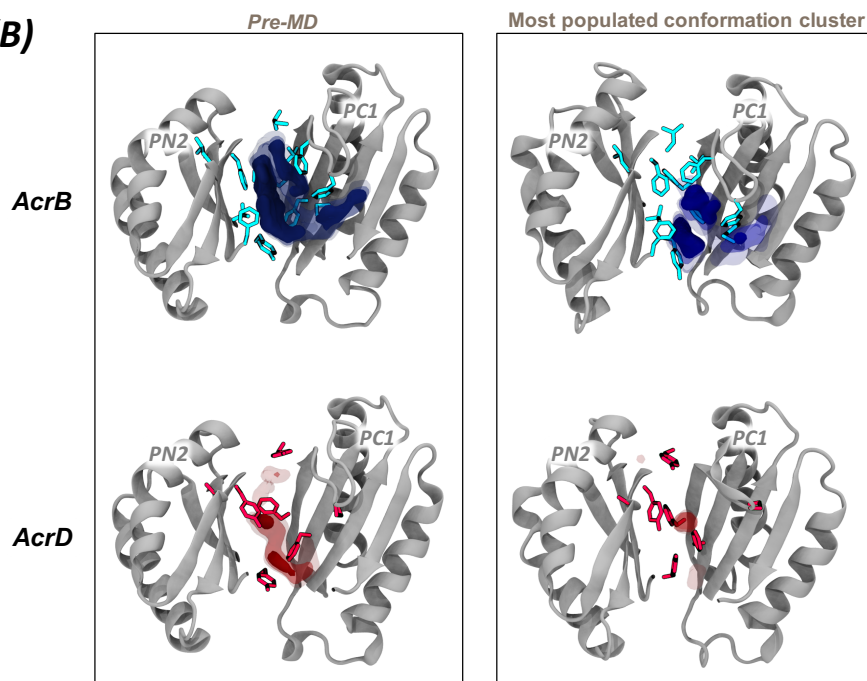
Figure 6. (A) Distribution of the lipophilic index for DP in T monomer of AcrB and AcrD over the clusters. (B) Molecular lipophilic potential surface observed within 4 Å of DP in the T monomer of AcrB (blue) and AcrD (red) in Pre-MD (left panel) and the representative of the most populated cluster (right panel) as seen from PC2-PN1 side. The hydrophobic/aromatic residues of the pocket are shown as sticks in the respective structures. Isosurfaces at 0.75 (solid), 0.5 and 0.25 (transparent) are shown

C. Electrostatic Potential

The electrostatic potential surface of the DP in AcrD and AcrB indicated a relatively dense positive environment in AcrD compared to AcrB (blue spots in Figures 7). Noticeable is that the difference was even emphasized when the electrostatic potential surfaces were compared for the representatives of the most populated clusters: in AcrB an extended surface area of negative potential appeared while the distribution of spots of negative and positive potentials did not change for AcrD with a more pronounced positive component.
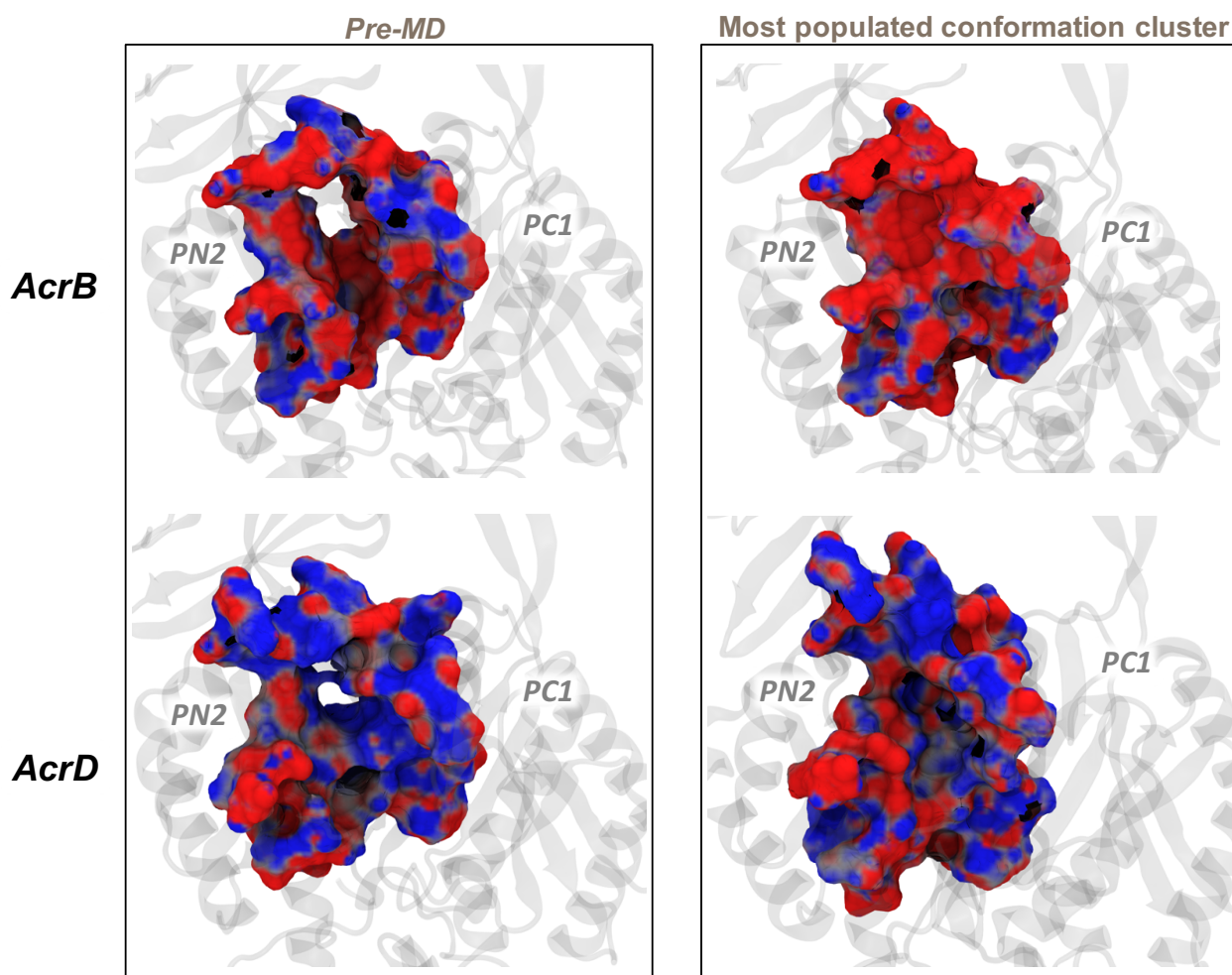
Figure 7. *Molecular surface representation of DP in the T monomer of AcrB and AcrD colored according to the electrostatic potential [red to blue from negative (-10 k_bT/e_c) to positive (+10 k_bT/e_c) potential]. The upper and lower images refer to AcrB and to AcrD, respectively. The left panel displays the potential of the regions before MD simulations, the right panel contains the potential of the most populated clusters extracted from the MD trajectories.*

D. Hydration Analysis

The RDF profile around the DP residues showed a minor difference in the intensity of the peak between AcrB and AcrD but a closer inspection revealed this difference to be essentially related to the difference in the so-called hydrophobic trap (HP-trap) region (Figure 8A) [277]. The HP-trap is lined by five hydrophobic phenylalanine residues in AcrB while the corresponding site in AcrD has only 2 of these phenylalanine residues (corresponding to F610 and F627 in AcrB) conserved. The RDF around the HP-trap residues showed a much greater difference in the intensity and distance of first solvation shell than that around the DP in *T* monomer. In the case of AcrB, a small peak corresponding to its first solvation shell around the HP-trap residues was observed close to 2 Å but is greatly suppressed showing the very low probability of hydration near this region as seen from the SDF in Figure 8B. On the other hand, this first peak was sharp and distinct at around 1.9 Å in the case of AcrD due to the presence of polar residues (N136, N178, S614).
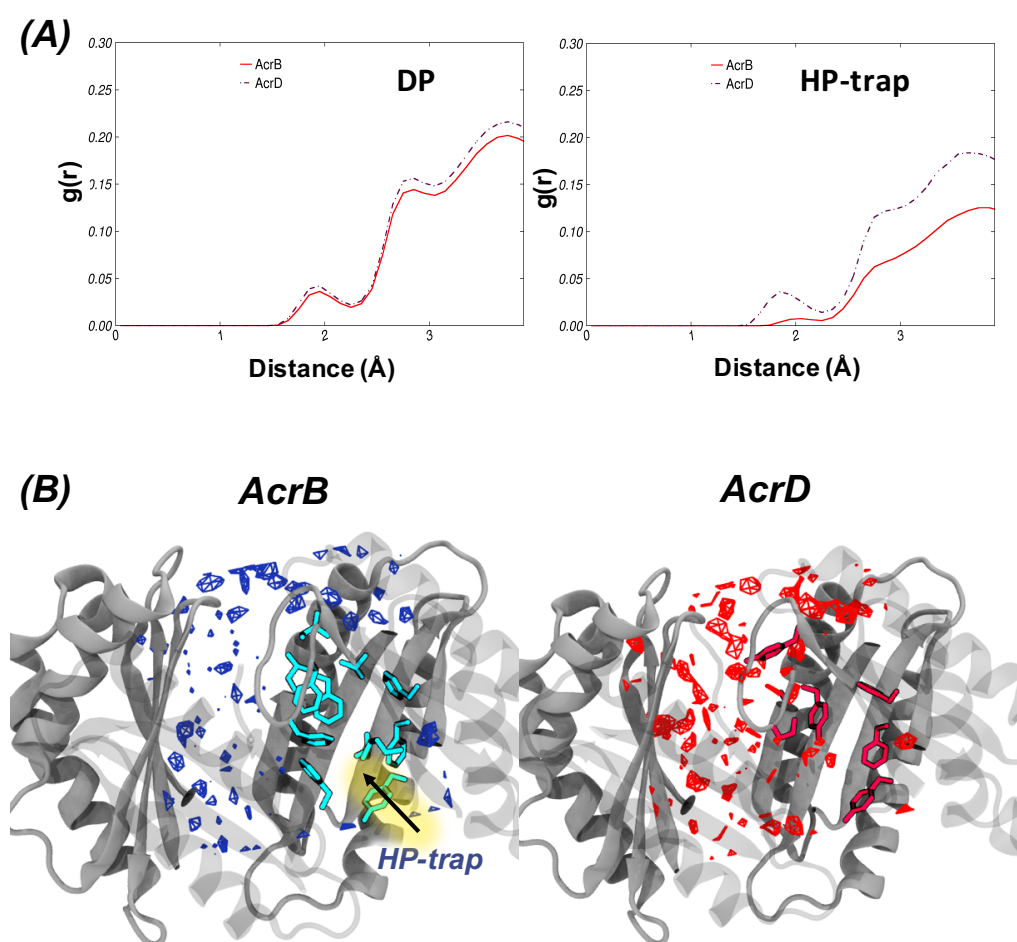
Figure 8. (A) Comparison of RDF profiles of water oxygen atoms around the DP and HP-trap of monomer T in AcrB (red solid line) and the corresponding regions in AcrD (brown dash-dotted line). (B) Comparison of SDF for waters in the DP of T monomer calculated on the configurations forming the most populated cluster of (left panel) AcrB and (right panel) AcrD illustrating the variation in the immediate environment of the hydrophobic residues. The isosurfaces are shown at density isovalue of 6. The position of the hydrophobic trap (HP-trap) in DP of AcrB is highlighted.

3.4. Discussion

A closer inspection of the mismatched residues underlined an evolutionary inclination towards populating the binding pockets of AcrD with more polar/charged and less voluminous residues than that in AcrB, likely altering the pocket environment to facilitate the recognition and transport of more hydrophilic molecules by the former (sequence alignment of AcrB with AcrD highlighting AP and DP is given as Figure 1 in Appendix). However, this conclusion should undergo a possible validation through the rationalization of the impact of the dynamics on a series of descriptors chosen to gain insights into the differences between AcrB and AcrD.

The first descriptor we considered is the volume of the pockets, although a pure steric filter of substrates by the different affinity sites is quite unlikely because of the large volumes of all the pockets considered in the present analysis. The mean volumes of AP and DP for AcrB are 2515 ±438 $Å^3$ and 2610±250 $Å^3$, respectively, having the corresponding values in AcrD at 3015±385 $Å^3$ and 2770±306 $Å^3$. The reported values point out that AP and DP in each transporter are essentially equally spacious. Surely, differences in the substrate specificities of AcrB and AcrD cannot be traced back to the size of such large pockets. Interestingly, in both transporters the dynamics produces a contraction of the pockets, if compared with the values evaluated for the crystal structures, when available, or in general for the configurations before MD simulations. For DP reduction of 30% (AcrB) and 28% (AcrD) with respect to the pre-MD structures, and for AcrB 35% with respect to the average value over a set of crystal structures are observed. This finding is in agreement with the conclusion of Fischer and Kandt [278], who noticed a closure of the DP in the *T* monomer of AcrB in absence of substrate during shorter MD simulations. In coherence with this hypothesis, we also found the DP volume to be at least 1000 $Å^3$ larger in substrate bound complexes [An unpublished study by our group found that benzene bound AcrB structures extracted from an MD simulation showed a consistently higher DP volume equivalent to that of crystal structures with doxorubicin and minocycline bound]. Moreover, population distributions over the clusters extracted from the MD simulations offer interesting insights into the different behavior of the two transporters. First, for both AP and DP, the first three clusters identified for AcrB cover roughly 90% and 64 % of the trajectories whilst the distribution is wider for AcrD, especially when AP is considered. Straightforwardly attributing this diversity to dissimilar flexibility might be not completely correct and could hide interesting features associated with the dynamics of the considered regions. We performed principal component analysis (PCA) on the MD trajectories considering the C-alpha atoms of corresponding Acr pumps to identify the dominant patterns of motions in the pockets. As visualized from the porcupine plots of the most prominent component (Figure 2 given in Appendix), it is evident that both the proteins have a flexible AP though with

differing magnitudes. For DP, the PCA data showed the pocket in AcrD with essential dynamics spread throughout the pocket unlike the less dynamic and more localized motions of DP in AcrB. Interestingly, the entire AP in AcrD exhibits almost a coherent motion with similar magnitude of eigenvector (depicted by length of the arrows) whereas in the case of AcrB, the loop residues 675-678 lining the base of AP show larger rearrangements. This loop represents a peculiar feature of the AP of AcrB, absent in AcrD. By analyzing the structures of the representatives of the AcrB clusters, it is immediate to recognize that, although different, all the clusters can be partitioned in two categories according to the behavior of the 675-678 loop (hereafter bottom-loop): the loop might be in a "up" or "down" conformation as depicted in Figure 9 below. The most populated cluster of AcrB is characterized by an up bottom-loop, the crystal structures exhibit only down configurations, as well as the second most populated cluster. A similar flip is not observed in AP of AcrD, and the analogous of the bottom-loop is always close to the Pre-MD arrangement. The lack of such a major conformational shift in addition to the substitution of the bulky Phenylalanine group in the switch loop of AcrB by Proline in AcrD, might possibly reduce steric hindrances for substrates of AcrD. The importance of this bottom-loop was already suggested by Fischer and Kandt by analizing their MD simulations [278]. Additionally, according to Kobayashi and co-workers the mutation E673G, in combination with Q569R and I626R, in AcrB transferred the β-lactam specificity of AcrD to AcrB. Thus, more than through the pure values of the volumes, specific structural features more directly involved in the entrance into and transport to a pocket might be of relevance in determining substrate specificity.
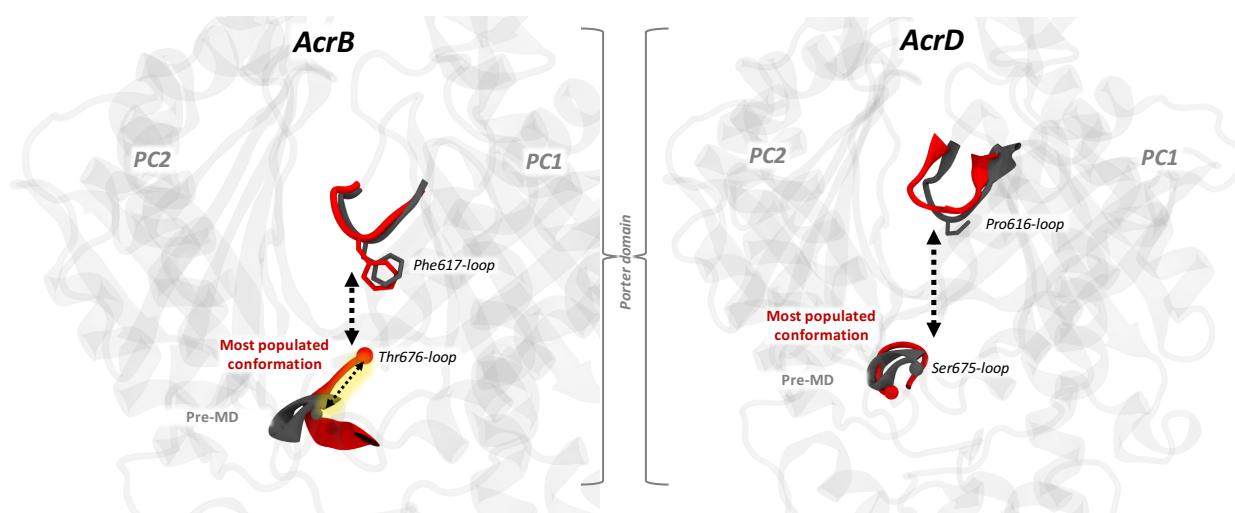
Figure 9. Main conformational states of the Thr676 loop (lining the base of AP) in AcrB (left panel) and of the corresponding Ser675 loop in AcrD (right panel). The configuration of the most populated clusters is in red and the pre-MD conformations in grey. The conformations of the switch loop are also reported.

Further, to quantify how different distribution of hydrophobic residues could affect substrate recognition and how this property is tuned by the dynamics of the protein, we calculated the lipophilic index of AP and DP of both transporters considering all of the cluster representatives obtained from MD simulations. This descriptor accounts for the sequence specificity and also includes an averaged dependence on the structure (see Section *Methods* for a more detailed discussion) and was evaluated over the ten clusters extracted from the MD trajectories. The pockets of AcrB are characterized by higher values of the lipophilic index than AcrD: AcrB has lipophilic indexes of 7.0 ± 1.0 and 13.1 ± 3.3 for AP and DP, respectively, to be compared with the values of 1.6 ± 0.6 and 1.9 ± 1.0 for the analogous pockets in AcrD. The higher lipophilic nature of AP in AcrB when compared to that in AcrD is required for AcrB to provide a favorable environment for its hydrophobic substrates to bind. However, the specific chemical environment of the access pocket is neither entirely hydrophobic nor entirely polar in both the proteins. Such a dispersed nature stabilizes binding of potential ligands by weak polar and hydrophobic interactions [279], while at the same time facilitates easy transport by preventing strong binding of the substrates to the pocket. Conversely, the lower lipophilic profile of AP in AcrD implies the presence of hydrophilic residues, which could be either positively or negatively charged, thereby influencing the substrate types recognized and transported by this pump. Note that the two values for AcrD are essentially identical whilst in AcrB there is a marked difference between the lipophilic indices of AP and DP, being the latter the more lipophilic. This could be an indication of DP being the site where substrates pre-selected at the AP might be differentiated in terms of their lipophilicity. The difference between the DP of AcrB and AcrD became more prominent on comparing their molecular lipophilic surfaces (Figure 6B). The LI isosurfaces are significantly larger in AcrB than in AcrD, which correlates well with the nature of the reported substrates transported by AcrB. Interestingly, the presence of phenylalanines only in the switch-loop of AcrB creates a large hydrophobic bridge between the DP and the AP, which would facilitate anchoring of aromatic compounds from the AP and their subsequent transport to the DP. The presence of polar/charged residues in the DP of AcrD would result in its increased hydration when compared to the DP of AcrB and the nature of water dynamics in this region would further influence the binding behavior of potential substrate molecules. Remarkably, the values reported in Tables 4 and 5 indicate that the dynamics affects more the DP than AP: the DP lipophilic index decreases from 16.3 (Pre-MD value) to 13.1 (from MD) for AcrB and from 4.9 to 1.9 for AcrD. This is a consequence of the overall shrinkage of the pockets and also of the rearrangements of the residues. Indeed, lipohilic indexes of AP do not exhibit the same reduction although the shrinkage of the volumes is essentially of the same size.

The local stereochemistry and distribution of functional groups in a region govern both the ordering of water molecules and their biologically important interactions in that region. The structure and the dynamics of the first water hydration shells around a putative binding pocket is of primary importance, given the relevance of water displacement for the free energy balance of the recognition event [241]. However, to insert this feature in a more accurate framework the radial distribution function (RDF) is not enough because it lacks the ability to present the information about the spatial positions of hydration zones. An example is the AP where the corresponding RDFs for AcrB and AcrD (Figure 4A) do not exhibit clear differences, which are evident in the spatial distribution function (SDF) reported in Figure 4B. Even for DP, RDF and SDF complement each other and provide a clear evidence of the contribution of the HP-trap in determining the hydration of the domain. In Figure 8A (left panel), the profiles for AcrB and AcrD overlap nicely but the same profiles differ remarkably as the calculation is restricted to the HP-trap of AcrB and the corresponding region of AcrD (right panel). Only in the case of AcrD the peak at ~1.9 Å remains. Figure 8B emphasizes even more the difference by comparing the spatial distribution. Several spots are homogeneously distributed for AcrD (right panel) whereas the DP of AcrB has several zones without spots, especially around the HP-trap.

An additional property that should be considered in this attempt to characterize the relevant pockets of AcrB and AcrD is the electrostatics. An electrostatic complementarity between the pocket and substrate molecules is essential for initial substrate recruitment and augmentation of their association rate [280, 281]. Concerning AP, a closer inspection of the electrostatic surface reveals positively charged residues like Arg568 and Arg625 of AcrD, that have been recently reported as key residues responsible for specificity to negatively charged molecules like the anionic β-lactams [108]. These residues were among those substitutions involving gain of electrostatic features in AcrD. For AcrB, the distribution of positively and negatively charged patches is more homogeneous without any predominance in the AP. However, a relatively dense positive environment in DP of AcrD compared to AcrB is identifiable. This higher positive charge originates from the electrostatic contributions of amino acids like Arg and Lys replacing their less polar counterparts in AcrB. In spite of this predominant positive nature, the presence of non-localized negatively charged and neutral patches allows for the binding of other substrates in AcrD. On the other side, the poor electrostatic environment of the DP in AcrB permits the binding of charged molecules like anionic beta-lactams but with far less affinity than that in AcrD.

3.5. Conclusions

In this study, we performed a comparative analysis of the physicochemical properties like volume, lipophilicity, electrostatic potential and hydration of AcrB and AcrD to rationalize the basis of differential substrate specificities shown by these homologous RND pumps of *E. coli*. Our results reveal that both the access and the deep pockets are considerably different between these two pumps with respect to the pocket descriptors computed here. Since the cavity volume essentially stays large enough to accommodate all potential substrate molecules, it possibly has an indirect effect on the lipophilic and electrostatic environment which altogether with the ensuing hydration within the pocket govern the recognition and transport of substrates by these pumps. More exhaustive studies including molecular docking and molecular dynamics simulations of selected substrates in the binding pockets of AcrB and AcrD are being considered to provide substantial information to further characterize these putative binding sites on the basis of substrate-protein interaction pattern.

In summary, our results suggest that features such as shape, lipophilicity, electrostatic potential and hydration in the vicinity of access and deep binding pockets are a few distinctive features between AcrB and AcrD. This is in agreement with the findings on other multidrug transporters where nonpolar and aromatic side chains impose specific prerequisites on drug size and shape [282].

Chapter 4

Molecular Rationale Behind the Differential Substrate Specificity of RND Transporters MexB and MexY

4.1. Introduction

Pseudomonas aeruginosa is an opportunistic human pathogen and a leading cause of nosocomial infections. This Gram-negative bacterium has been recognized as an increasingly important and worrisome species in health care-associated infections due to the emergence and spread of multidrug-, extensive drug-, and pandrug-resistant *P. aeruginosa* infections susceptible to very few antimicrobial agents [283, 284]. The intrinsic resistance of *P. aeruginosa* to multiple clinically relevant antibiotics results from the synergy between its low permeable outer membrane and the action of (chromosomally encoded) multidrug efflux systems like the ones constituted by the RND superfamily of secondary transporters [285, 286]. The exceptionally low outer membrane permeability (~12-100-fold lower than that of *E. coli [287]*) is, however, ineffectual without a secondary resistance mechanism such as efflux [288]. The RND transporters in *P. aeruginosa* contribute to both intrinsic as well as acquired resistance in this bacteria [289, 290]. Several RND type efflux systems have been identified in *P. aeruginosa* PAO1 [27, 89, 106, 263, 291-299], the most significant determinants of multidrug resistance being MexAB-OprM [300-302] and MexXY-OprM [90, 303, 304]. These two composite Mex systems contribute in an additive manner to the resistance to antibiotics [305, 306]; moreover, their different specificities (viz. MexAB-OprM for β-lactams and MexXY for aminoglycosides) extend the resistance profile of the infectious strains to numerous chemically diverse classes of antibiotics [306]. Such changes in the antibiotic susceptibility of these organisms can result in an increase in the minimum inhibitory concentration (MIC) of a drug to a level greater than clinically achievable [286].

      The MexAB-OprM tripartite system was the first RND-type multidrug efflux system to be discovered in *P. aeruginosa* at approximately the same time as the AcrAB system of *E. coli* in 1993 [302]. The MexXY system, identified later, is the only one among the 12 identified RND systems that mediates aminoglycoside resistance in *P. aeruginosa* PAO1 with numerous reports of clinical isolates during the last decade [106, 284, 307]. In fact, MexY is one of the relatively few bacterial transporters of the RND superfamily known to carry aminoglycosides [308] (along with AcrD of *E.*

coli). Owing to comparable homology of MexY to MexB and AcrB (Table 1) on which extensive studies have been carried out, MexY is expected to resemble them in the global features like structural fold, location of binding pockets and functional mechanism [117, 309, 310] (The general structural and functional aspects of RND transporters are discussed in Chapter 1).

Table 1. *Sequence identity (similarity) between RND transporters of E. coli and P. aeruginosa calculated using EMBOSS Stretcher [128] (All values are in percentages)*

Sequence	MexB	MexY	AcrB	AcrD
MexB	-	46.9 (65.5)	69.8 (83.2)	61.1 (76.5)
MexY	-	-	47.9 (67.0)	48.2 (66.7)
AcrB	-	-	-	65.3 (79.7)
AcrD	-	-	-	-

Despite the important degree of homology of MexB and MexY between them and with AcrB and AcrD of *E. coli* (Table 1), these transporters show certain similarities and differences in their substrate specificities. The differences between the Mex systems are slight yet significant. For instance, the substrate specificity of MexB is very similar to that of AcrB (Table 2). On the other hand, MexY specificity is extremely broad and is one of the principal determinants of aminoglycoside (gentamicin, tobramycin, amikacin, isepamycin) resistance in *P. aeruginosa* but might not contribute to the intrinsic resistance to beta-lactams, despite its potency in the extrusion of several beta-lactams [91]. Previous studies on the Mex pumps focused on identifying substrate recognition sites and the importance of the periplasmic loop regions by chimeric (domain swapping) studies [123, 311]. In addition, a few studies attempted to identify the substrates of Mex pumps [91, 224], the residues involved in substrate selectivity [121, 129] and also to explain the structural basis for the differential binding of inhibitors to MexB and MexY [277].

Table 2. *Substrate specificities of homologous MexB and MexY transporters in P. aeruginosa [277, 301, 306, 312-318].*

Transporter(s)	MexB	MexY	MexB and MexY
Substrates	Most beta-lactams (except imipenem), Quinolones, chloramphenicol, novobiocin	Aminoglycosides (gentamicin, tobramycin, amikacin), Quinolones, chloramphenicol, Penicillins (except carbenicillin and sulbenicillin), Cephems (except cefsulodin and ceftazidime)	Macrolides, Fluoroquinolones and Tetracyclines

The molecular basis for the diversity in the substrate profile of these Mex pumps, still largely remains unclear. In fact, one of the most obvious possibilities would be the differences in some

structural and chemical features within the putative substrate-binding pockets of which little is known in these transporters. The absence of an experimentally solved structure of MexY and the availability of only one drug (inhibitor) bound MexB structure have further hindered such structural studies. In addition, as these biological systems are not static *in vivo*, understanding their dynamics especially in terms of the preferred conformation, of interaction with the solvent, and of the nature of the putative pockets in this preferred state is essential for a more reasonable comparison. In this chapter, we characterize and compare the molecular properties (pocket descriptors) like accessible binding volume, lipophilic index, electrostatic potential, and hydration profile of the putative binding pockets (AP and DP in the *L* and *T* monomers, respectively) on structures derived from X-ray crystallography, homology modeling, and extensive MD simulations. Our study highlights the similarities and the differences in these features that could affect the molecular recognition process in these pumps in addition to their primary amino acid sequence. In particular, this study would be informative to new drug design attempts to correlate the different specificity patterns of these Mex transporters to the physicochemical as well as topographical properties calculated on (or projected onto) the molecular surface of their multifunctional recognition sites.

4.2. Methods

4.2.1. Homology modeling of MexY

A reliable three-dimensional atomic structure of the system of interest is essential for any structure-based computational study. Since no experimental structure of MexY has been solved yet, we built a hypothetical model of its asymmetric trimer structure by multiple template-based homology modeling. The amino acid sequence of full length MexY transporter protein from *P. aeruginosa* was retrieved from the UniProt database [265] (UNIPROT ID: Q9ZNG8), and subsequently searched for the best available template structures bearing homologous relationship to the query sequence using the NCBI-BLAST tool [161] against the Protein Data Bank (PDB) (www.rcsb.org) [266]. Both MexB of *P. aeruginosa* and AcrB of *E. coli* sequence showed a comparable identity (~47% and ~48%, respectively) and similarity (~66% and ~67%, respectively) with MexY sequence with least gaps (none in the binding pockets) over maximum sequence coverage. We therefore chose the high resolution crystal structure of AcrB (PDB ID: 4DX5, 1.9Å [111]) and MexB (PDB ID: 3W9I, 2.7Å [277]) for multiple-template based modeling of MexY. The protein sequences were optimally aligned by ClustalOmega [162] and the results were visually inspected to ensure the absence of gaps in important secondary structure regions. Modeller 9.13 [152] was used to generate a total of 100 asymmetric models of MexY based on AcrB and MexB templates using

an optimization method combining slow MD with very thorough variable target function method (VTFM) through 300 iterations, and this whole cycle was repeated twice unless the objective function MOLPDF was greater than 10^6. The resulting models were ranked using DOPE [267] score values, and the top 5 models (with the lowest DOPE score) were selected for individual structure quality checks. Each model was further subjected to loop refinement using Modeller, and to overall structure relaxation by energy minimizations using AMBER14 [191]. The most reliable model was then selected based on various geometric and stereochemical quality [268] factors calculated for backbone angles, side chains flips, rotamers, steric clashes etc. evaluated using PROCHECK [169], ERRAT [269], ProSA [165], Verify3D [167] programs available in MolProbity [270] and Structure Analysis and Verification Server tools (SAVES).

We then performed comparative structural evaluation by superimposition of the modeled MexY structures over the experimentally determined X-ray crystal structures of AcrB and MexB used as templates. Likewise, the template structures were also evaluated with the same programs to serve as reference for the results obtained for MexY model. Visual inspections were performed with VMD1.9.1 [271] and PyMOL [272].

4.2.2. Molecular dynamics simulations of MexB and MexY

MD simulations of the crystal structure of MexB (PDB ID: 3W9I) and of the homology model of MexY were carried out using AMBER14 molecular modeling software [191]. The topology and the initial coordinate files for these apo-protein structures were created using the *LEaP* module of AmberTools14 after suitably embedding them in POPE bilayer patches, solvating with explicit TIP3P water model, and neutralizing the residual charge of the system with required number of randomly placed K^+/Cl^- ions [208-211]. The ions count was suitably adjusted to account for an osmolarity of 0.15M KCl. Embedding of the protein into a pre-equilibrated POPE bilayer patch was done using the PPM server [212] and subsequently the CharmmGUI tool [213]. The lipid residue nomenclature was converted from the CHARMM to AMBER format using the *charmmlipid2amber.py* python script provided with AmberTools. The central pore lipids were added after calculating the number of lipids to be added to each leaflet by dividing the approximate area of the central pore by the standard area per lipid of POPE molecules [192]. Periodic boundary conditions were used and the distance between the protein and the edge of the box was set to be at least 15 Å in each direction.

Multi-step energy minimization with a combination of steepest descent and conjugate gradient methods was carried out using the *pmemd* program implemented in AMBER14 to relax internal constrains of the systems by gradually releasing positional restraints. Following this, the

systems were heated from 0 to 310 K by a 1 ns heating (0-100 K) under constant volume (NVT) followed by 5 ns of constant pressure heating (NPT) (100-310 K) with the phosphorous heads of lipids restrained along the Z-axis to allow membrane merging and to bring the atmospheric pressure of the system to 1 bar. Langevin thermostat [215] (collision frequency of 1 ps^{-1}) was used to maintain a constant temperature. Multiple short equilibration steps of 500 ps under anisotropic pressure scaling (Berendsen barostat) in NPT conditions were performed to equilibrate the box dimensions. A time step of 2 fs was used during all these runs, while post-equilibrium MD simulations were performed with a time step of 4 fs under constant volume conditions after hydrogen mass repartitioning [216]. The particle-mesh Ewald (PME) algorithm [203] was used to evaluate long-range electrostatic forces with a non-bonded cutoff of 9 Å. During the MD simulations, the length of all R-H bonds was constrained with SHAKE algorithm [186]. Coordinates were saved every 100 ps. The ff14SB [214] version of the all-atom Amber force field was used to represent the protein systems while lipid14 [192] parameters were used for the POPE bilayer. To improve statistical reliability of the results, four nearly 1 μs-long simulations were performed for each transporter, adding to a total simulation time of around 8 μs. Graphs were generated using the xmgrace [218] plotting tool. Trajectory analysis was performed using *cpptraj* module [217] of AmberTools14 and VMD1.9.1.

4.2.3. Clustering of MD trajectories

Cluster analysis of the MD trajectories was performed using the average linkage hierarchical agglomerative clustering method implemented in *cpptraj* module of AMBER. This helps to reduce the number of structures for analysis still retaining the large conformational space sampled during the MD runs. In this approach, we clustered the trajectory independently based on RMSD (cutoff set to 3 Å) of the AP in *L* monomer and the DP in *T* monomer for both the proteins. The representative structure of each of the 10 top clusters generated in each case was taken for the calculation of volumes, the molecular lipophilicity indexes and the electrostatic potential.

All non-protein molecules were stripped from the trajectory during post-processing to reduce additional memory usage and to speed up file processing.

4.2.4. Pocket descriptors

The list of the pocket descriptors identified for the present study contains:

1. Cavity volume;
2. Molecular lipophilicity potential;
3. Electrostatic potential;
4. Hydration.

The various pocket descriptors used to characterize the binding site [273] were calculated using specific programs after validating their applicability to RND systems by assessing results against available crystal structures and experimental data

4.2.4.1. Cavity Volume

Evolution of size and shape of the AP and DP during the MD simulations was examined using the two-probe sphere method of *rbcavity* program bundled in rDock suite [231]. This allows obtaining detailed information on the pocket volume and plasticity of the site. In this method, the potential binding site volume was identified by a fast grid-based cavity detection algorithm [232] within a sphere of radius 14 Å, centered over the pockets, using large and small probe radii of 6.0 Å and 1.5 Å, respectively.

4.2.4.2. Molecular Lipophilicity Potential

Lipophilicity is one of the indicators of the various intermolecular interactions [234] (viz. steric, hydrophobic and electrostatic) that the molecule can elicit in a receptor-bound state. The three-dimensional distribution of lipophilicity in space or on a molecular surface can be described using Molecular Lipophilicity Potential (MLP), which represents the influence of all lipophilic fragmental contributions of a molecule on its environment. The lipophilicity of AP in *L* monomer and DP in *T* monomer were qualitatively and quantitatively estimated in this way using MLP Tools [237] plugin available for PyMOL.

4.2.4.3. Electrostatic Potential

The electrostatic potential surface maps were computed by APBS [239], after preprocessing the structures of MexB and MexY to assign charges and radii using PDB2PQR server [240]. All electrostatic potential calculations were performed at 0.15 M physiologic salt concentration, with a solvent probe of radius 1.4 Å, a solvent dielectric constant of 78.5, a biomolecular dielectric constant of 2.0, a temperature of 310 K, a minimum grid spacing of 0.5 Å while keeping the other Poisson-Boltzmann parameters at default.

4.2.4.4. Hydration Analysis

The radial distribution function (RDF) profiles indicate the probability of finding water molecules at a certain distance from a region or residue of interest and is commonly used to analyze the solution structure revealed from either experimental or computer simulations data. The RDF analysis of water oxygen atoms was performed using *cpptraj* module of AMBER14, in which the RDF of all water oxygen atoms to each atom of AP in *L* monomer and of DP in *T* monomer was computed over the entire length of the simulation.

Though RDF clearly shows a difference in the water distribution around the desired regions, it lacks the ability to present the information about the spatial positions of these differences. Hence,

spatial distribution function (SDF) [242] of waters around the whole protein was calculated on the configurations extracted from MD simulations and forming the most populated cluster using the Gromacs utility *g_spatial* [274] to determine the three-dimensional density distribution of aqueous solution around the binding pockets of these transporters. RDF and SDF together highlight the hydration around the binding pockets of these proteins, which can be effectively used to understand the molecular mechanism of interaction of water molecules penetrating the pocket in a dynamic manner.

4.3. Results

4.3.1. Primary sequence analysis

On comparing MexB to MexY (sequence alignment given as figure 3 in Appendix), we noticed that both AP and DP were less conserved than the overall proteins, sharing only around 35% and 33% identities, respectively. However, the DP showed a better conservation in its residue type (78% vs. 67% for the AP), which is predominantly hydrophobic compared to the AP in these Mex transporters. Most of the residues identified as essential to establish interactions with the substrates and/or the inhibitors in AcrB [108, 109, 111, 116, 118, 211, 275, 276] were well conserved in MexB. The characteristic hydrophobic trap (HP-trap) sitting within the DP and rich in phenylalanine residues was completely conserved in MexB but not MexY. The HP-trap does interact only smoothly with the transported drugs [319], but is likely a preferred target site for inhibitor designing in AcrB [133, 225, 275] and has also been shown in AcrB to indirectly affect drug binding by site-directed mutagenesis [320] and molecular dynamics simulations [211]. Nevertheless, MexY appeared to retain most of the hydrophobic nature in this region by substitutions involving other hydrophobic residues except for a polar uncharged residue Y605 (corresponding to F610 in MexB).
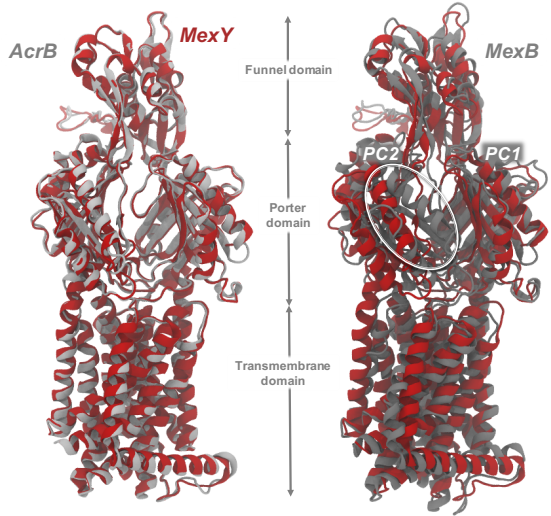
4.3.2. Structure of MexY predicted by homology modeling

Visual inspection of the top 5 models of MexY generated by multiple template based homology modeling revealed an overall identical 3D structural fold characteristic of RND transporters. The final MexY model featured 94.3% and 99.2% of residues in the favored and allowed regions of the Ramachandran plot, respectively. The overall quality factor for non-bonded atomic interactions assessed with ERRAT plot further confirmed the high quality of our model with a score of 93.34%, which was in the range of high-resolution experimental structures. The ProSA evaluation showed that the overall model quality (Z-score) of the homology model (-11.6) and the template crystal structures (-12.4; -12.3) were comparable and within the range of scores typically found for native

proteins of similar size. The local quality of the model based on interaction energies for each residue as evaluated with ProSA served as an additional verification of model quality with most of the residues falling in the negative energy scale and only a minor set on the positive side.

The evaluation results are summarized in Table 3, which also shows the structure of the MexY model superposed to its AcrB and MexB template structures, highlighting the similarity in their general structural fold. Note that in the MexB crystal structure, the PC2 subdomain is significantly shifted towards the transmembrane domain and towards PC1, which was also discussed by Sennhauser *et al.* based on their first published X-ray structure of MexB [135]. This conformational shift constricted the substrate entrance channel towards the periplasmic side formed by the PC domains in MexB. This channel between PC1 and PC2 is open in AcrB, continuing to the access pocket in the *L* and *T* monomers. The MexY model chosen in our case resembles AcrB conformation in this aspect.

Table 3. *Evaluation results of the multiple-template based MexY model. (Left panel) Structural superposition of MexY (red) onto MexB and AcrB (both in gray in the respective inset) showing only a single monomer for clarity.*

Evaluation Criteria	MexY model
RMSD (Cα-atoms)	3.6 Å (MexB) 1.0 Å (AcrB)
TM-score	0.9
Ramachandran favored	94.3%
Errat	93.34%
Verify-3D	Pass

4.3.3. Molecular dynamics simulations of MexB and MexY

The stability of the MexY model and its suitability for subsequent analyses were further validated by performing 4 independent μs-long MD simulations. In particular, monomer-wise RMSD of the protein backbone and of the whole protein with reference to the initial structure established the stability and suitability of the MexY model for subsequent analyses. The cluster representatives from the equilibrated trajectories of MexB and MexY were used to characterize the distribution of accessible binding volumes, molecular lipophilicity indexes and electrostatic potential, while the trajectories themselves were analyzed for hydration patterns within the AP and DP, to identify

similarities and differences in the binding pockets of these two transporters. In the following, we present the results from these characterizations on the two aforementioned sites similar to that of Acr transporters discussed in Chapter 3.

4.3.3.1.   Access Pocket

A. Pocket Volume

The accessible binding volume provided to a ligand is the first of the many factors governing the optimal binding in addition to shape and electrostatic complementarity [231]. Promiscuous RND transporters were earlier identified to have a large binding site with a reasonable degree of plasticity to facilitate binding of molecules of a wide range of sizes [321-323]. In order to assess the impact of available binding volume on substrate recognition and binding, we calculated the volume of AP within the *L* monomer for all of the cluster representatives.

The exhibited volumes of the AP in *L* monomer were large enough for their corresponding substrates to bind in different orientations. The AP of MexB featured an average value of around 1278 $Å^3$ with the most populated cluster (about 28% of the simulation time) showing a pocket volume of around 1565 $Å^3$. In the case of MexY, the AP of *L* monomer showed a larger average volume of 2192 $Å^3$ with the most populated cluster showing around 2575 $Å^3$ for about 19% of the simulation time (Figure 1). The average volume was slightly reduced in both cases from the corresponding initial (Pre-MD) structures of MexB (1336 $Å^3$) and MexY (2222 $Å^3$), probably due to a partial closure of the binding pocket in the absence of bound substrates [278] (Table 4). It is to be recalled here that the analogous pocket in AcrB and AcrD showed average volumes of 2515 $Å^3$ (±438 $Å^3$) and 3015 $Å^3$ (±385 $Å^3$), respectively, during MD and volumes of 3620 $Å^3$ and 3760 $Å^3$, respectively, in their Pre-MD structures.
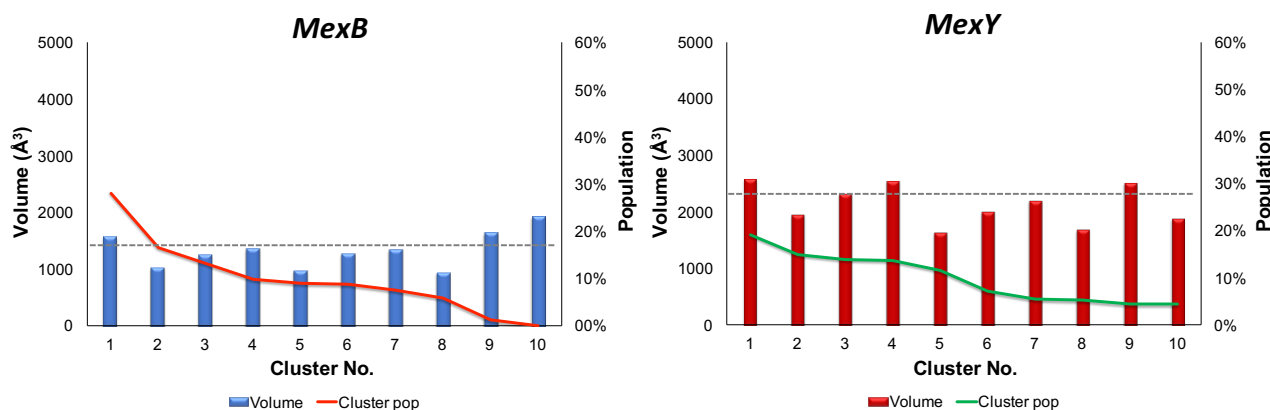
Figure 1. Volume distribution of AP in the L monomer of MexB and MexY over the simulation timescale. The corresponding volume in their Pre-MD structures are shown as dotted line.

Table 4. Pocket volume of AP in the L monomer of MexB and MexY.

System	Volume (Å^3)	
	Pre-MD	MD (Average ± SD)
MexB	1336	1278 ± 228
MexY	2222	2192 ± 340

Since the values of MexB, which is supposed to share a high homology with AcrB, showed significantly low volumes of the AP, we examined the available crystal structures of MexB (drug-free and inhibitor bound) and also homology models built based on AcrB with open AP and found their volume to be larger than the pre-MD crystal structure in this case. The constricted nature of MexB with narrower pockets in comparison to AcrB was also reported by Sennhauser *et al.* [135].

B. Molecular Lipophilicity Potential and Lipophilicity Index

In order to quantify the lipophilicity of the pocket and estimate its changes during protein dynamics, we calculated the lipophilic index (LI) of AP in *L* monomer of both transporters considering all of the cluster representatives obtained from MD simulations.

The results showed the AP of MexY to be slightly more lipophilic than that of MexB, which is quite unexpected based on the composition of pocket residues (Table 5). However, since the LI estimation method implemented here is limited by the fact that it has to identify a pocket within the defined grid, a likely chance that the collapsed pocket in the Pre-MD conformation of MexB could have influenced the LI calculation can not be neglected. To evaluate the level of impact on LI due to the collapse of the pocket, we calculated the LI on the available crystal structures of MexB and also structures built from AcrB as templates to have AP in an open state. The LI values in all these structures were much higher than the Pre-MD and a maximum LI value of 9.9 was seen in the crystal structure identified by the PDB code 2V50 [135].

Table 5. Lipophilic index of AP in the L monomer of MexB and MexY.

System	Lipophilic Index (LI)	
	Pre-MD	MD (Average ± SD)
MexB	2.7	2.7 ± 0.9
MexY	4.2	4.5 ± 1.2

As expected, the specific chemical environment of the pocket was neither entirely hydrophobic nor entirely polar in both the proteins (Figure 2) which is partly essential for weak binding with dispersed interactions for potential substrates to be transported further. Additionally, there was no

difference between Pre-MD and MD values of MexB. This is different from what we observed for AcrB and AcrD.
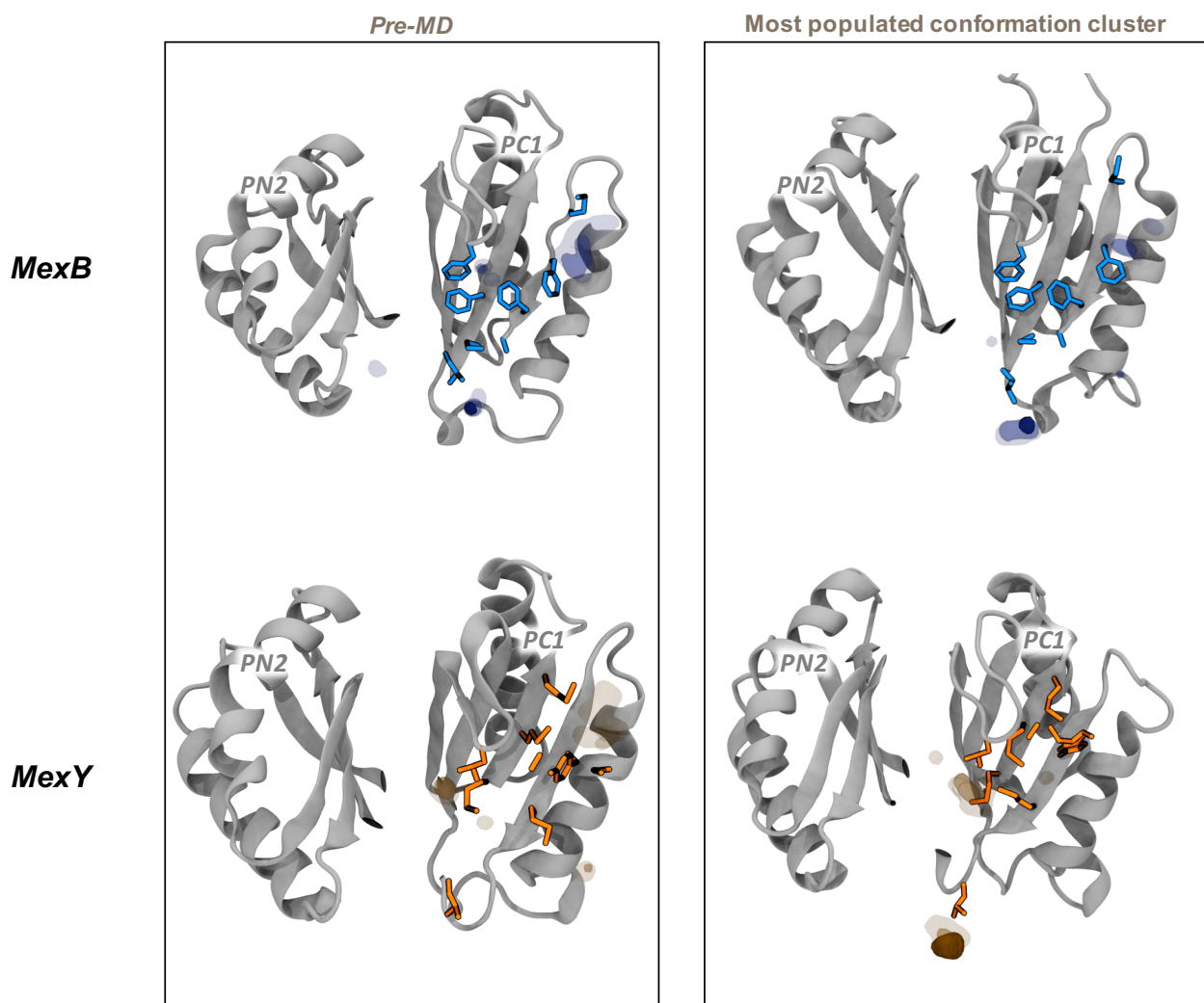
Figure 2. *MLP surface observed within 4 Å of the AP in the L monomer of MexB (blue) and MexY (orange) as seen from PC2-PN1 side. The hydrophobic/aromatic residues of the pocket are shown as sticks in the respective structures (isosurfaces at 0.75 (solid), 0.5 (dark transparent) and 0.25 (light transparent) are shown).*

C. *Electrostatic Potential*

The AP of MexY has a slightly greater population of polar and charged residues compared to that of MexB. In order to assess the contribution of this difference on the electrostatic potential of the pocket, we calculated the electrostatic potential using Adaptive Poisson-Boltzmann Solver (APBS) and projected it onto the solvent accessible surface area of the two proteins. The electrostatic feature in MexB appeared to be localized as seen by the negative patch near the base of AP and positive patches near the AP exposed to the periplasmic cleft entrance. In the case of MexY, this

localization was less intense and an overall greater distribution of electrostatically positive patches within the AP relative to MexB was observed (Figure 3).
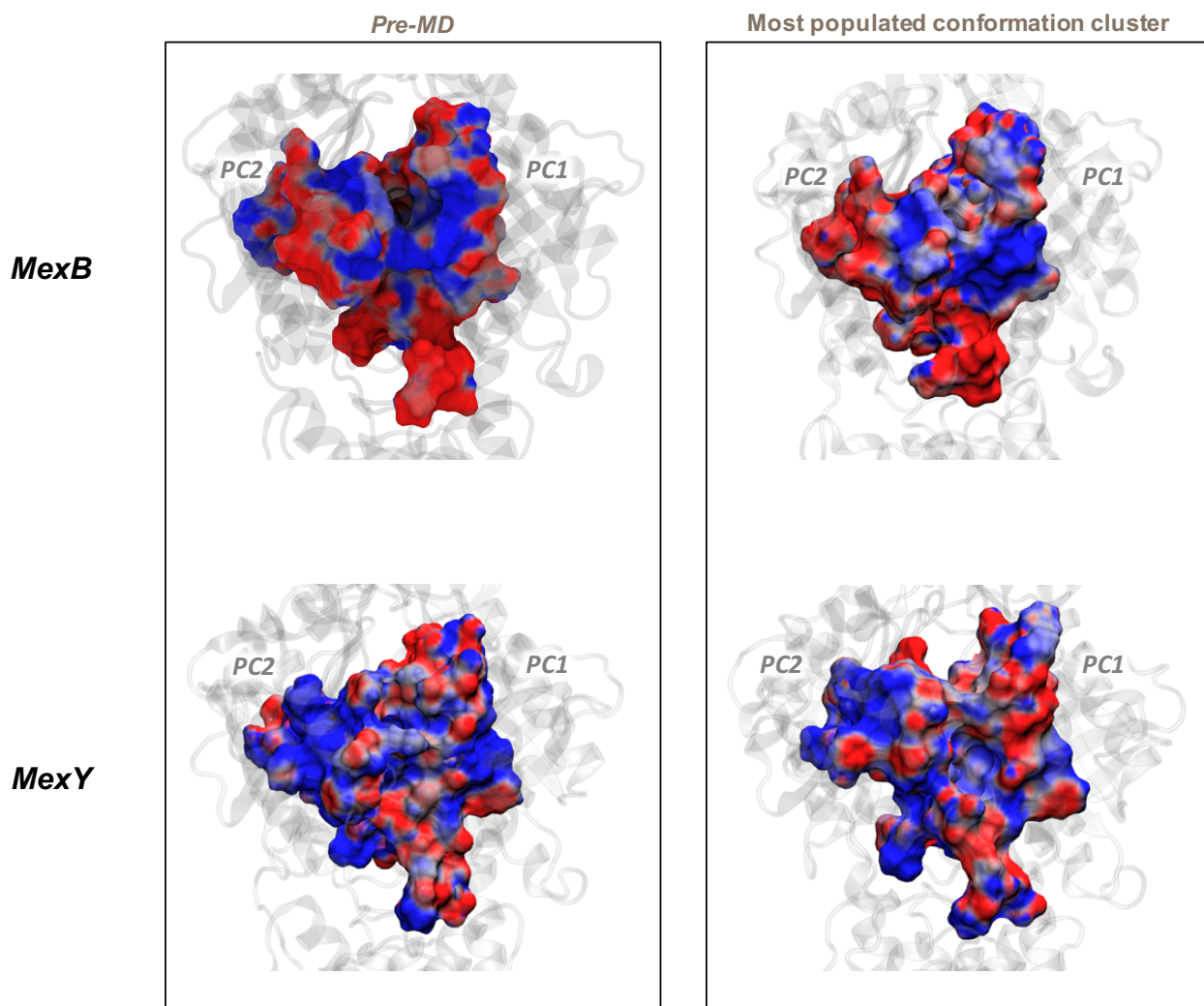
Figure 3. Electrostatic potential surface of the AP in the L monomer of MexB and MexY as seen from the periplasmic cleft entrance. (RWB color scale from negative (-10 k_bT/e_c) to positive potential (+10 k_bT/e_c)). Regions not part of the AP are shown as cartoon and are colorless irrespective of their electrostatic potential.

D. Hydration Analysis

The RDF and SDF profiles of water oxygen atoms around the AP residues of *L* monomer in MexB and MexY were assessed for any possible difference in the intensity of hydration, which could arise from the differences in the nature of residues lining the pocket as well as from the pressure exerted by the dynamic motions of the pocket. The first solvation shell was found at around 2 Å from any residue lining the AP in both the proteins, and displayed a slightly lower intensity in MexB (Figure 4A). The SDF was calculated on the configurations extracted from MD simulations and forming the

most populated cluster in MexB and MexY to get more insights into the spatial distribution of water in the pocket (Figure 4B).
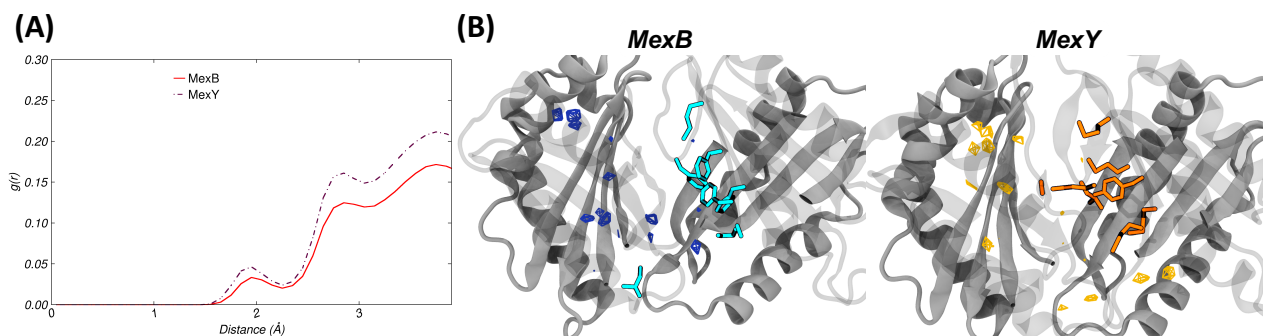
Figure 4. *(A) Comparison of RDF profiles of water oxygen atoms around the AP of monomer L in MexB and MexY. (B) Comparison of SDF for waters in AP of L monomer calculated on the configurations belonging to the most populated cluster of MexB and MexY illustrating the variation in the immediate environment of the hydrophobic residues. The isosurfaces are shown at density isovalue of 6.*

4.3.3.2.   Deep Pocket

The DP (likely the recognition site for low molecular mass compounds [109]), especially in AcrB and MexB is separated into two parts: a relatively hydrophilic main translocation channel and a phenylalanine rich hydrophobic trap (HP-trap). The channel is known to exhibit sufficient space for multisite drug binding, while the HP-trap forms a deep and narrow fissure. The DP exists in a collapsed state in *L* and *O* monomers but is supposed to be wide open in the *T* monomer; therefore, all the analyses concerning this site were performed on the *T* monomers of MexB and MexY.

A. Pocket Volume

The various conformations sampled during the MD simulations exhibited volumes in the range of 2000-2800 $Å^3$ with the most populated conformations in MexB (33.7%) and MexY (19.5%) presenting a volume of around 2500 $Å^3$ (Figure 5). The average volumes of the DP as extracted from the MD simulations of MexB and of MexY were quite similar (Table 6), and correspond to a large collapse of the pocket with respect to the conformations in the initial (Pre-MD) structures.
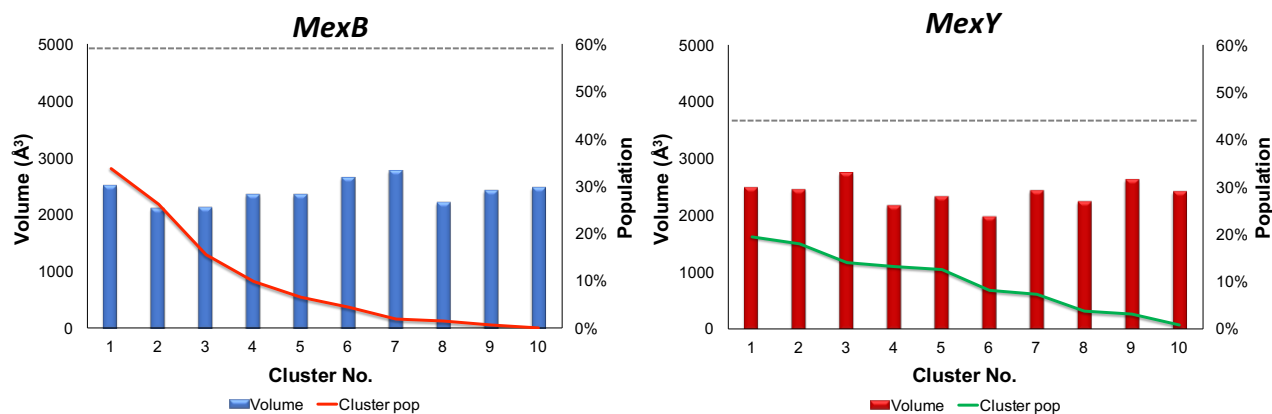
Figure 5. *Volume distribution of DP in the T monomer of MexB and MexY over the simulation timescale. The corresponding volume in their Pre-MD structures are shown as dotted line.*

Table 6. *Pocket volume of DP in the T monomer of MexB and MexY.*

System	Volume (Å^3)	
	Pre-MD	MD (Average ± SD)
MexB	5120	2310 ± 196
MexY	3840	2400 ± 210

B. Molecular Lipophilicity Potential

The difference between the nature of DP in MexB and MexY became more noticeable from their molecular lipophilic surfaces (Figure 6) and the LI values (Table 7). With its phenylalanine rich hydrophobic region wide open in MexB, we observed a relatively higher LI in the Pre-MD structure of MexB compared to that of MexY.

Table 7. *Lipophilic index of DP in the T monomer of MexB and MexY.*

| System	Lipophilic Index (LI)	
	Pre-MD	MD (Average ± SD)
MexB	20.1	4.1 ± 2.3
MexY	15.9	8.9 ± 2.2
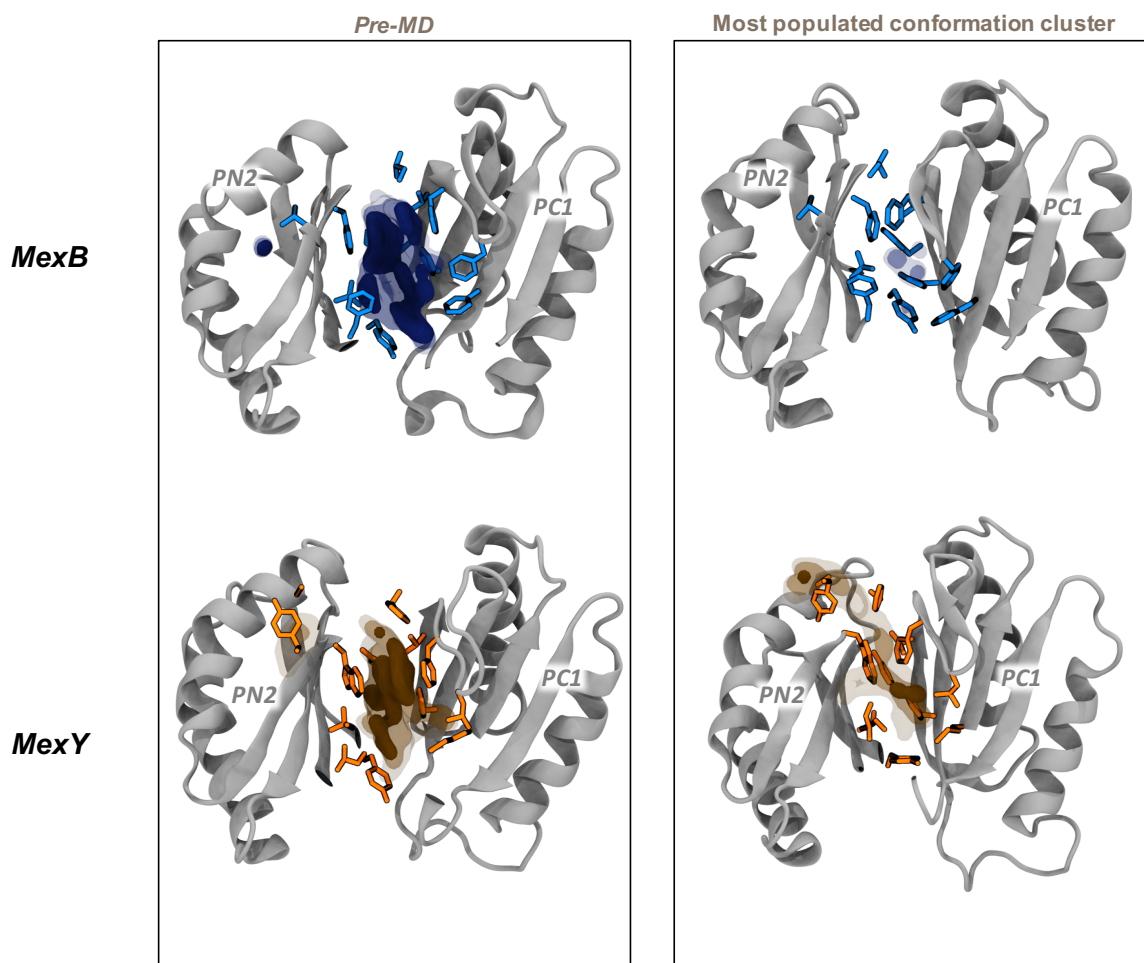
Figure 6. *Molecular lipophilic potential surface observed within 4Å of DP in the T monomer of MexB (blue) and MexY (orange) as seen from PC2-PN1 side. The hydrophobic/aromatic residues of the pocket are shown as sticks in the respective structures (Isosurfaces at 0.75 (solid), 0.5 and 0.25 (transparent) are shown).*

C. Electrostatic Potential

The electrostatic potential surface of the DP in *T* monomer of MexB and MexY indicated a significantly greater positive environment in MexB compared to the more negative pocket of MexY (Figure 7). The composition analysis of the pocket showed that the DP in MexY is composed of approximately 12% negatively charged residues while it is only around 4% in MexB. Likewise, the positive charged residues made up close to 6% of the DP in MexB in contrast to a mere 2% in MexY. This is in good agreement with the desired complementarity needed to accommodate the charged substrates transported by these proteins (e.g. zwitterionic or negatively charged β-lactams; positively charged aminoglycosides). The differences in the electrostatic potential between the DP pockets of MexB and MexY appear to be the most striking distinctive tracts of the two transporters.
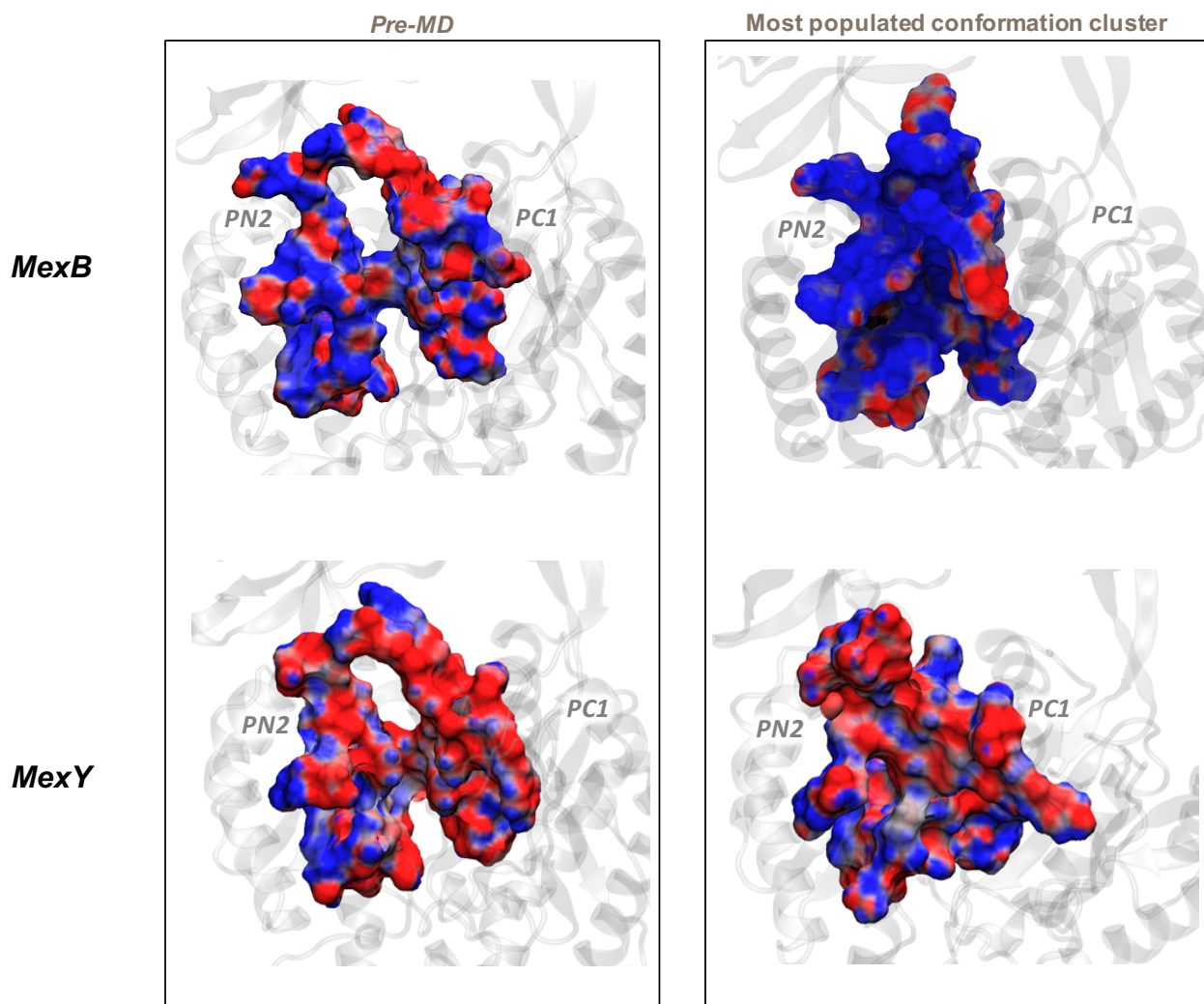
Figure 7. *Electrostatic potential surface of DP in the T monomer of MexB and MexY as seen from PC2-PN1 side. The color scale adopted is: red, negative (-10 k_bT/e_c) to blue, positive potential (+10 k_bT/e_c)). Regions not part of the DP are shown as cartoon and are colorless irrespective of their electrostatic potential.*

D. Hydration Analysis

The overall hydration of the DP as reflected by the RDF plot was not very different between the two Mex proteins with the first solvation shell appearing at around 2 Å from any residue lining the pocket (Figure 8). The hydration around the corresponding HP-trap regions was also analyzed and we found a similar trend as seen for the DP though the intensity was far lower in this case. The spatial positions of water dense regions as seen from the SDF (Figure 8) highlighted the dehydration of the HP-trap region in both the proteins.
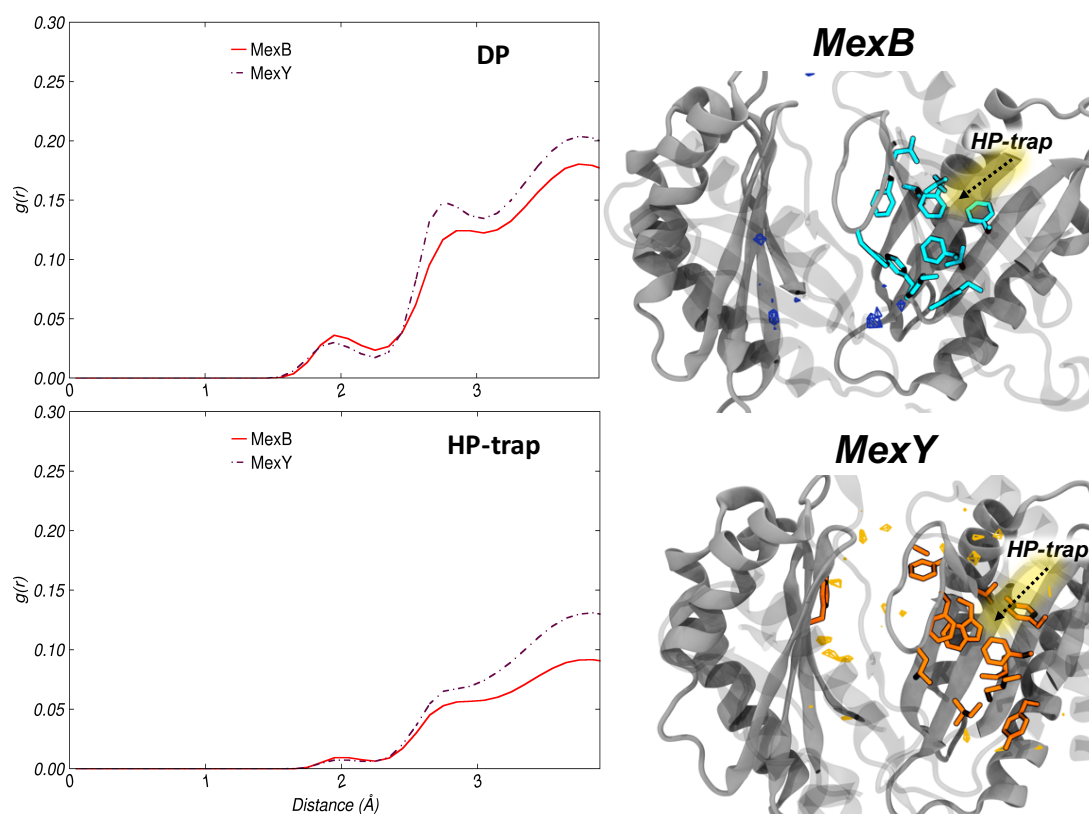
Figure 8. Comparison of hydration profiles of DP in monomer T of Mex transporters of P. aeruginosa. (Left panel) RDF profile of water oxygen atoms around the DP of monomer T in MexB and MexY. (Right panel) Comparison of SDF for waters in DP of T monomer calculated on the configurations belonging to the most populated cluster of MexB and MexY illustrating the variation in the immediate environment of the hydrophobic residues. The isosurfaces are shown at density isovalue of 6.

4.4. Discussion

RND transporters possess the ability to bind and transport several chemically distinct molecules. This becomes clinically significant in the case of bacteria such as *P. aeruginosa*, where active efflux of compounds via Mex pumps is one of the principal resistance mechanisms adopted to reduce antibiotic susceptibility. Although these transporters are identified to be promiscuous, they are selective towards certain classes of antibiotics and to certain specific antibiotics within a class. For example, MexB does not transport aminoglycosides whereas MexY does; likewise, MexB transports penicillins including carbenicillin and sulbenicillin whereas MexY transports all but these two penicillins. Although a few published studies have focused on identifying and understanding substrate selectivity in these pumps, no molecular rationale has been identified so far. We, therefore, adopted a full *in silico* atomistic approach to provide a possible explanation for the substrate selectivity of MexB and MexY transporters based on various physicochemical features (alternatively, pocket descriptors) of the two main putative substrate binding pockets (AP and DP). The pocket descriptors chosen include accessible binding volume, lipophilic index, electrostatic

potential and hydration, which are influenced by the primary sequence, the tertiary structure as well as dynamics of the protein. In addition, these descriptors should highlight the complementarity in the physicochemical properties of the pocket and the substrates transported by these pumps.

Since these descriptors include structural properties, which need a three-dimensional structure for their estimation and no experimentally solved structure for MexY has been available yet, we built and validated a hypothetical model of MexY based on the high-resolution experimental structures of its homologues MexB and AcrB, sharing a nearly equal percentage of identical residues with MexY. Moreover, we performed µs-long molecular dynamics simulations of both proteins solvated and embedded in a model membrane, and we extracted from the resulting trajectories the most-sampled conformations of the AP in *L* monomer and of the DP in *T* monomer for estimation of the pocket descriptors.

For a substrate molecule to get accommodated into a binding pocket, a primary criterion is the accessible volume inside the pocket. In accordance to one of the general features of promiscuous drug binding pocket, these transporters exhibited pocket volumes much larger than the volume of its largest substrate. Such large volumes might not screen the incoming molecules based on geometric criteria but indicate the possibility of the potential substrates to bind in more than one orientation. The AP showed a considerably smaller volume than DP in contrast to what we observed for Acr pumps. This is most likely due to its already constricted state in the crystal structure as also reported by Sennhauser *et al.* from their crystallographic studies [135]. In addition, since the pocket was already closed to a large extent in the Pre-MD structure, no further closure of the pocket as high as seen in the case of Acr transporters described in Chapter 3 occurs. In the case MexY, the sterically bulky side chain of W177 (corresponding to F178 in MexB) oriented into the DP kept the volumes low in both its Pre-MD as well as MD structures with only a fractional closure of the DP in contrast to the much wide open Pre-MD structure of MexB resembling its homologue AcrB.

Secondly, chemical descriptors like lipophilicity and electrostatic potential come into play. The lipophilic profile is essential to provide to a hydrophobic or amphipathic molecule sufficient hydrophobic interaction for suitable binding inside the pocket. Both these Mex transporters are known to favor binding of hydrophobic macrolides (viz. erythromycin, spiramycin) and the predominant number of hydrophobic residues in the binding pockets support this binding. However, the lipophilic index of both the AP and the DP in MexB during MD are lower than in MexY despite the fact that the former has a larger number of hydrophobic residues in the pockets. This result is possibly due to the limitation of the method used to estimate LI based on pocket identification within a grid, which is smaller in MexB than in MexY and both pockets are more collapsed in the

former protein. In spite of this limitation, there was no difference in LI of AP between Pre-MD and MD values of MexB. This is different from what we observed for AcrB and AcrD which originated from Pre-MD structures possessing wide open pockets. This indicates the impact of closed pocket on LI values. In the case of DP, even though the corresponding HP-trap region is conserved in its overall hydrophobic nature, the residues W177 and Y605 in MexY are less hydrophobic than their phenylalanine counterparts in MexB. In addition, the presence of the bulky side chain of W177 in DP further hinders the available pocket space, and eventually the LI of DP also. Nevertheless, the difference in LI between MexB and MexY is greater in the lesser-conserved DP than that in the AP, as observed for the homologous Acr transporters. This may indicate DP to be the site where substrates pre-selected at the AP are differentiated in terms of their lipophilicity.

Subsequently, for the recognition and binding of charged substrates (viz. polycationic aminoglycosides by MexY and zwitterionic or anionic β-lactams by MexB), an electrostatic complementarity is highly desired and essential for initial substrate recruitment and augmentation of their association rate [280, 281]. It is interesting to note that the electrostatic nature of MexY and MexB seen here is comparable to that of AcrD (more positive) and AcrB (more negative). However, based on homology to Acr pumps of *E. coli* in which residues essential for specificity to anionic β-lactams in AcrD were recently identified [108], we found the corresponding residues (Q in MexB/MexY at position corresponding to R568 in AcrD; M in MexB/MexY at position corresponding to R625 in AcrD) to differ between MexB and MexY. This may indicate the presence of a different selection filter for charged substrates in these Mex transporters. The DP, on the other hand, showed its electrostatic feature to be in good agreement with the desired complementarity expected in these proteins for their charged substrates. The greater negative charge in the DP of MexY may favor positively charged aminoglycosides (in the hydrophilic substrate-translocation channel part of DP) and disfavor negatively charged molecules; however, along with the scattered positive charges, the DP in MexY may feebly favor beta-lactams (especially zwitterionic). Likewise, MexB with its predominant positive electrostatic potential within the DP may attract negatively charged as well as zwitterionic β-lactams, and extrude them with greater efficiency along with weakly acidic quinolones such as cinoxacin and nalidixic acid, in comparison to their lesser efficiency in pumping out cationic antibiotics (oleandomycin, erythromycin, and puromycin). The reason for imipenem not being a substrate of these transporters could not be explained from our results but according to another study comparing the behavior of imipenem and meropenem [224], the authors identified that imipenem not only has a low affinity to both AP and DP but also has a slightly enhanced propensity of forming interactions with water molecules than meropenem (which is a good substrate of MexB).

Lastly, the hydration profile characterized from the trajectories further confirms the difference in the nature of residues lining the pocket. The hydration in AP did not show a very significant difference between MexB and MexY as was the case with lipophilicity and electrostatic potential. In contrast, the DP of MexY showed many high-density water regions than that of MexB, possibly due to greater number of charged residues in the translocation channel part of DP along with lesser geometric pressure by the pore domains. At a global level, the RDF plots of MexB and MexY are comparable to that of AcrB and AcrD, respectively, which can be related to the similarity in the overall physicochemical makeup of their putative binding pockets. But, we noticed that at identical isovalues where many water dense regions were found in AcrB and AcrD, very few such regions were featured in MexB than MexY. The observed difference in the SDF between the Mex and Acr pumps might likely be due to the pressure exerted by constriction of the AP in the former, inducing a larger and fairly continuous displacement of water molecules. In the case of DP, the presence of a polar residue (Y605) might have had a minor influence on the hydration of corresponding HP-trap region in MexY, provided it was less buried by the hydrophobic bulky side chain of W177. The observed differences in the hydration of these putative binding pockets obtained from RDF and SDF complement each other and present effectively the statistical structural information on pocket-water interactions obtained from MD simulations of such complex systems. Further characterization of hydration profile in terms of water residence time and hydrogen bond lifetimes within the DP of these proteins could give useful insights on any possible role of water in substrate transport [209, 324].

Collectively, these structural adaptations enhance protein-substrate binding and specificity, but the dynamic profile of the pocket and the resulting changes in the pocket descriptors in these Mex pumps also upsurges the corresponding dissociation constant, eventually creating a transient complex for effective transport.

4.5. Conclusions

We compared various physicochemical pocket descriptors between MexB and MexY and highlighted certain features, which could explain the similarities and differences in substrate selectivity within the multidrug binding property of these promiscuous transporters. Overall, based on the results from the comparison of various physicochemical descriptors of the putative binding pockets in MexB and MexY, it seems more likely that the DP could be the major substrate selection site though partial pre-selection may occur in the AP too. It is not to be forgotten that the multidrug-

binding pockets are large and flexible with multiple points of interaction for the substrate. In addition, previous studies on AcrB have shown their binding sites to be extremely adaptable to many substrates [223] and minor substitutions would still not influence the global properties of the binding pockets being compared in our study. More exhaustive studies including molecular docking and MD simulations of selected substrates in the binding pockets of MexB and MexY are being considered to provide substantial information to further characterize these putative binding sites on the basis of substrate-protein interaction pattern.

Chapter 5

Substrate-Transporter Interaction Analysis in AcrB and AcrD

5.1. Introduction

Many studies have characterized the molecular mechanism of recognition of substrate and inhibitor molecules by AcrB and reported certain residues and regions as important for substrate recognition and subsequent transport [102, 115, 116, 122, 211, 223, 224, 226, 264, 276, 279, 319, 322, 325-327]. In the case of AcrD, the other major transporter responsible for specificity towards hydrophilic compounds and a better transporter of anionic β-lactams than its promiscuous homologue AcrB, only a few studies have been reported till date where the authors attempted to identify substrate recognition site(s) [108, 114, 118, 130]. Of particular importance is the work of Kobayashi *et al.* [108], where they identified certain key residues in the AP as potential determinants of specificity towards negatively charged β-lactams (aztreonam, carbenicillin, and sulbenicillin) in these two Acr transporters. However, the various atomic level interactions underlying these recognition events remain unclear. In addition, having characterized the global differences in the putative binding pockets between AcrB and AcrD (Chapter 3), it could be interesting to determine if there are any preferred sub-sites for a specific substrate type, and/or the involvement of specific residues or residue type within the binding sites which influence the recognition pattern in these transporters.

To address this issue and provide interaction level insights responsible for the differential binding of certain substrates to these transporters, we performed molecular docking of AcrB and AcrD with two molecules which act as substrate, non-substrate or poor substrate of these pumps. The ligands chosen were minocycline (MIN) and aztreonam (AZT) (Figure 1) which belong to different classes of antibiotics and are structurally and chemically dissimilar. Minocycline, a tetracycline class antibiotic, is known to be transported by AcrB and has also been co-crystallized in the DP of the *T* monomer in AcrB [102, 111] but is not transported by AcrD [257, 260]. Aztreonam, an anionic β-lactam class antibiotic, is known to be a specific substrate of AcrD and a poor or non-substrate of AcrB [108, 261]. These two molecules also exhibit different physicochemical

properties such as charge (AZT is negatively charged while MIN is neutral under physiological conditions), logP (MIN is more hydrophobic than AZT), H-bond donors, H-bond acceptors etc. (Table 1).
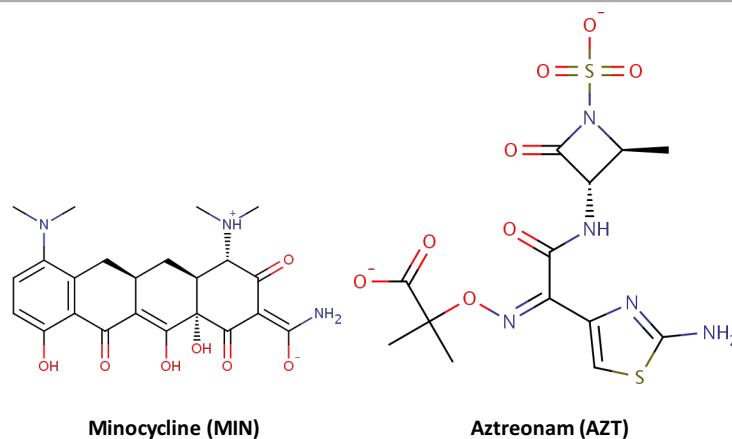
Figure 1. *The two-dimensional structures of the antibiotics used in the current study*

Table 1. *Physicochemical properties of the selected antibiotics used for molecular docking study on AcrB and AcrD. (Taken from antimicrobial compounds database [328])*

Property		Minocycline (MIN)	Aztreonam (AZT)
Molecular weight		457.5	433.4
van der Waals volume (Å^3)		403.7	331.6
logP (XlogP3)		1.8	0.3
Polar surface area (Å^2)		211.5	226.0
Physiological Charge		0	-2
H-bond acceptors		9	11
H-bond donors		5	2
Substrate	AcrB	Yes	Yes (poor)
	AcrD	No	Yes (better than AcrB)

Even though, a full- or semi-flexible docking in which flexibility is attributed to either both ligand and protein or just the ligand, should be the best option [219, 222, 329], we employed an alternative approach that has been previously tested and found to be reliable, and extensively used by our group. In this approach of rigid ensemble (cross-) docking, an ensemble of conformations of the ligand and receptor generated externally by MD simulations are docked against each other [221] with AutoDock VINA [220]. By adopting this protocol, possible unphysical configurations assumed by the ligand during flexible docking can be avoided. The results obtained provided better

insights into the ligand-binding modes adopted by MIN and AZT in AcrB and AcrD, in addition to a few characteristic drug-protein interactions that distinguish the preferred binding site of these compounds.

5.2. Methods

5.2.1. Receptor preparation

The ensemble of receptor structures consisted a total of 50 structures each for AcrB and AcrD extracted from clustering of ~6 μs-long MD simulation of their apo-form (as discussed in Chapter 2). Also, since apo-form extract from MD mostly have the pockets in a collapsed state (as seen in Chapter 3), additionally 10 crystal structures of AcrB (PDB IDs: 2DHH [102], 2GIF [103], 2J8S [330], 3W9H [277], 4DX5 [111], 4DX7 [111], 4U8V [127], 4U8Y [127], 4U95 [127], 4U96 [127]; resolution <3Å) were added to the ensemble of AcrB to include different conformations of the receptor. In the case of AcrD, since no crystal structure was available, homology models were built based on each of the ten AcrB crystal structures and subsequently included in the ensemble of AcrD. Thus, the receptor ensemble had a total of 60 structures each of AcrB and AcrD.

5.2.2. Ligand preparation

The ensemble of ligand structures consisted of cluster representatives of MIN and AZT taken from the antimicrobial compounds database [328] (http://www.dsf.unica.it/translocation/db/). The clusters of these molecules were determined using a hierarchical agglomerative algorithm from their μs-long MD simulations in explicit solvent conditions. In total the ensemble of ligand included two structures of minocycline and three of aztreonam.

5.2.3. Molecular Docking

Ensemble dockings of minocycline and aztreonam on AcrB and AcrD were performed using AutoDock Vina [220]. AutoDock Tools [331] was utilized to prepare the input *pdbqt* file for receptor (AcrB and AcrD) and ligand (MIN and AZT) structures. The size of the grid box was set to 125.0 × 125.0 × 110.0 Å and was centered such that the grid enclosed the whole protein for carrying out blind docking runs [332, 333]. The exhaustiveness parameter, which quantifies the accuracy of the docking search, was set to 8192 (the default value is 8), and the maximum number of binding modes to output was assumed to be 100. The predicted binding affinity (kcal/mol), which is a measure of how strongly a ligand binds to the receptor, is calculated based on the scoring function used in AutoDock Vina. A more negative binding affinity indicates stronger binding. The reliability of affinity scores predicted by Vina was tested by comparing its correlation to RMSD

with respect to the co-crystal MIN and also with binding free energy (ΔG) calculated using MM-GBSA method [252, 334].

5.2.4. Clustering of docked poses

All the docked poses of a ligand were initially categorized to identify poses within the putative binding pockets (AP of *L* monomer and DP of *T* monomer) and were then subjected to RMSD based clustering (cutoff set to 3 Å) using c*pptraj* module [217] of AmberTools14. Clustering was done to obtain the most populated and preferred binding modes per ligand to be analyzed for their interaction pattern with AcrB and AcrD proteins.

Various programs like ligPlot+ [335], PLIP [336] and VMD [271] were used for the analysis of docking results.

5.3. Results

We focus here on the docking results related to poses identified within AP of the *L* monomer and DP of the *T* monomer, which are the two main putative binding pockets in RND transporters suggested to be involved in the recognition and selectivity of different types of substrate molecules (details of the structural and functional aspects of RND transporters are discussed in Chapter 1 of this thesis).

5.3.1. Validation of the docking setup

The first step in molecular docking is to determine the reliability of the molecular docking strategy for our case. We validated our docking setup by docking MIN into an ensemble of AcrB crystal structures used here and evaluated it against the best resolution AcrB-MIN complex available (PDB ID: 4DX5 [111]). Among the poses of MIN identified within the DP of *T* monomer in AcrB ensemble, 10 poses had a root-mean-square-deviation (RMSD) below 2 Å and 16 poses in total showed an RMSD below 4 Å from the binding mode of the co-crystallized MIN. The lowest RMSD observed were 0.91 Å and 0.96 Å (Figure 2A) with Vina predicted binding affinities of -10.7 and -11.1 kcal/mol, respectively, and were ranked 12th and 4th, respectively, among the poses identified within the DP (Figure 2B). Note that the poses with better affinity than those mentioned above are not fully inside the DP. The affinity scores of all the poses showing an RMSD below 3 Å from the co-crystallized MIN ranged from -10.2 to -11.1 kcal/mol.
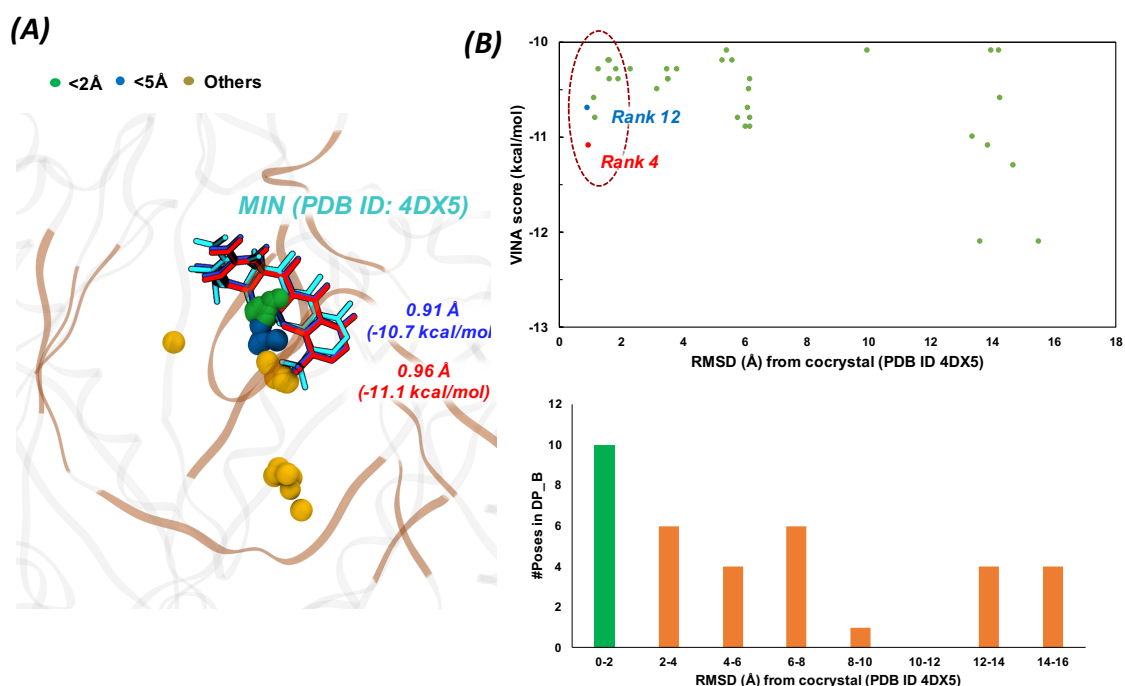
Figure 2. *Docking results of minocycline (MIN) into AcrB by blind docking using Vina. (A) Superposition of docked poses with the co-crystallized conformation of MIN (cyan colored licorice). The lowest RMSD docked pose of MIN is shown in blue licorice while the second in red. Also shown are their predicted Vina affinity values within parenthesis. Other poses identified within the DP of T monomer are shown as beads and colored differently based on their RMSD from the co-crystallized MIN [PDB ID: 4DX5] (green for <2 Å; blue for <5 Å; yellow for all other poses). The DP region is shown in brown ribbon. (B) Upper panel shows the scatter plot of Vina affinity scores (kcal/mol) against RMSD while the lower panel shows the distribution of docked poses within the DP of T monomer with respect to their RMSD. The column colored green corresponds to the poses represented as green beads in panel (A) while other columns are colored orange by default. All RMSDs reported here are with respect to the binding mode of the co-crystallized MIN [PDB ID: 4DX5]*

Further, the affinity scoring method of Vina was evaluated against the free energy of binding (ΔG) calculated using MM-GBSA implemented in AMBER14 [191] after running a short restrained optimization (with restraints on all heavy atoms) of the docked complexes obtained. The results showed a good and comparable correlation between Vina affinity scores as well as ΔG from MM-GBSA with the RMSD for docked poses falling within 3 Å from co-crystallized ligand (correlation coefficient of 0.71 for Vina score and 0.82 for ΔG). These results seem to reinforce the idea that AutoDock Vina predictions are reliable for our system. Hence, we continue using Vina affinity scores rather than the more expensive free energy of binding method for this study.

5.3.2. Access pocket

Having validated our docking setup with the crystal structures above, we discuss here the results obtained from the entire ensemble of 60 receptor structures. In the AP of *L* monomer in both AcrB and AcrD, we found several binding poses (shown as beads in Figure 3) for MIN and AZT

dispersed all over the pocket space. In the case of MIN, both AcrB and AcrD showed majority of the poses in the score range of -10 kcal/mol to -11 kcal/mol (Figure 4). The average affinity for AcrB-MIN binding modes was -10.6 (±0.04) kcal/mol while for AcrD it was -10.4 (±0.03) kcal/mol.

AZT, however, showed a lower affinity than minocycline in the AP of both AcrB and AcrD with an average affinity of -9.2 (±0.04) kcal/mol for AcrB and -9.2 (±0.03) for AcrD.
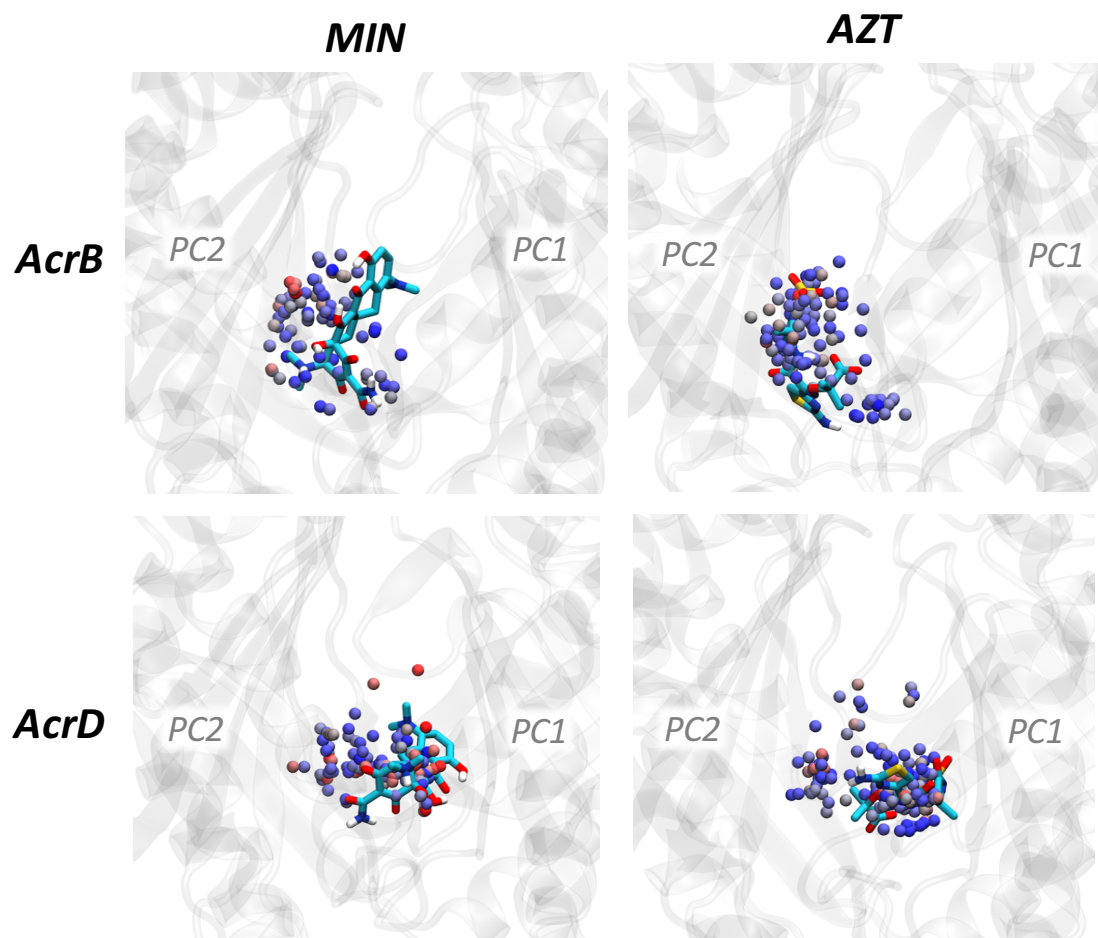
Figure 3. The binding mode (as viewed from the periplasmic entrance) of the most populated cluster pose (shown as licorice), predicted for MIN and AZT in AP of L monomer in AcrB and AcrD as predicted from blind docking with Vina. Also shown are the different binding sites of MIN and AZT as beads marking the center of mass of the ligand pose and colored according to their normalized Vina affinity scores (blue to red from lowest to highest affinity).
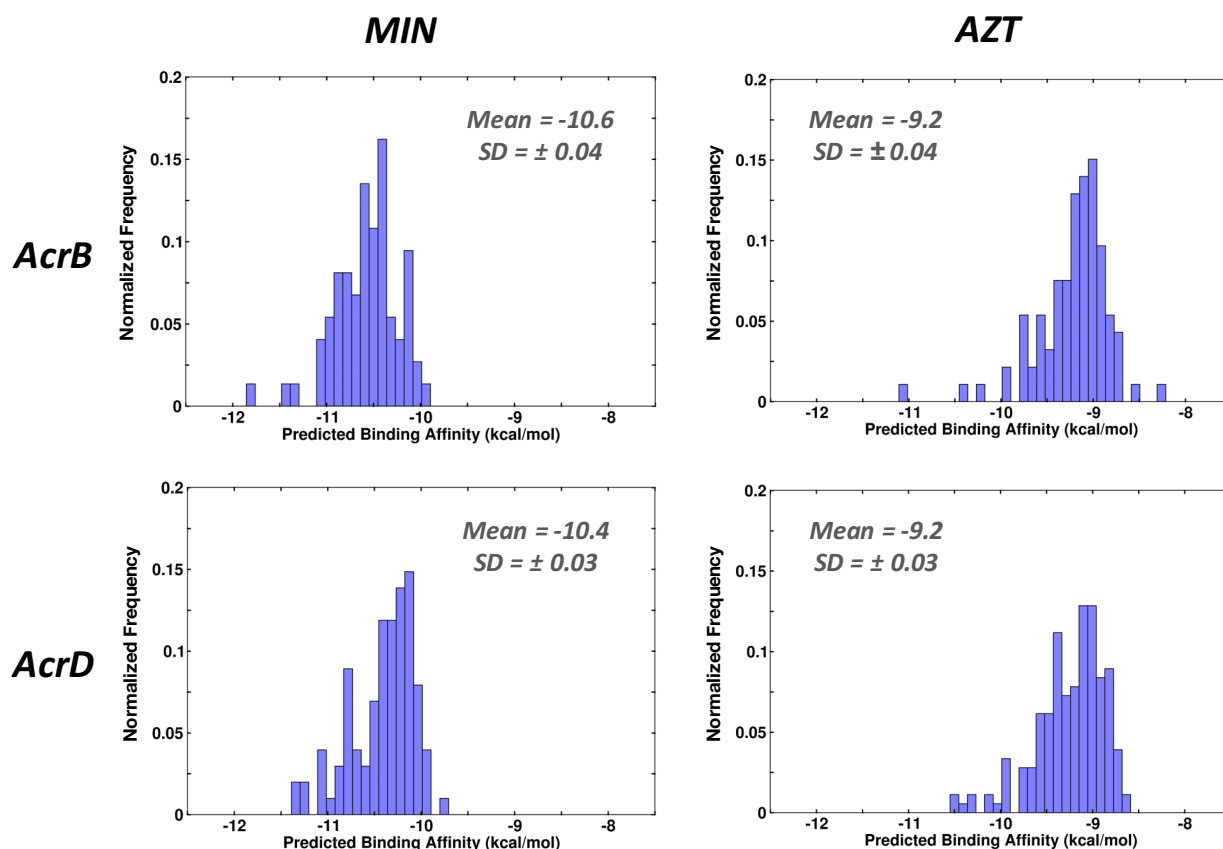
Figure 4. Histogram distribution of the binding affinities shown by MIN and AZT considering all poses within AP of L monomer in AcrB and AcrD.

We then analyzed the interactions made by the most populated cluster representative obtained from the clustering of docked poses in the AP of *L* monomer to determine the type of residues and the major type of interaction shown by these molecules (Table 2). Though the predicted Vina affinity scores of the corresponding ligands in AcrB and AcrD were quite similar and the AP in the two transporters shares an identity of 60% and similarity of 75%, the residues that characterized the binding were very different (Figure 5). For instance, the hydrophobic core of MIN showed hydrophobic interactions with F615 and F617 of the switch loop, which separates the AP from the DP [111] in AcrB, but was devoid of any such hydrophobic interactions in AcrD. However, in both proteins MIN showed several hydrogen bonds as expected from its major polar domain (containing the acidic OH, a dimethylamino group, and an amide group).

In the case of AZT, although the Vina affinity of the poses centered at essentially the same value, the distribution of poses inside AP of AcrB seems to be narrower than that of AcrD. The binding of AZT with its polar (sulphonic and carboxylic acid, amino) groups was characterized with H-bonds most of which came from polar residues of the pocket in addition to a salt bridge formed

by the sulfonic acid group of AZT with R625 which is also located at the interface between AP and DP.

Table 2. *Interactions shown by the cluster representative of the most populated cluster obtained from docked poses in the AP of L monomer in AcrB and AcrD. Interactions were calculated using PLIP [336].*

	MIN	AZT
AcrB	**Hydrophobic**: Phe615, Phe617 **H-bonds**: Ser134, Thr676, Ala677, Arg815, Arg815, Glu826, **Salt bridge:** Glu826 (with tertamine group of MIN)	**H-bonds**: Ser79, Asn81, Thr91, Thr93, Ser134, Glu673, Thr676, Ala677
AcrD	**H-bonds**: Thr572, Ser665, Ser676, Asp680, Asn717	**H-bonds**: Thr135, Asn136, Ser665, Ser675, Glu824 **Salt bridge**: Arg625 (with sulfonic acid group of AZT)

Figure 5. *Two-dimensional schematic plots showing the interactions between MIN and AZT with AP of L monomer in AcrB and AcrD. Plots were generated using LigPlot+ [335].*

5.3.3. Deep pocket

The ligand poses sampled in the DP of *T* monomer in AcrB and AcrD obtained from the entire ensemble of 60 receptor structures, were also similarly analyzed to determine the presence of any preferred binding modes for MIN and AZT. We found several binding modes for MIN and AZT (shown as beads in Figure 6), even in DP of AcrB and AcrD, but were lesser dispersed than the poses in AP. Several of these poses, especially the most populated cluster, was found close to the binding mode exhibited by the co-crystallized MIN (PDB ID: 4DX5). The average affinity for AcrB-MIN binding modes in the DP was -10.5 (±0.04) kcal/mol and for AcrD-MIN it was -10.5 (±0.03) kcal/mol. AZT showed a lower affinity than MIN in the DP of AcrB and AcrD with an average value of -9.1 (±0.03) kcal/mol in both the proteins (Figure 7).
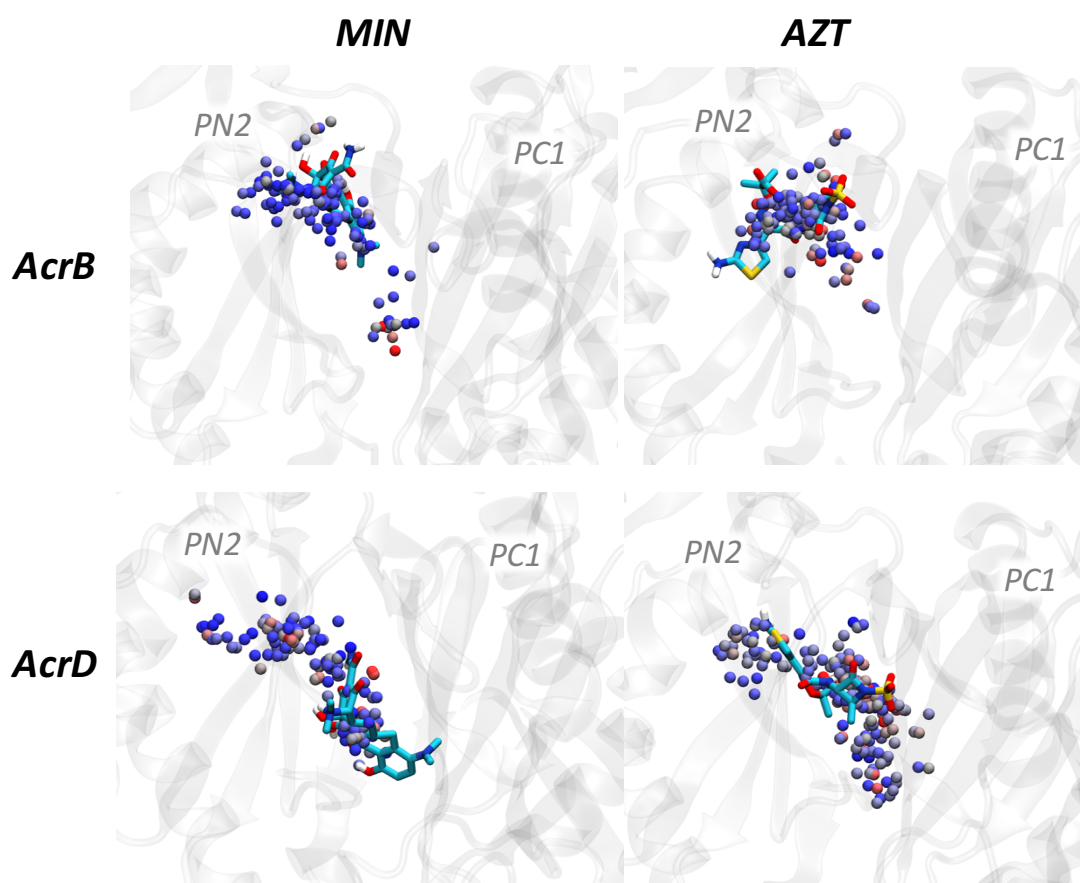
Figure 6. *The binding mode (as viewed from the PC2-PN1 side) of the most populated cluster pose (shown as licorice), predicted for MIN and AZT in DP of T monomer in AcrB and AcrD as predicted from blind docking with Vina. Also shown are the different binding sites of MIN and AZT as beads marking the center of mass of the ligand pose and colored according to their normalized Vina affinity scores (blue to red from lowest to highest affinity).*

The top populated clusters of docked poses of MIN in the DP of *T* monomer of AcrB were found in the groove of the DP making hydrophobic interactions using the hydrophobic core of MIN. The polar domain (acidic OH, a dimethylamino group, and an amide group) of MIN was in almost all cases oriented towards the wide translocation channel lined with numerous polar groups. However, in the case of AcrD, the MIN poses were not as buried in the groove as in AcrB and were found to be closer to the translocation channel.

AZT sampled poses away from the groove region of DP and closer to the exit of the translocation channel in AcrB. The poses in AcrB were also found to be restricted in the pocket space sampled. In contrast to this, AZT was found to prefer a greater region of the DP in AcrD with its polar groups finding complementarity due to the less hydrophobic and more polar DP of AcrD (as seen in Chapter 3).
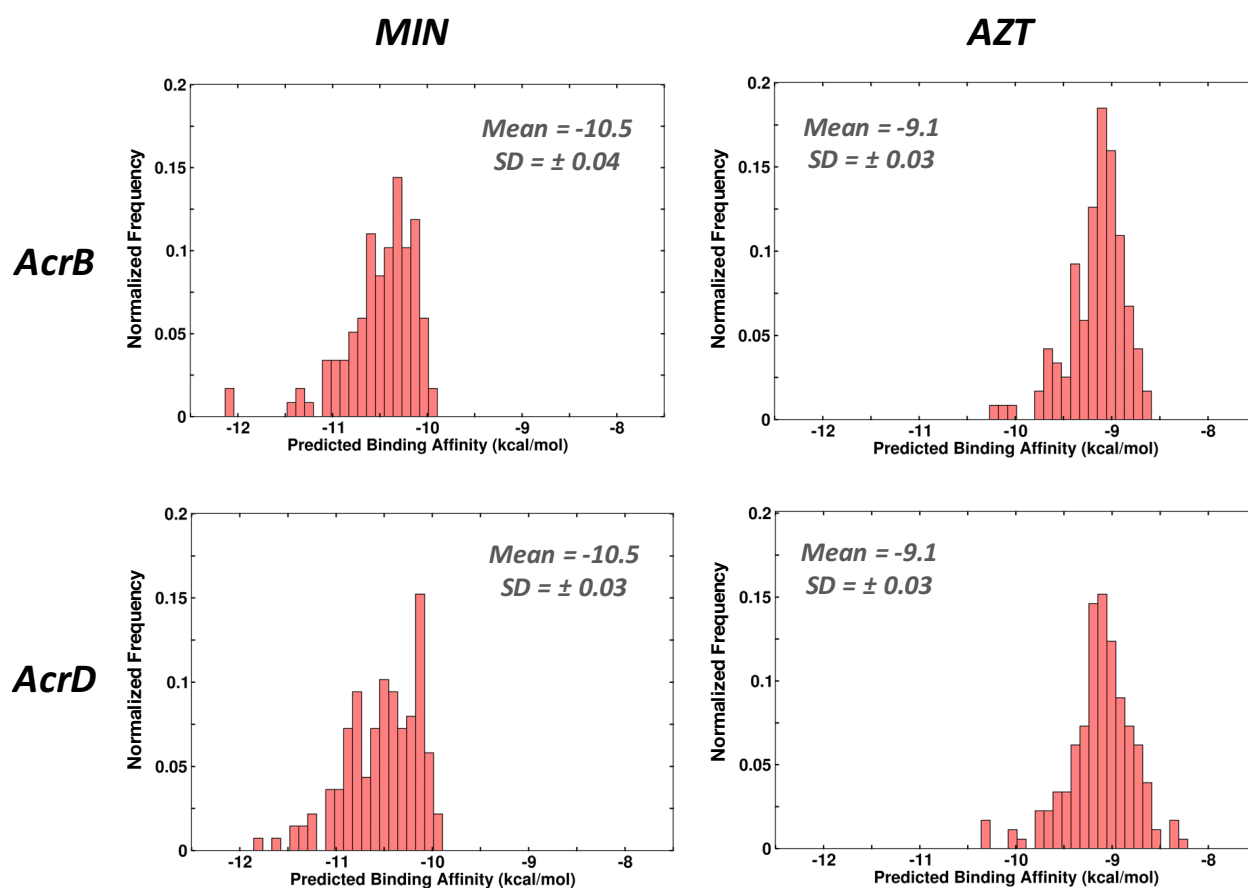
Figure 7. Histogram distribution of the binding affinities shown by MIN and AZT considering all poses within DP of T monomer in AcrB and AcrD.

The interactions shown by the most populated cluster representative obtained from the clustering of docked poses in the DP included hydrophobic interactions for MIN in both AcrB and AcrD though with different residues (Table 3). AZT showed mainly H-bonds with residues at

corresponding locations of AcrB and AcrD viz. Ser46, Gln89 in AcrB and Thr46, Ser89 in AcrD. A salt bridge formed by AZT using its sulfonic acid group was also identified in both proteins (Figure 8).

Table 3. Interactions shown by the cluster representative of the most populated cluster obtained from docked poses in the DP of T monomer in AcrB and AcrD. Calculated using PLIP [336].

	MIN	AZT
AcrB	**Hydrophobic**: Phe178, Ile277, Ala279, Val612 **H-bonds**: Gly179, Arg620,	**H-bonds**: Ser46, Ser48, Gln89, Ser128, Lys163, Glu273 **Salt bridge**: ARG620 (with Sulfonic acid group of AZT)
AcrD	**Hydrophobic**: Pro667 **H-bonds**: Thr135, Asn136, Lys292, Tyr327, Ser614 **pi-stacking**: Phe627	**H-bonds**: Thr46, Ser89, Ser614, Asn619, Arg625 **Salt bridge**: Arg44 (with Sulfonic acid group of AZT)
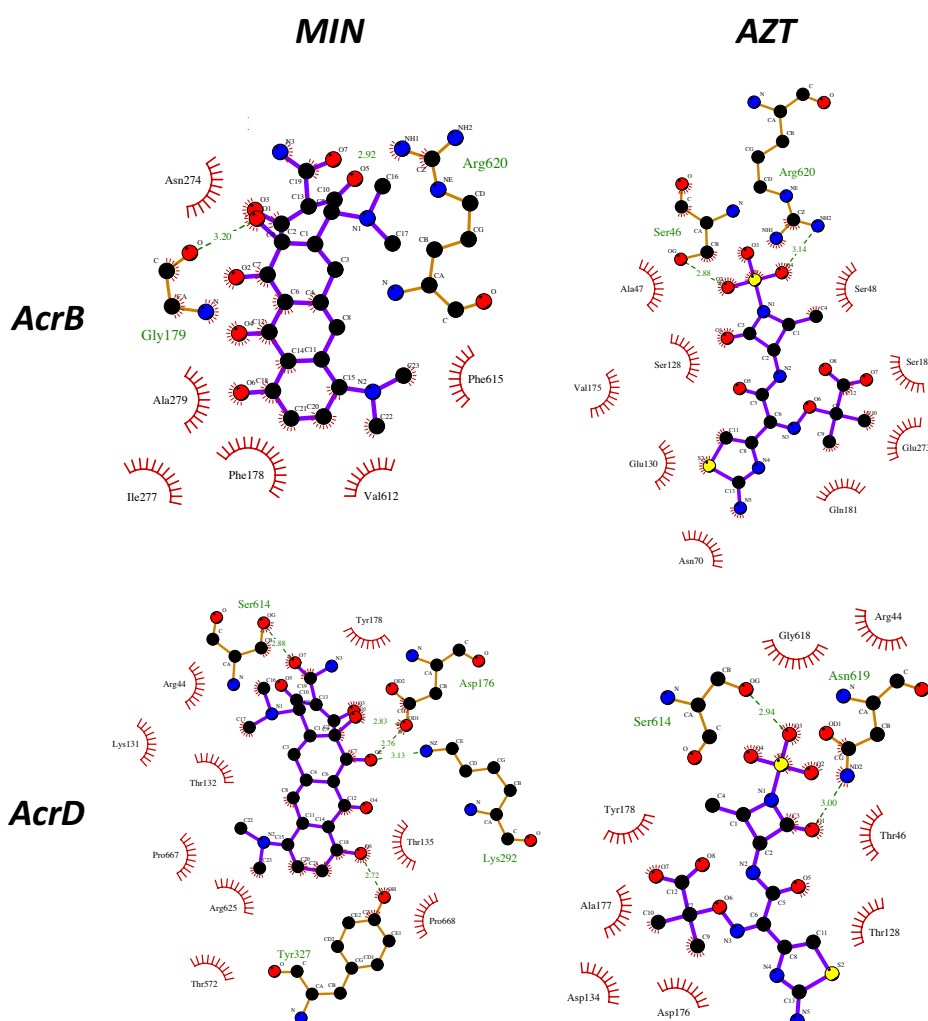
Figure 8. Two-dimensional schematic plots showing the interactions between MIN and AZT with DP of T monomer in AcrB and AcrD. Plots were generated using LigPlot+ [335].

5.4. Discussion

The substrate specificity of RND transporters is known to be extremely diverse, especially considering the notorious MDR transporter AcrB of *E. coli*. However, these transporters are also known to exhibit some preference in the substrate type recognized and transported by them. For instance, AcrB prefers more hydrophobic molecules while AcrD prefers more hydrophilic. In addition, AcrD is found to be a better transporter of anionic β-lactams like aztreonam than AcrB. In spite of the extensive information available from numerous studies on these transporters, the concept of differential substrate selectivity shown by these pumps remains largely unresolved. Based on available crystal structures of AcrB, the binding modes of only a few ligands like minocycline, doxorubicin, erythromycin, rifampicin have been identified in either one of the two putative binding pockets. The binding mode and interaction features of these ligands in the other pocket and for numerous other ligands in either of the binding pockets remain unknown. Hence, here we attempted to predict the binding mode and interaction pattern of MIN and AZT using molecular docking.

We first evaluated the applicability of the docking program, Autodock Vina, for our system and found its reliability in reproducing the binding mode of MIN as well as in predicting the affinity score for the docked poses on par with high-level approaches like the MM-GBSA for poses < 3Å RMSD from co-crystallized MIN (PDB ID: 4DX5). However, a direct correlation between the predicted affinity values by Vina and the transport behavior of these proteins should not be drawn and we need to be careful in interpreting the results here. Having said this, the use of these values for comparative examination of binding modes of different ligands can be informative.

There does not appear to be a large difference between the preferred binding of MIN and AZT, at least from the predicted binding affinity results of the poses inside the large voluminous AP, in both AcrB and AcrD. The AP of *L* monomer is very large (mean volume of 2515±438 Å^3 for AcrB and 3015±385 Å^3 for AcrD from MD) with dispersed chemical features like polar, charged and hydrophobic residues lining it (as seen in Chapter 3). This property of AP was proposed by us, based on the results discussed in Chapter 3, and also others [323, 337] to allow the binding of potential substrates with multiple binding modes. Our molecular docking results support this hypothesis by showing the numerous possible binding modes at different sites with little differences in binding affinities for MIN and AZT in the AP of *L* monomer in AcrB and AcrD. However, certain indispensable features are required for the binding of these ligands. For instance, numerous polar groups inside the pocket provide H-bonding complementarity for these molecules, which is an essential interaction shown by both MIN and AZT in their binding with both the Acr transporters. Additionally, in the case of MIN, hydrophobic residues may favor the accommodation of its

hydrophobic core partly flanked by polar groups. Likewise, the presence of a positively charged residue like Arg or Lys may favorably strengthen the recognition of AZT or similar molecules possessing negatively charged groups like sulphonic and carboxylic acid groups.

In the case of DP, due to the presence of a well-defined groove-like binding pocket (still being voluminous, 2610±250 Å^3 in AcrB and 2770±306 Å^3 in AcrD from MD), a more organized or in other terms a less diffused-type of binding is seen. The presence of numerous hydrophobic Phe groups in the DP of AcrB offers MIN a hydrophobic complementarity for its core and at the same time the polar groups at the translocation channel side sufficiently complement the polar domains of MIN. The absence of such hydrophobic complementarity, as seen in the DP of AcrD, might not completely hinder the binding of MIN but may result in a more peripheral binding by satisfying its polar interactions alone. In the case of AZT, however, in terms of chemical complementarity such a hydrophobic feature is unfavorable for accommodating the charged domains of the molecule. Thus, AZT prefers the channel more in AcrB than the groove unlike in AcrD, but the reason for it being a poor substrate of AcrB than AcrD is unclear from the docking results obtained. Perhaps MD simulations could explain this behavior of AZT in AcrB by rationalizing on the grounds of the existent dynamics between the transporter and its substrate.

In addition to gaining a better understanding of the differences in the binding modes and interacting residues of AZT and MIN in AcrB and AcrD, our study also identified R625 in AcrD which is located in proximity of the DP to be involved in making charged contacts with AZT in nearly 69% of the poses identified in the AP of *L* monomer in AcrD. This residue was recently identified by Kobayashi *et al.* [108] as possibly the most important residue for the recognition of anionic β-lactams based on mutation studies to transfer AcrD substrate specificity to AcrB. This finding, validates both our docking approach as well as the hypotheses drawn in Chapter 3 regarding the correlations prevailing between the physicochemical properties of the binding pockets and the substrate nature. Nevertheless, we would not rule out the possibility of limitations in our approach and are currently improving the predictions and scoring by using better ligand parameters and other docking programs like HADDOCK [338, 339] where protein flexibility and water are considered in such predictions.

Chapter 6

Choice of A Better Ligand-Parameterization Approach for HADDOCK

6.1. Introduction

In our earlier docking approach aimed at gaining a better understanding of the binding modes and interaction pattern of selected ligands in the putative binding pockets of AcrB and AcrD, we observed that the binding poses were scored unsatisfactorily given the presence of significant differences in their chemistry and susceptibility to the efflux pumps. This could be partly due to the lack of full protein flexibility during docking [329, 331, 340, 341], the lack of explicit waters [209, 224, 342] and the fact that the scoring functions are derived mostly from the complexes whose binding pockets exclude water [220]. These approximations can have a large impact on searching and on scoring particularly with MDR efflux systems like the RND transporters, known to have extremely flexible binding pockets [209, 211, 321] and presumably water-filled channels [209, 226, 343]. To address these limitations, we started working on HADDOCK [338, 339], a top ranked docking program based on CASP (Critical Assessment of Structure Prediction)/CAPRI (Critical Assessment of PRedicted Interactions) [344-347].

HADDOCK is being widely used for protein-protein docking but its potential in protein-ligand (small-molecule) docking remains largely unexplored. Its multi-stage docking protocol involves an initial rigid body docking, followed by simulated annealing of semi-flexible refinement in torsion angle space and a final explicit water refinement, which could in principle address, at least partly, the limitations discussed above. In addition, its scoring function combining OPLS [348] intermolecular van der Waals and Coulomb electrostatics energies, an empirically-derived desolvation energy term [349], and one or more restraints energy terms reflecting the agreement between the generated models and experimental/prediction information, has also been proven to be robust in ranking and selecting the best models [344].

In collaboration with the developers of HADDOCK we assessed the performance of the program in protein-ligand docking, identifying an optimal protocol suitable for our docking calculations to be used for RND transporters. One of the primary concerns to be addressed in this regard was related to the ligand input file. By default, the parameters for small molecules in

HADDOCK are taken from PRODRG [350, 351], which is based on the concept of united-atoms force field with parameters for all atoms except non-polar hydrogens. The major advantage of such force fields is the gain in computational speed due to reduced number of atoms being considered in the calculations. However, PRODRG is accompanied by 2 major limitations which could be of particular concern in the type of molecules we study. One of these is related to the way it assigns atomic charges and the other is related to the exclusion of non-polar hydrogens. PRODRG works on the concept of charge groups for assigning partial charges. According to this, an entire group (viz. -COO) must be first recognized as a charge group to assign atomic charges. This recognition and assigning of charges is known to be a major weakness especially for ring systems with nitrogens, and several charge group definitions (e.g. C-Cl) thereby allowing for a probable source of error. Not considering all hydrogens is accompanied with a major drawback of missing quadrupolar charge distribution in aromatic rings which hold an effective positive charge near the hydrogens and an effective negative charge nearer to the middle of the ring [352, 353]. This is undesirable for studying the drug-receptor interactions such as those stabilizing binding between AcrB and its antibiotic substrates featuring aromatic rings (for instance minocycline is clearly stabilized by weak aromatic stacking interactions [223, 354]) and charge groups (aztreonam has charge groups). Hence, we first proceeded by evaluating the impact of using an all-atom force field based on General AMBER Force Field (GAFF) [355] parameters with charges derived from QM calculations against the default united-atom force field based PRODRG approach.

6.2. Method

To use a different set of parameters than the default in HADDOCK, they must be converted into CNS format. We performed such conversion for the GAFF derived parameters of minocycline (MIN) (used as test case here), available in the antimicrobial compounds database developed in our group [328] and featuring quantum derived (B3LYP/6-31G**) atomic partial charges. These parameters (hereafter named AMBER-like parameters) were converted to CNS format using ACPYPE [356] and were implemented into the development version of HADDOCK webserver for our testing purpose.

Ab-initio surface-based docking runs of MIN structures (namely cluster representatives taken from the antimicrobial compounds database) with PRODRG derived parameters and AMBER-like parameters were performed on an ensemble of 10 crystal structures of AcrB (PDB IDs: 2DHH [102], 2GIF [103], 2J8S [330], 3W9H [277], 4DX5 [111], 4DX7 [111], 4U8V [127], 4U8Y [127], 4U95 [127], 4U96 [127]; resolution <3Å) using the settings recommended for small-

molecule docking with HADDOCK [357]. The *ab-initio* docking run (comparable to blind docking) using surface restraints randomly selected from all accessible residues was adopted here to sample the entire surface of the receptor and try to identify the putative binding sites and subsequently the ligand binding modes in the receptor. An overview of the settings used in the docking runs is tabulated below (Table 1).

A total of 50000 docked poses were generated per run in the first stage (it0) of which 2000 were taken to the subsequent stages (it1 and water refinement).

Table 1. Overview of the settings/parameters used in the HADDOCK docking calculation.

Receptor		AcrB crystal structures (10) (Resolution <= 3Å) [PDB IDs: 4U8V, 4U8Y, 4U95, 4U96, 2GIF, 2DHH, 2J8S, 3W9H, 4DX5, 4DX7]	
Ligand		MIN cluster representatives (c0, c1)	
Sampling		Random sampling of the entire surface of the receptor	
Parameterization		PRODRG	AMBER-like
Total docked complexes generated	it0 (Rigid-body EM)	50000	
	it1 (Semi-flexible refinement)	2000	
	water (Explicit solvent refinement)	2000	
Reference complex AcrB-MIN		4DX5 (1.9 Å)	

Scoring function: The HADDOCK scores (in arbitrary units) are defined by the equations below,

it0: score $= 0.01\ E_{air} + 0.01\ E_{vdW} + 1.0\ E_{elec} + 1.0\ E_{desolv} - 0.01\ BSA$

it1: score $= 0.1\ E_{air} + 1.0\ E_{vdW} + 1.0\ E_{elec} + 1.0\ E_{desolv} - 0.01\ BSA$

where, the energy terms E_{air} is for the ambiguous interaction restraint, E_{vdW} for the van der Waals, E_{elec} for the electrostatic and E_{desolv} is the desolvation while *BSA* corresponds to the buried surface area.

6.3. Results

In the rigid body EM stage (it0) of HADDOCK, docked poses are initially generated and subsequently taken to the refinement stages based on their HADDOCK score. Hence, it would be interesting to assess the initial sampling stage (it0) and subsequent semi-flexible refinement (it1) stage to determine the differences in the nature of the poses generated and their scores arising from

the different parameters used. Also, as the co-crystallized MIN (PDB ID: 4DX5) is found in the DP of the *T* monomer of AcrB (structural aspects of RND transporters have been discussed in Chapter 1) which would be our reference structure in this study, all the comparative analysis between the behavior of the different parameter-based approaches would be restricted to the DP region alone considering it0 and it1 stages.

To examine for the differences in the overall sampling of poses generated in it0 stage, we performed a statistical contact analysis to identify all residues of the pocket making contacts with the ligand poses. In terms of the residues making contacts with the ligand, as seen from the molecular surface shown in Figure 1, no significant differences were observed in the DP. Since this was only a qualitative analysis, we examined the distribution of poses sampled in the DP with respect to their RMSD from the co-crystallized MIN. This revealed a greater difference between the poses sampled in the different approaches (Figure 2).
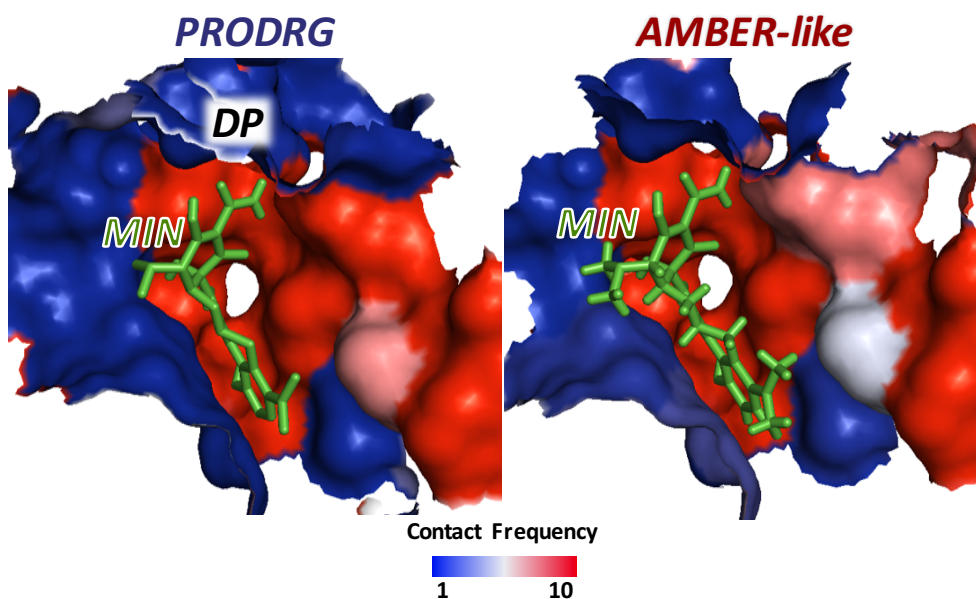
Figure 1. The surface of the DP in T monomer of AcrB colored according to the contact frequency (blue to red from 1 to 10) comparing the results of contact statistics from ab-initio surface based docking (blind docking) using PRODRG and AMBER-like parameters for MIN with HADDOCK. Also shown is the co-crystallized MIN in the reference structure (PDB ID: 4DX5).
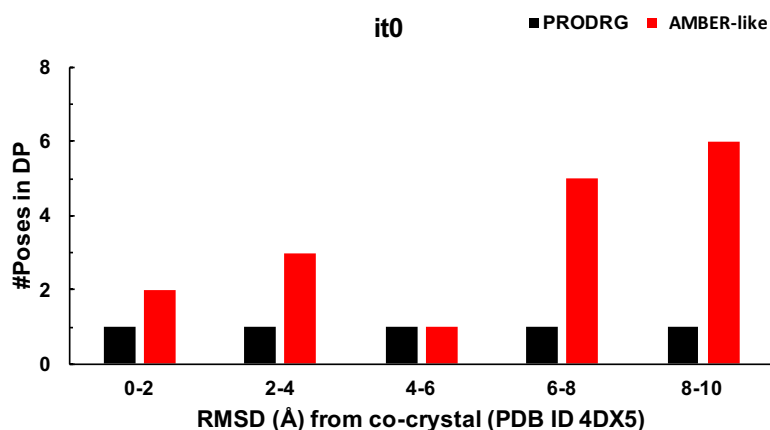
Figure 2. *Histogram distribution of the binding poses obtained from it0 stage for MIN in the DP of T monomer of AcrB with respect to their RMSD from the co-crystallized MIN in the reference structure (PDB ID: 4DX5). The black columns in the histogram correspond to PRODRG-based results while the red columns correspond to the AMBER-like parameters-based results.*

The HADDOCK scores consider contributions from electrostatics and van der Waals components, both of which differ between PRODRG and AMBER-like parameters used in our case. Hence, the poses though similar in their binding mode should be scored differently based on the difference in their components. On analyzing the difference in the HADDOCK scores we found that PRODRG gave better scores than the AMBER-like parameters in the it0 stage but not in it1. After refinement in it1 stage, both the poses with AMBER-like parameters which were within 2 Å from the co-crystallized MIN scored better (-51.5 and -51.7 a.u.) than the single pose obtained after refinement in PRODRG approach (-44.7 a.u.). The other poses though ranked in the same range were not fully inside DP and hence are not relevant here.
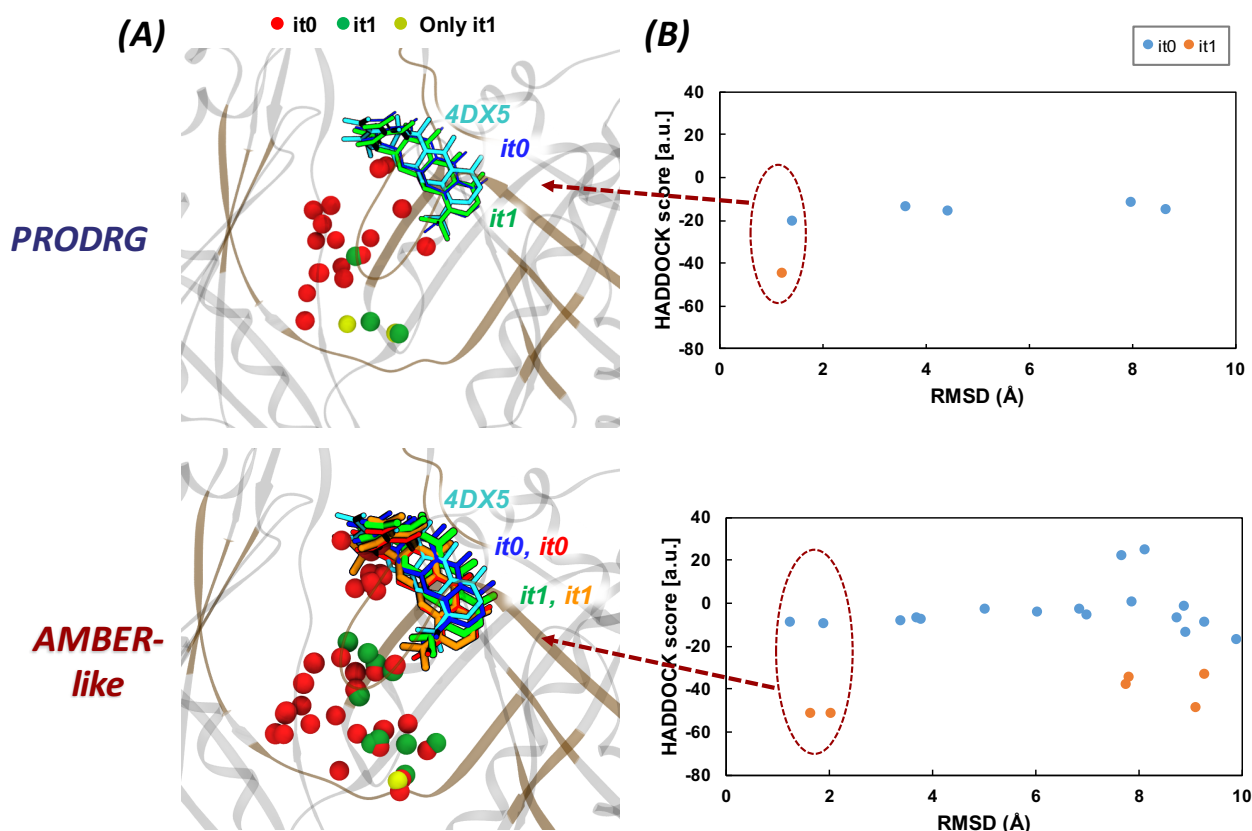
Figure 3*. Comparison of the binding modes and HADDOCK scores (in arbitrary units, a.u.) between PRODRG and AMBER-like parameters. **(A)** Binding mode of the ligands within 2 Å of co-crystallized MIN (cyan) as licorice while representing all other binding poses as beads colored differently based on the stage they were identified in (it0 in red, it1 in green, only it1 in yellow). The lowest RMSD pose from the it0 and its corresponding it1 pose are shown in blue and green licorice, respectively, while the second lowest pose (in the case of AMBER-like parameters in the lower panel) is colored red and orange for it0 and it1, respectively. **(B)** HADDOCK scores (in arbitrary units, a.u.) plotted against RMSD from the co-crystallized MIN. The dots are colored blue for it0 and orange for it1 poses. Also, encircled are the dots corresponding to the poses shown in in panel (A).*

Lastly, the fraction of native contacts conserved in the lowest RMSD poses obtained from both approaches was calculated, and the best poses from both approaches showed the same level of conservation (Table 2). The common hydrophobic interactions shown by poses obtained from PRODRG and AMBER-like approaches differed. For instance, Phe178 found to show hydrophobic interaction with MIN was found in both the poses from AMBER-like approach but not in that from PRODRG, where Ala279 was found instead. Similarly, the residues forming H-bonds were quite similar between the poses from AMBER-like but differed from that of PRODRG.

Table 2. Fraction of native contacts conserved in the lowest RMSD poses obtained from PRODRG and AMBER-like parameters are shown along with the various hydrophobic and H-bond interactions made by them. The corresponding interactions observed in the reference crystal structure are also shown. The residues showing common interactions as the reference are underlined. Calculated using PLIP with maximum threshold for the contacts identified set to 4 Å [336].

Structure	Hydrophobic interactions	H-bonds	Fraction of total native contacts
4DX5-MIN	Phe178, Ala279	Ser48, Gln151, Gly179, Asp276, Ile277, Arg620	Reference
PRODRG pose	Asn274, Ile277, Ala279, Phe610, Val612	Gln151, Ser180, Asn274, Asp276, Ile277, Ala286, Arg620	0.625
AMBER-like pose - 1	Phe178, Asn274, Ile277	Gly179, Asn274, Asp276, Ile277, Arg620	0.625
AMBER-like pose - 2	Ile277, Phe610; Phe178 (pi-Stacking)	Asn274, Asp276, Ile277, Arg620	0.5

6.4. Discussion

Both approaches involving the use of PRODRG and AMBER-like parameters for MIN were successful in producing near native crystal poses with equal conservation in the fraction of native contacts, though the interacting residues differed. However, a difference in the contribution from the additional hydrogens and the atomic partial charges on the atoms in the ligand with AMBER-like parameters appeared to score better after refinement than the PRODRG parameterized ligand. This may eventually help in a better prediction and better ranking of poses. We are currently optimizing the individual terms (like the E_{vdw} and E_{elec}) in the scoring function for a much better predictive power and applicability towards AcrB-like RND systems.

Conclusions and Future Perspectives

The main objective of this thesis was to provide better insights into the substrate-transporter complementarity underlying the recognition and transport events in the principal RND transporters of *E. coli* (AcrB, AcrD) and *P. aeruginosa* (MexB, MexY). In order to address this, we had to first build homology models of AcrD and MexY, which were meticulously validated and suitably refined to obtain a quality comparable to that of their crystallographic template structures. Since these transporter proteins are not static in biological samples, we performed μs-long MD simulations to obtain data related to their dynamic nature in apo-form. Subsequently, we chose a set of important descriptors like pocket volume, molecular lipophilic potential, electrostatic potential and hydration to analyze the putative binding sites in these transporters. The analysis of these parameters at the putative binding sites of these RND transporters allowed to correlate their different substrate specificity patterns to the physicochemical as well as topographical properties calculated on (or projected onto) the molecular surface of their multifunctional recognition sites. Our results suggested that the interactions of ligands with and their affinity to these transporters are resulting from the interplay between properties, such as volume, electrostatic potential, and specific features, such as the position of the base loop in AP of AcrB, and all together tuned by the dynamics of the systems. Thus, our work contributed to furnish important indications on the molecular determinants behind the partially different substrate specificity of the two couples of proteins in *E. coli* and *P. aeruginosa*. Such a rationalization in terms of molecular descriptors constitutes useful information to be exploited in new drug design attempts. To our knowledge this is the first attempt to insert all these descriptors in the setup of microscopically well-funded picture of these transporters aiming at a detailed characterization of their different specificity.

The results from molecular docking study with AutoDock Vina underlined the importance of certain key interaction types needed for a substrate to bind to its transporter and/or for a transporter to recognize its substrate. In addition, it complemented the findings from the descriptor based characterization of the putative binding pockets in these transporters. Important for this docking study was the availability of several structures of the receptors extracted from our extensive MD simulations to be combined, when available, with crystal structures. Additionally, we included dynamical properties of the ligands by using the conformations resulting from long MD simulations of the ligands in water boxes. These large sets of ligand and receptor configurations will be exploited more deeply in the future to analyze, from the statistical point of view, the results of the docking runs, in terms of frequency of contacts and distribution of the poses. Thus, the docking procedure will acquire, in our opinion, a further value beside the prediction of single poses. Surely, a future perspective will be also the validation of the identified poses by MD simulations,

including enhanced sampling techniques (for example, to mimic capture events and functional rotation), especially to gain insights into the transport properties of AcrB, AcrD, MexB, and MexY.

As a further research objective, it would be valuable to use the information gained from this study for docking of ligands with these transporters using information-driven programs like HADDOCK to obtain much better understanding of the interaction pattern between substrates and their transporter. In collaboration with the developers of HADDOCK, we are currently testing the impact of different parametrizations and settings on the scoring function. To make a more accurate assessment thereon more ligands will be considered in the future.

Appendix
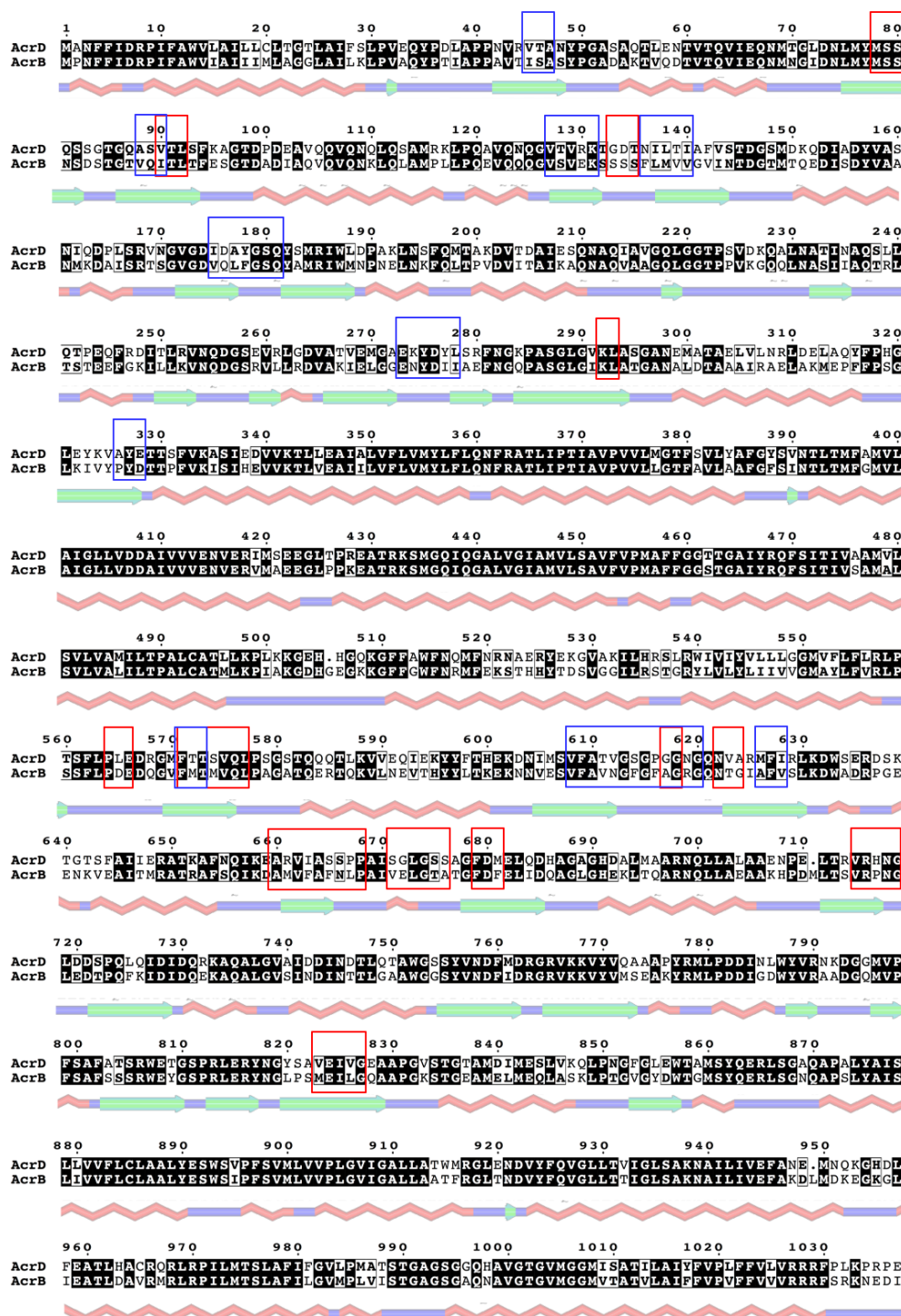
Figure 1. *Sequence alignment of AcrD with AcrB. The regions corresponding to access pocket (AP) and deep pocket (DP) are marked in red and blue boxes, respectively. The secondary structure of the sequence is also shown. Identical residues are highlighted with black filled boxes, similar with color less boxes while all others are mismatches*
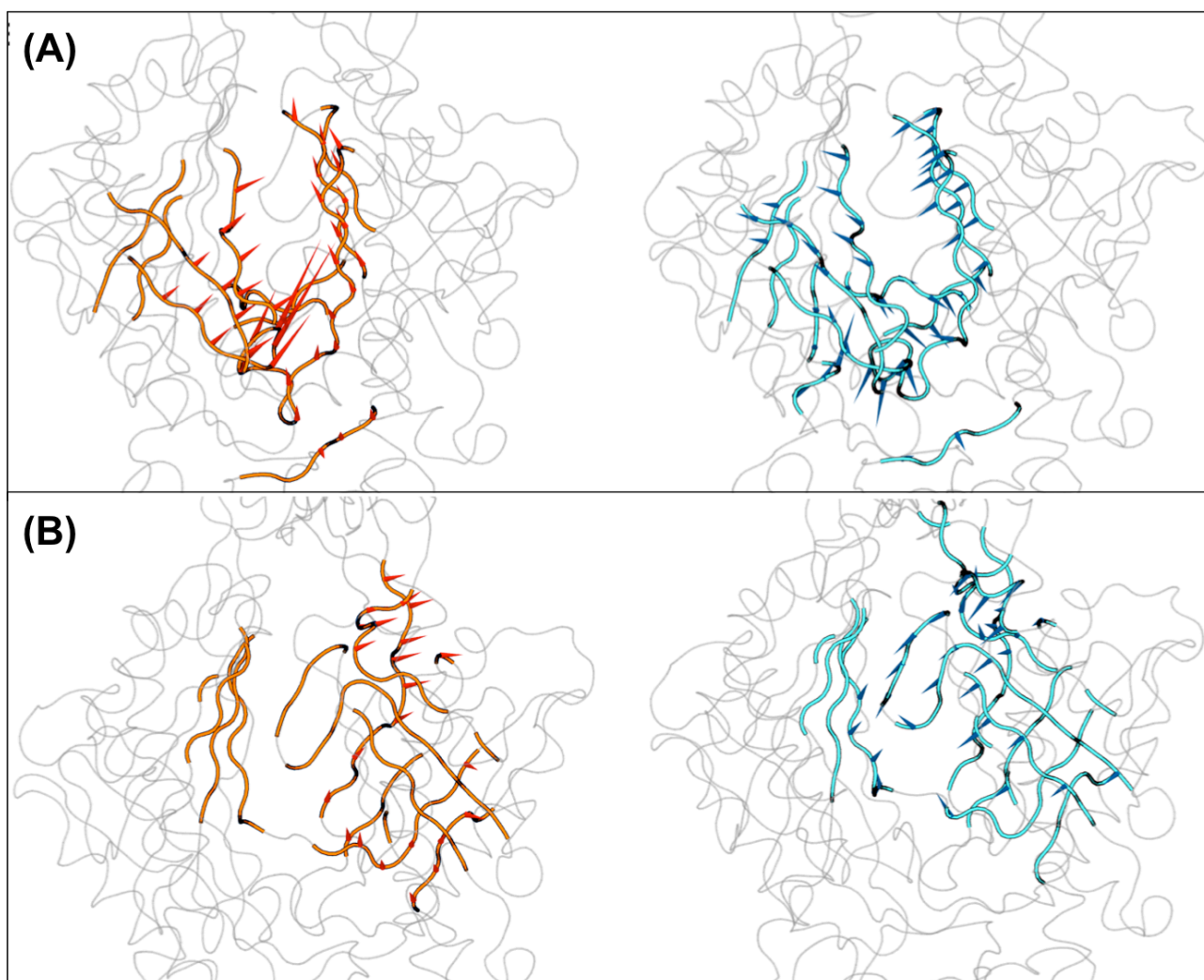
Figure 2. *Porcupine plots of the first eigenvector for (A) AP in the L monomer and (B) DP in the T monomer of (left panels) AcrB and (right panels) AcrD simulations shown as arrows (>2Å) attached to the backbone Cα atoms indicating the magnitude of the corresponding eigenvalues*
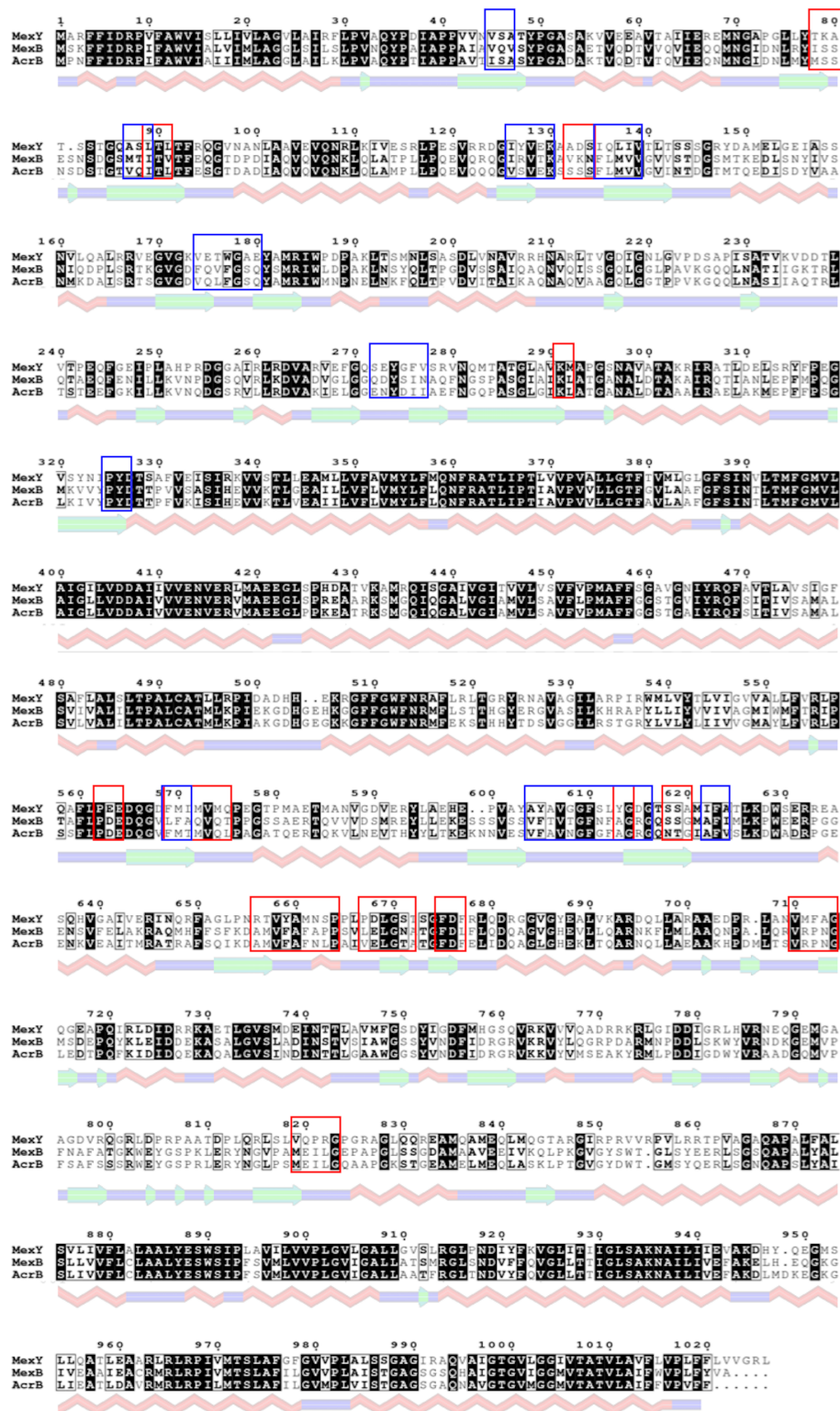
Figure 3. Sequence alignment of MexY with MexB and AcrB. The regions corresponding to access pocket (AP) and deep pocket (DP) are marked in red and blue boxes, respectively. The secondary structure of the sequence is also shown. Identical residues are highlighted with black filled boxes, similar with color less boxes while all others are mismatches

Significant Achievements

PUBLICATIONS

- **Ramaswamy, V. K.**, Cacciotto, P., Malloci, G., Ruggerone, P., & Vargiu, A. V. "Computational modelling of efflux pumps and their inhibitors". (Accepted in Essays in Biochemistry)
- Cacciotto, P., **Ramaswamy, V. K.**, Malloci, G., Ruggerone, P., & Vargiu, A. V. "Molecular modeling of multi-drug properties of RND transporters". Bacterial Multidrug Exporters of the series Methods in Molecular Biology, Ed. A. Yamaguchi (Springer Verlag, Heidelberg, New York) *(In Press)*
- **Ramaswamy, V. K.**, Cacciotto, P., Malloci, G., Ruggerone, P., & Vargiu, A. V. "Multidrug efflux pumps and their inhibitors characterized by computational modeling". In Efflux-Mediated Antimicrobial Resistance in Bacteria. Springer International Publishing, 797-831, 2016. (doi:10.1007/978-3-319-39658-3_30)

*Manuscripts on Chapters 3 and 4 ready to submit

ORAL PRESENTATIONS

- **Venkata Krishnan R.,** Vargiu, A. V. and Ruggerone, P., "Molecular rationale behind the differential substrate specificity of RND transporters AcrB and AcrD", *ITN Translocation Workshop*, Aix Marseille University, Marseille, France, 2016
- **Venkata Krishnan R.,** Vargiu, A. V. and Ruggerone, P., "Molecular rationale behind the differential substrate specificity of RND transporters AcrB and AcrD", *IMI Translocation Meeting*, Jacobs University, Bremen, Germany, 2016
- **Venkata Krishnan R.,** Vargiu, A. V. and Ruggerone, P., "Comparative study of RND efflux pumps: Insights from molecular docking and MD simulations", *Marie Curie ITN Training*, University of Cagliari, Cagliari, Italy, 2015
- **Venkata Krishnan R.,** Vargiu, A. V. and Ruggerone, P., "Computational study of MexY efflux pump: Insights from homology modeling and MD simulations", *Gordon Research Seminar on Multi-Drug Efflux Systems (GRS)*, Lucca (Barga), Italy, 2015
- **Venkata Krishnan R.,** Vargiu, A. V. and Ruggerone, P., "Molecular modeling of RND and related single-components efflux pumps", *Marie Curie ITN Training*, Jacobs University, Bremen, Germany, 2014

POSTER PRESENTATIONS

- **Venkata Krishnan R.,** Malloci, G., Vargiu, A. V. and Ruggerone, P., "Molecular rationale behind the differential substrate specificity of RND transporters AcrB and AcrD", *IMI Translocation Meeting*, Jacobs University, Bremen, Germany, 2016
- **Venkata Krishnan R.,** Malloci, G., Vargiu, A. V. and Ruggerone, P., "Comparative study of RND efflux pumps: Insights from MD simulations and binding pocket characterization", *60th Annual Meeting of the Biophysical Society*, Los Angeles, California, 2016
- **Venkata Krishnan R.,** Vargiu, A. V. and Ruggerone, P., "Comparative study of RND efflux pumps: Insights from molecular docking and MD simulations", *Marie Curie ITN Training*, University of Cagliari, Cagliari, Italy, 2015
- **Venkata Krishnan R.,** Vargiu, A. V. and Ruggerone, P., "Computational study of MexY efflux pump: Insights from homology modeling and MD simulations", *Gordon Research Conference on Multi-Drug Efflux Systems*, Lucca (Barga), Italy, 2015
- **Venkata Krishnan R.,** Vargiu, A. V. and Ruggerone, P., "Computational study of MexY efflux pump: Insights from homology modeling and MD simulations", *Gordon Research Seminar on Multi-Drug Efflux Systems*, Lucca (Barga), Italy, 2015
- **Venkata Krishnan R.,** Vargiu, A. V. and Ruggerone, P., "Computational study of MexY efflux pump: Insights from homology modeling and MD simulations", *Computational Molecular Science 2015*, The University of Warwick, Coventry, United Kingdom, 2015
- **Venkata Krishnan R.,** Vargiu, A. V. and Ruggerone, P., "Molecular modeling of RND and related single-components efflux pumps", *Marie Curie ITN Training and IMI Translocation Meeting*, Jacobs University, Bremen, Germany, 2014

Bibliography

1. Bassett, E.J., et al., *Tetracycline-labeled human bone from ancient Sudanese Nubia (AD 350)*. Science, 1980. **209**(4464): p. 1532-1534.
2. Nelson, M.L., et al., *Brief communication: mass spectroscopic characterization of tetracycline in the skeletal remains of an ancient population from Sudanese Nubia 350–550 CE*. Am J Phys Anthropol, 2010. **143**(1): p. 151-154.
3. Cook, M., E. Molto, and C. Anderson, *Fluorochrome labelling in Roman period skeletons from Dakhleh Oasis, Egypt*. Am J Phys Anthropol, 1989. **80**(2): p. 137-143.
4. Armelagos, G.J., *Disease in ancient Nubia*. Science, 1969. **163**(3864): p. 255-259.
5. Aminov, R.I., *A brief history of the antibiotic era: lessons learned and challenges for the future*. Front Microbiol, 2010. **1**: p. 134.
6. Aldridge, S. and J. Sturichio, *The Discovery and Development of Penicillin: 1928-1945*. 1999: London: American Chemical Society and Royal Society of Chemistry.
7. *American Chemical Society International Historic Chemical Landmarks. Discovery and Development of Penicillin.* [cited 2016 November 25]; Available from: http://www.acs.org/content/acs/en/education/whatischemistry/landmarks/flemingpenicillin.html
8. Walsh, C.T. and T.A. Wencewicz, *Prospects for new antibiotics: a molecule-centered perspective*. J Antibiot, 2014. **67**(1): p. 7-22.
9. Davies, J., *Where have all the antibiotics gone?* Can J Infect Dis Med Microbiol, 2006. **17**(5): p. 287-290.
10. Li, X.-Z., C.A. Elkins, and H.I. Zgurskaya, *Efflux-Mediated Antimicrobial Resistance in Bacteria*. 2016: Springer.
11. Davies, J. and D. Davies, *Origins and evolution of antibiotic resistance*. Microbiol Mol Biol Rev, 2010. **74**(3): p. 417-433.
12. Falagas, M.E. and D.E. Karageorgopoulos, *Pandrug resistance (PDR), extensive drug resistance (XDR), and multidrug resistance (MDR) among Gram-negative bacilli: need for international harmonization in terminology*. Clin Infect Dis, 2008. **46**(7): p. 1121-1122.
13. Magiorakos, A.P., et al., *Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance*. Clin Microbiol Infect, 2012. **18**(3): p. 268-281.
14. O'Neill, J., *Tackling drug-resistant infections globally: Final report and recommendations*. 2016, London: Wellcome Trust & HM Government
15. Blair, J.M., et al., *Molecular mechanisms of antibiotic resistance*. Nat Rev Microbiol, 2015. **13**(1): p. 42-51.
16. Dax, S.L., *Antibacterial chemotherapeutic agents*. 2012: Springer Science & Business Media.
17. George, A.M., *Multidrug resistance in enteric and other Gram-negative bacteria*. FEMS Microbiol Lett, 1996. **139**(1): p. 1-10.
18. Pagès, J.-M. and L. Amaral, *Mechanisms of drug efflux and strategies to combat them: challenging the efflux pump of Gram-negative bacteria*. Biochim Biophy Acta 2009. **1794**(5): p. 826-833.
19. Upadhyay, R., *Emergence of drug resistance in microbes, its dissemination and target modification of antibiotics: a life threatening problem to human society*. Int J Pharm Biol Res, 2011. **2**(5): p. 119-126.
20. Tanwar, J., et al., *Multidrug resistance: an emerging crisis*. Interdiscip Perspect Infect Dis, 2014. **2014**: p. 7.

21. Utsui, Y. and T. Yokota, *Role of an altered penicillin-binding protein in methicillin-and cephem-resistant Staphylococcus aureus.* Antimicrob Agents Chemother, 1985. **28**(3): p. 397-403.

22. Sherrard, L.J., M.M. Tunney, and J.S. Elborn, *Antimicrobial resistance in the respiratory microbiota of people with cystic fibrosis.* The Lancet, 2014. **384**(9944): p. 703-713.

23. Poole, K., *Efflux-mediated antimicrobial resistance.* J Antimicrob Chemother, 2005. **56**(1): p. 20-51.

24. Piddock, L.J., *Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria.* Clin Microbiol Rev, 2006. **19**(2): p. 382-402.

25. Nikaido, H. and J.M. Pagès, *Broad-specificity efflux pumps and their role in multidrug resistance of Gram-negative bacteria.* FEMS Microbiol Rev, 2012. **36**(2): p. 340-63.

26. Lomovskaya, O. and K.A. Bostian, *Practical applications and feasibility of efflux pump inhibitors in the clinic-a vision for applied use.* Biochem Pharmacol, 2006. **71**(7): p. 910-8.

27. Blair, J.M., G.E. Richmond, and L.J. Piddock, *Multidrug efflux pumps in Gram-negative bacteria and their role in antibiotic resistance.* Future Microbiol, 2014. **9**(10): p. 1165-1177.

28. Li, X.-Z. and H. Nikaido, *Efflux-mediated drug resistance in bacteria: an update.* Drugs, 2009. **69**(12): p. 1555-1623.

29. Nikaido, H., *Prevention of drug access to bacterial targets: permeability barriers and active efflux.* Science, 1994. **264**(5157): p. 382-8.

30. Piddock, L.J., *Multidrug-resistance efflux pumps - not just for resistance.* Nat Rev Microbiol, 2006. **4**(8): p. 629-36.

31. Rosner, J.L. and R.G. Martin, *An excretory function for the Escherichia coli outer membrane pore TolC: upregulation of marA and soxS transcription and Rob activity due to metabolites accumulated in tolC mutants.* J Bacteriol, 2009. **191**(16): p. 5283-92.

32. Paul, S., et al., *A single-component multidrug transporter of the major facilitator superfamily is part of a network that protects Escherichia coli from bile salt stress.* Mol Microbiol, 2014. **92**(4): p. 872-84.

33. Guelfo, J.R., et al., *A MATE-family efflux pump rescues the Escherichia coli 8-oxoguanine-repair-deficient mutator phenotype and protects against H_2O_2 killing.* PLoS Genet, 2010. **6**(5): p. e1000931.

34. Bogomolnaya, L.M., et al., *The ABC-type efflux pump MacAB protects Salmonella enterica serovar Typhimurium from oxidative stress.* mBio, 2013. **4**(6): p. e00630-13.

35. Poole, K., *Stress responses as determinants of antimicrobial resistance in Gram-negative bacteria.* Trends Microbiol, 2012. **20**(5): p. 227-34.

36. Li, X.-Z., P. Plésiat, and H. Nikaido, *The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria.* Clin Microbiol Rev, 2015. **28**(2): p. 337-418.

37. Podnecky, N.L., K.A. Rhodes, and H.P. Schweizer, *Efflux pump-mediated drug resistance in Burkholderia.* Front Microbiol, 2015. **6**: p. 305.

38. Baugh, S., et al., *Inhibition of multidrug efflux as a strategy to prevent biofilm formation.* J Antimicrob Chemother, 2014. **69**(3): p. 673-81.

39. Matsumura, K., et al., *Roles of multidrug efflux pumps on the biofilm formation of Escherichia coli K-12.* Biocontrol Sci, 2011. **16**(2): p. 69-72.

40. Saier, M.H., Jr., et al., *The transporter classification database.* Nucleic Acids Res, 2014. **42**: p. D251-8.

41. Ren, Q., K. Chen, and I.T. Paulsen, *TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels.* Nucleic Acids Res, 2007. **35**(suppl 1): p. D274-D279.

42. Eswaran, J., et al., *Three's company: component structures bring a closer view of tripartite drug efflux pumps.* Curr Opin Struct Biol, 2004. **14**(6): p. 741-7.

43.   Delmar, J.A. and E.W. Yu, *The AbgT family: A novel class of antimetabolite transporters.* Protein Sci, 2016. **25**(2): p. 322-337.

44.   Hassan, K.A., et al., *Homologs of the Acinetobacter baumannii AceI transporter represent a new family of bacterial multidrug efflux systems.* MBio, 2015. **6**(1): p. e01982-14.

45.   Du, D., et al., *Structure, mechanism and cooperation of bacterial multidrug transporters.* Curr Opin Struct Biol, 2015. **33**: p. 76-91.

46.   Higgins, C.F., *ABC transporters: physiology, structure and mechanism--an overview.* Res Microbiol, 2001. **152**(3-4): p. 205-10.

47.   Shapiro, A.B., et al., *Stimulation of P-glycoprotein-mediated drug transport by prazosin and progesterone.* Eur J Biochem, 1999. **259**(3): p. 841-850.

48.   Martin, C., et al., *Communication between multiple drug binding sites on P-glycoprotein.* Mol Pharmacol, 2000. **58**(3): p. 624-632.

49.   Higgins, C.F. and K.J. Linton, *The ATP switch model for ABC transporters.* Nat Struct Mol Biol, 2004. **11**(10): p. 918-926.

50.   Martin, C., et al., *Drug binding sites on P-glycoprotein are altered by ATP binding prior to nucleotide hydrolysis.* Biochemistry, 2000. **39**(39): p. 11901-11906.

51.   McDevitt, C.A., et al., *Is ATP binding responsible for initiating drug translocation by the multidrug transporter ABCG2?* FEBS J, 2008. **275**(17): p. 4354-4362.

52.   Yan, N., *Structural advances for the major facilitator superfamily (MFS) transporters.* Trends Biochem Sci, 2013. **38**(3): p. 151-9.

53.   Lee, A., et al., *Interplay between efflux pumps may provide either additive or multiplicative effects on drug resistance.* J Bacteriol, 2000. **182**(11): p. 3142-50.

54.   Tal, N. and S. Schuldiner, *A coordinated network of transporters with overlapping specificities provides a robust survival strategy.* Proc Natl Acad Sci U S A, 2009. **106**(22): p. 9051-6.

55.   Lewinson, O., et al., *Promiscuity in multidrug recognition and transport: the bacterial MFS Mdr transporters.* Mol Microbiol, 2006. **61**(2): p. 277-284.

56.   Fluman, N., et al., *Dissection of mechanistic principles of a secondary multidrug efflux protein.* Mol Cell, 2012. **47**(5): p. 777-787.

57.   Sapunaric, F.M., M. Aldema-Ramos, and L.M. McMurry, *Tetracycline resistance: efflux, mutation, and other mechanisms.*, in *Frontiers in Antimicrobial Resistance, a tribute to Stuart B. Levy*, D.G. White, M.N. Alekshun, and P.F. McDermot, Editors. 2005, ASM Press: Washington DC. p. 3-18.

58.   Nikaido, H. and H.I. Zgurskaya, *Antibiotic efflux mechanisms.* Curr Opin Infect Dis, 1999. **12**(6): p. 529-536.

59.   Kuroda, T. and T. Tsuchiya, *Multidrug efflux transporters in the MATE family.* Biochim Biophys Acta, 2009. **1794**(5): p. 763-8.

60.   He, X., et al., *Structure of a cation-bound multidrug and toxic compound extrusion transporter.* Nature, 2010. **467**(7318): p. 991-4.

61.   Lu, M., et al., *Structures of a Na^+-coupled, substrate-bound MATE multidrug transporter.* Proc Natl Acad Sci U S A, 2013. **110**(6): p. 2099-104.

62.   Tanaka, Y., et al., *Structural basis for the drug extrusion mechanism by a MATE multidrug transporter.* Nature, 2013. **496**(7444): p. 247-51.

63.   Paulsen, I.T., et al., *The SMR family: a novel family of multidrug efflux proteins involved with the efflux of lipophilic drugs.* Mol Microbiol, 1996. **19**(6): p. 1167-75.

64.   Schuldiner, S., *EmrE, a model for studying evolution and mechanism of ion-coupled transporters.* Biochim Biophys Acta, 2009. **1794**(5): p. 748-62.

65.   Pornillos, O., et al., *X-ray structure of the EmrE multidrug transporter in complex with a substrate.* Science, 2005. **310**(5756): p. 1950-1953.

66.   Chen, Y.J., et al., *X-ray structure of EmrE supports dual topology model.* Proc Natl Acad Sci U S A, 2007. **104**(48): p. 18999-9004.

67. Schuldiner, S., *Undecided membrane proteins insert in random topologies. Up, down and sideways: it does not really matter.* Trends Biochem Sci, 2012. **37**(6): p. 215-9.

68. Korkhov, V.M. and C.G. Tate, *Electron crystallography reveals plasticity within the drug binding site of the small multidrug transporter EmrE.* J Mol Biol, 2008. **377**(4): p. 1094-103.

69. Morrison, E.A., et al., *Antiparallel EmrE exports drugs by exchanging between asymmetric structures.* Nature, 2011. **481**(7379): p. 45-50.

70. Rotem, D. and S. Schuldiner, *EmrE, a multidrug transporter from Escherichia coli, transports monovalent and divalent substrates with the same stoichiometry.* J Biol Chem, 2004. **279**(47): p. 48787-93.

71. Korkhov, V.M. and C.G. Tate, *Electron crystallography reveals plasticity within the drug binding site of the small multidrug transporter EmrE.* J Mol Biol, 2008. **377**(4): p. 1094-1103.

72. Srinivasan, V.B., G. Rajamohan, and W.A. Gebreyes, *Role of AbeS, a novel efflux pump of the SMR family of transporters, in resistance to antimicrobial agents in Acinetobacter baumannii.* Antimicrob Agents Chemother, 2009. **53**(12): p. 5312-6.

73. Lytvynenko, I., et al., *Molecular basis of polyspecificity of the Small Multidrug Resistance Efflux Pump AbeS from Acinetobacter baumannii.* J Mol Biol, 2016. **428**(3): p. 644-657.

74. Srinivasan, V.B. and G. Rajamohan, *KpnEF, a new member of the Klebsiella pneumoniae cell envelope stress response regulon, is an SMR-type efflux pump involved in broad-spectrum antimicrobial resistance.* Antimicrob Agents Chemother, 2013. **57**(9): p. 4449-62.

75. Tseng, T.-T., et al., *The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins.* J Mol Microbiol Biotechnol, 1999. **1**(1): p. 107-125.

76. Venter, H., et al., *RND-type drug efflux pumps from Gram-negative bacteria: molecular mechanism and inhibition.* Front Microbiol, 2015. **6**: p. 377.

77. Saier, M., et al., *Two novel families of bacterial membrane proteins concerned with nodulation, cell division and transport.* Mol Microbiol, 1994. **11**(5): p. 841-847.

78. Nikaido, H., *Multidrug efflux pumps of Gram-negative bacteria.* J Bacteriol, 1996. **178**(20): p. 5853-9.

79. Dreier, J. and P. Ruggerone, *Interaction of antibacterial compounds with RND efflux pumps in Pseudomonas aeruginosa.* Front Microbiol, 2015. **6**: p. 660.

80. Li, X.-Z., L. Zhang, and H. Nikaido, *Efflux pump-mediated intrinsic drug resistance in Mycobacterium smegmatis.* Antimicrob Agents Chemother, 2004. **48**(7): p. 2415-2423.

81. Milano, A., et al., *Azole resistance in Mycobacterium tuberculosis is mediated by the MmpS5–MmpL5 efflux system.* Tuberculosis, 2009. **89**(1): p. 84-90.

82. Serizawa, M. and J. Sekiguchi, *The Bacillus subtilis YdfHI two-component system regulates the transcription of ydfJ, a member of the RND superfamily.* Microbiology, 2005. **151**(6): p. 1769-1778.

83. Quiblier, C., et al., *Contribution of SecDF to Staphylococcus aureus resistance and expression of virulence factors.* BMC Microbiol, 2011. **11**(1): p. 1.

84. Sulavik, M.C., et al., *Antibiotic susceptibility profiles of Escherichia coli strains lacking multidrug efflux pump genes.* Antimicrob Agents Chemother, 2001. **45**(4): p. 1126-36.

85. Lu, M., et al., *Structures of a Na+-coupled, substrate-bound MATE multidrug transporter.* Proc Natl Acad Sci U S A, 2013. **110**(6): p. 2099-2104.

86. Schuldiner, S., *Undecided membrane proteins insert in random topologies. Up, down and sideways: it does not really matter.* Trends Biochem Sci, 2012. **37**(6): p. 215-219.

87. Symmons, M.F., et al., *The assembled structure of a complete tripartite bacterial multidrug efflux pump.* Proc Natl Acad Sci U S A, 2009. **106**(17): p. 7173-8.

88. Piddock, L.J., *Multidrug-resistance efflux pumps - not just for resistance.* Nat Rev Microbiol, 2006. **4**(8): p. 629-636.

89. Lister, P.D., D.J. Wolter, and N.D. Hanson, *Antibacterial-resistant Pseudomonas aeruginosa: clinical impact and complex regulation of chromosomally encoded resistance mechanisms.* Clin Microbiol Rev, 2009. **22**(4): p. 582-610.

90. Aires, J.R., et al., *Involvement of an active efflux system in the natural resistance of Pseudomonas aeruginosa to aminoglycosides.* Antimicrob Agents Chemother, 1999. **43**(11): p. 2624-2628.

91. Masuda, N., et al., *Substrate specificities of MexAB-OprM, MexCD-OprJ, and MexXY-oprM efflux pumps in Pseudomonas aeruginosa.* Antimicrob Agents Chemother, 2000. **44**(12): p. 3322-3327.

92. Morita, Y., et al., *Roles of MexXY-and MexAB-multidrug efflux pumps in intrinsic multidrug resistance of Pseudomonas aeruginosa PAO1.* J Gen Appl Microbiol, 2001. **47**(1): p. 27-32.

93. Hobbs, E.C., et al., *Conserved small protein associates with the multidrug efflux pump AcrB and differentially affects antibiotic resistance.* Proc Natl Acad Sci U S A, 2012. **109**(41): p. 16696-701.

94. Du, D., et al., *Structure of the AcrAB-TolC multidrug efflux pump.* Nature, 2014. **509**(7501): p. 512-5.

95. Murakami, S., et al., *Crystal structure of bacterial multidrug efflux transporter AcrB.* Nature, 2002. **419**(6907): p. 587-93.

96. Sennhauser, G., et al., *Crystal structure of the multidrug exporter MexB from Pseudomonas aeruginosa.* J Mol Biol, 2009. **389**(1): p. 134-45.

97. Yin, Y., et al., *Structure of the multidrug transporter EmrD from Escherichia coli.* Science, 2006. **312**(5774): p. 741-744.

98. Chen, Y.-J., et al., *X-ray structure of EmrE supports dual topology model.* Proc Natl Acad Sci U S A, 2007. **104**(48): p. 18999-19004.

99. Dawson, R.J. and K.P. Locher, *Structure of a bacterial multidrug ABC transporter.* Nature, 2006. **443**(7108): p. 180-185.

100. Murakami, S., et al., *Crystal structure of bacterial multidrug efflux transporter AcrB.* Nature, 2002. **419**(6907): p. 587-593.

101. Ruggerone, P., et al., *RND efflux pumps: structural information translated into function and inhibition mechanisms.* Curr Top Med Chem, 2013. **13**(24): p. 3079-100.

102. Murakami, S., et al., *Crystal structures of a multidrug transporter reveal a functionally rotating mechanism.* Nature, 2006. **443**(7108): p. 173-179.

103. Seeger, M.A., et al., *Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism.* Science, 2006. **313**(5791): p. 1295-1298.

104. Seeger, M.A., et al., *The AcrB efflux pump: conformational cycling and peristalsis lead to multidrug resistance.* Curr Drug Targets, 2008. **9**(9): p. 729-49.

105. Murakami, S., *Multidrug efflux transporter, AcrB-the pumping mechanism.* Curr Opin Struct Biol, 2008. **18**(4): p. 459-65.

106. Nikaido, H. and J.-M. Pagès, *Broad-specificity efflux pumps and their role in multidrug resistance of Gram-negative bacteria.* FEMS Microbiol Rev, 2012. **36**(2): p. 340-363.

107. Nakashima, R., et al., *Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket.* Nature, 2011. **480**(7378): p. 565-9.

108. Kobayashi, N., et al., *β-Lactam selectivity of multidrug transporters AcrB and AcrD resides in the proximal binding pocket.* J Biol Chem, 2014. **289**(15): p. 10680-10690.

109. Nakashima, R., et al., *Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket.* Nature, 2011. **480**(7378): p. 565-569.

110. Murakami, S., *Structures and Transport Mechanisms of RND Efflux Pumps*, in *Efflux-Mediated Antimicrobial Resistance in Bacteria.* 2016, Springer. p. 3-28.

111. Eicher, T., et al., *Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop.* Proc Natl Acad Sci U S A, 2012. **109**(15): p. 5687-5692.

112. Yamaguchi, A., R. Nakashima, and K. Sakurai, *Structural basis of RND-type multidrug exporters.* Front Microbiol, 2015. **6**.

113. Lomovskaya, O. and M. Totrov, *Vacuuming the periplasm.* J Bacteriol, 2005. **187**(6): p. 1879-1883.

114. Aires, J.R. and H. Nikaido, *Aminoglycosides are captured from both periplasm and cytoplasm by the AcrD multidrug efflux transporter of Escherichia coli.* J Bacteriol, 2005. **187**(6): p. 1923-1929.

115. Yao, X.-Q., et al., *Drug export and allosteric coupling in a multidrug transporter revealed by molecular simulations.* Nat Commun, 2010. **1**: p. 117.

116. Yao, X.-Q., et al., *Drug uptake pathways of multidrug transporter AcrB studied by molecular simulations and site-directed mutagenesis experiments.* J Am Chem Soc, 2013. **135**(20): p. 7474-7485.

117. Eda, S., H. Maseda, and T. Nakae, *An elegant means of self-protection in gram-negative bacteria by recognizing and extruding xenobiotics from the periplasmic space.* J Biol Chem, 2003. **278**(4): p. 2085-2088.

118. Elkins, C.A. and H. Nikaido, *Substrate specificity of the RND-type multidrug efflux pumps AcrB and AcrD of Escherichia coli is determined predominately by two large periplasmic loops.* J Bacteriol, 2002. **184**(23): p. 6490-6498.

119. Mao, W., et al., *On the mechanism of substrate specificity by resistance nodulation division (RND)-type multidrug resistance pumps: the large periplasmic loops of MexD from Pseudomonas aeruginosa are involved in substrate recognition.* Mol Microbiol, 2002. **46**(3): p. 889-901.

120. Franke, S., et al., *Molecular analysis of the copper-transporting efflux system CusCFBA of Escherichia coli.* J Bacteriol, 2003. **185**(13): p. 3804-3812.

121. Middlemiss, J.K. and K. Poole, *Differential impact of MexB mutations on substrate selectivity of the MexAB-OprM multidrug efflux pump of Pseudomonas aeruginosa.* J Bacteriol, 2004. **186**(5): p. 1258-1269.

122. Murakami, S., et al., *Extramembrane central pore of multidrug exporter AcrB in Escherichia coli plays an important role in drug transport.* J Biol Chem, 2004. **279**(5): p. 3743-3748.

123. Tikhonova, E.B., Q. Wang, and H.I. Zgurskaya, *Chimeric analysis of the multicomponent multidrug efflux transporters from gram-negative bacteria.* J Bacteriol, 2002. **184**(23): p. 6499-6507.

124. Pos, K.M., *Drug transport mechanism of the AcrB efflux pump.* Biochim Biophys Acta, 2009. **1794**(5): p. 782-93.

125. Nikaido, H., et al., *Multidrug efflux pump AcrAB of Salmonella typhimurium excretes only those β-lactam antibiotics containing lipophilic side chains.* J Bacteriol, 1998. **180**(17): p. 4686-92.

126. Husain, F., M. Bikhchandani, and H. Nikaido, *Vestibules are part of the substrate path in the multidrug efflux transporter AcrB of Escherichia coli.* J Bacteriol, 2011. **193**(20): p. 5847-9.

127. Eicher, T., et al., *Coupling of remote alternating-access transport mechanisms for protons and substrates in the multidrug efflux pump AcrB.* Elife, 2014. **3**: p. e03145.

128. Rice, P., I. Longden, and A. Bleasby, *EMBOSS: the European molecular biology open software suite.* Trends Genet, 2000. **16**(6): p. 276-277.

129. Wehmeier, C., et al., *Site-directed mutagenesis reveals amino acid residues in the Escherichia coli RND efflux pump AcrB that confer macrolide resistance.* Antimicrob Agents Chemother, 2009. **53**(1): p. 329-330.

130. Rosenberg, E.Y., D. Ma, and H. Nikaido, *AcrD of Escherichia coli is an aminoglycoside efflux pump.* J Bacteriol, 2000. **182**(6): p. 1754-1756.

131. Morita, Y., J. Tomida, and Y. Kawamura, *MexXY multidrug efflux system of Pseudomonas aeruginosa.* Front Microbiol, 2012. **3**: p. 408.

132. Morita, Y., M.L. Sobel, and K. Poole, *Antibiotic inducibility of the MexXY multidrug efflux system of Pseudomonas aeruginosa: involvement of the antibiotic-inducible PA5471 gene product.* J Bacteriol, 2006. **188**(5): p. 1847-1855.

133. Ramaswamy, V.K., et al., *Multidrug Efflux Pumps and Their Inhibitors Characterized by Computational Modeling*, in *Efflux-Mediated Antimicrobial Resistance in Bacteria*. 2016, Springer. p. 797-831.

134. Kandt, C. and L. Monticelli, *Membrane protein dynamics from femtoseconds to seconds.* Methods Mol Biol, 2010. **654**: p. 423-440.

135. Sennhauser, G., et al., *Crystal structure of the multidrug exporter MexB from Pseudomonas aeruginosa.* J Mol Biol, 2009. **389**(1): p. 134-145.

136. Schwede, T. and M. Peitsch, *Computational structural biology: methods and applications.* 2008, Singapore: World Scientific Publishing Co. Pte. Ltd. .

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

138. Hvidsten, T.R., et al., *A comprehensive analysis of the structure-function relationship in proteins based on local structure similarity.* PLoS ONE, 2009. **4**(7): p. e6266.

139. Engh, R.A. and R. Huber, *Accurate bond and angle parameters for X-ray protein structure refinement.* Acta Crystallogr, Sect A: Found Crystallogr, 1991. **47**(4): p. 392-400.

140. Liapunov, A.M., *Stability of Motion by AM Liapunov*. 1 ed. Vol. 30. 2000: Elsevier. 1-322.

141. Garman, E.F., *Developments in x-ray crystallographic structure determination of biological macromolecules.* Science, 2014. **343**(6175): p. 1102-1108.

142. Drenth, J., *Principles of protein X-ray crystallography*. 2007: Springer Science & Business Media.

143. Van Ingen, H. and A.M. Bonvin, *Information-driven modeling of large macromolecular assemblies using NMR data.* J Magn Reson, 2014. **241**: p. 103-114.

144. Harris, S.A., et al., *Cooperativity in drug-DNA recognition: a molecular dynamics study.* J Am Chem Soc, 2001. **123**(50): p. 12658-12663.

145. Jackman, L.M. and S. Sternhell, *Application of Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry: International Series in Organic Chemistry*. 2013: Elsevier.

146. Orlova, E.V. and H.R. Saibil, *Structural analysis of macromolecular assemblies by electron microscopy.* Chem Rev, 2011. **111**(12): p. 7710-7748.

147. Goldstein, J., et al., *Scanning electron microscopy and X-ray microanalysis: a text for biologists, materials scientists, and geologists*. 2012: Springer Science & Business Media.

148. Hayat, M.A., *Principles and techniques of scanning electron microscopy. Biological applications. Volume 1*. 1974: Van Nostrand Reinhold Company.

149. Gordon, R., R. Bender, and G.T. Herman, *Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography.* J Theor Biol, 1970. **29**(3): p. 471-481.

150. Boeckmann, B., et al., *The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003.* Nucleic Acids Res, 2003. **31**(1): p. 365-370.

151. Westbrook, J., et al., *The protein data bank and structural genomics.* Nucleic Acids Res, 2003. **31**(1): p. 489-491.

152. Šali, A. and T.L. Blundell, *Comparative protein modelling by satisfaction of spatial restraints.* J Mol Biol, 1993. **234**(3): p. 779-815.

153. Bower, M.J., F.E. Cohen, and R.L. Dunbrack, *Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool.* J Mol Biol, 1997. **267**(5): p. 1268-1282.

154. Schwede, T., et al., *SWISS-MODEL: an automated protein homology-modeling server.* Nucleic Acids Res, 2003. **31**(13): p. 3381-3385.

155. Xu, D., Y. Xu, and C. Uberbacher, *Computational tools for protein modeling.* Curr Protein Pept Sci, 2000. **1**(1): p. 1-21.

156. Martí-Renom, M.A., et al., *Comparative protein structure modeling of genes and genomes.* Annu Rev Biophys Biomol Struct, 2000. **29**(1): p. 291-325.

157. Šali, A., et al., *Evaluation of comparative protein modeling by MODELLER.* Proteins: Struct, Funct, Bioinf, 1995. **23**(3): p. 318-326.

158. Webb, B. and A. Sali, *Comparative protein structure modeling using Modeller.* Curr Protoc Bioinformatics, 2014: p. 5.6. 1-5.6. 32.

159. Eswar, N., et al., *Comparative protein structure modeling using Modeller.* Curr Protoc Bioinformatics, 2006. **15**(5.6): p. 5.6.1-5.6.30.

160. Rost, B., *Twilight zone of protein sequence alignments.* Protein Eng, 1999. **12**(2): p. 85-94.

161. Madden, T., *The BLAST sequence analysis tool.* 2013: The NCBI Handbook [Internet]. 2nd edition. Bethesda (MD): National Center for Biotechnology Information (US).

162. Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.* Mol Syst Biol, 2011. **7**(1).

163. ŠAli, A. and J.P. Overington, *Derivation of rules for comparative protein modeling from a database of protein structure alignments.* Protein Sci, 1994. **3**(9): p. 1582-1596.

164. MacKerell Jr, A.D., et al., *All-atom empirical potential for molecular modeling and dynamics studies of proteins†.* J Phys Chem B, 1998. **102**(18): p. 3586-3616.

165. Wiederstein, M. and M.J. Sippl, *ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.* Nucleic Acids Res, 2007. **35**(suppl 2): p. W407-W410.

166. Sippl, M.J., *Recognition of errors in three-dimensional structures of proteins.* Proteins: Struct, Funct, Bioinf, 1993. **17**(4): p. 355-362.

167. Eisenberg, D., R. Lüthy, and J.U. Bowie, *VERIFY3D: assessment of protein models with three-dimensional profiles.* Methods Enzymol, 1997. **277**: p. 396.

168. Hooft, R., et al., *Errors in protein structures.* Nature, 1996. **381**(6580): p. 272-272.

169. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures.* J Appl Crystallogr, 1993. **26**(2): p. 283-291.

170. Galeazzi, R., *Molecular dynamics as a tool in rational drug design: current status and some major applications.* Curr Comput Aided Drug Des, 2009. **5**(4): p. 225-240.

171. Ruggerone, P., et al., *Molecular dynamics computer simulations of multidrug RND efflux pumps.* Comput Struct Biotechnol J, 2013. **5**: p. e201302008.

172. Schlitter, J., M. Engels, and P. Krüger, *Targeted molecular dynamics: a new approach for searching pathways of conformational transitions.* J Mol Graph, 1994. **12**(2): p. 84-89.

173. Izvekov, S. and G.A. Voth, *A multiscale coarse-graining method for biomolecular systems.* J Phys Chem B, 2005. **109**(7): p. 2469-2473.

174. Takada, S., *Coarse-grained molecular simulations of large biomolecules.* Curr Opin Struct Biol, 2012. **22**(2): p. 130-137.

175. Parkin, J., M. Chavent, and S. Khalid, *Molecular simulations of Gram-negative bacterial membranes: a vignette of some recent successes.* Biophys J, 2015. **109**(3): p. 461-468.

176. Grossfield, A., S.E. Feller, and M.C. Pitman, *Convergence of molecular dynamics simulations of membrane proteins.* Proteins: Struct, Funct, Bioinf, 2007. **67**(1): p. 31-40.

177. Frenkel, D. and B. Smit, *Understanding molecular simulation: from algorithms to applications.* Computational Sciences Series. Vol. 1. 2002: Elsevier. 1-638.

178. Lebowitz, J., J. Percus, and L. Verlet, *Ensemble dependence of fluctuations with application to machine computations.* Phys Rev, 1967. **153**(1): p. 250.

179. Nosé, S., *A molecular dynamics method for simulations in the canonical ensemble.* Mol Phys, 1984. **52**(2): p. 255-268.

180. Berendsen, H.J., et al., *Molecular dynamics with coupling to an external bath.* J Chem Phys, 1984. **81**(8): p. 3684-3690.

181. Nosé, S., *A unified formulation of the constant temperature molecular dynamics methods.* J Chem Phys, 1984. **81**(1): p. 511-519.

182. Hoover, W.G., *Canonical dynamics: equilibrium phase-space distributions.* Phys Rev A, 1985. **31**(3): p. 1695.

183. Andersen, H.C., *Molecular dynamics simulations at constant pressure and/or temperature.* J Chem Phys, 1980. **72**(4): p. 2384-2393.

184. Verlet, L., *Computer" experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules.* Phys Rev, 1967. **159**(1): p. 98.

185. Hockney, R.W., *Potential Calculation and Some Applications.* Methods in Computational Physics. Vol. 9. 1970: Academic Press, Inc., .

186. Ryckaert, J.-P., G. Ciccotti, and H.J. Berendsen, *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes.* J Comput Phys, 1977. **23**(3): p. 327-341.

187. Tuckerman, M.E. and G.J. Martyna, *Understanding modern molecular dynamics: techniques and applications.* J Phys Chem B, 2000. **104**(2): p. 159-178.

188. Tuckerman, M.E., et al., *Non-Hamiltonian molecular dynamics: Generalizing Hamiltonian phase space principles to non-Hamiltonian systems.* J Chem Phys, 2001. **115**(4): p. 1678-1702.

189. Cornell, W.D., et al., *A second generation force field for the simulation of proteins, nucleic acids, and organic molecules.* J Am Chem Soc, 1995. **117**(19): p. 5179-5197.

190. Pearlman, D.A., et al., *AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules.* Comput Phys Commun, 1995. **91**(1): p. 1-41.

191. Case, D., et al., *AMBER 14.* 2014: University of California, San Francisco.

192. Dickson, C.J., et al., *Lipid14: the amber lipid force field.* J Chem Theory Comput, 2014. **10**(2): p. 865-879.

193. Brooks, B.R., et al., *CHARMM: the biomolecular simulation program.* J Comput Chem, 2009. **30**(10): p. 1545-1614.

194. Brooks, B.R., et al., *CHARMM: a program for macromolecular energy, minimization, and dynamics calculations.* J Comput Chem, 1983. **4**(2): p. 187-217.

195. Van Gunsteren, W. and H. Berendsen, *Groningen molecular simulation (GROMOS) library manual.* Biomos, Groningen, 1987. **24**(682704): p. 13.

196. Soares, T.A., et al., *An improved nucleic acid parameter set for the GROMOS force field.* J Comput Chem, 2005. **26**(7): p. 725-737.

197. Christen, M., et al., *The GROMOS software for biomolecular simulation: GROMOS05.* J Comput Chem, 2005. **26**(16): p. 1719-1751.

198. Marrink, S.J. and D.P. Tieleman, *Perspective on the Martini model.* Chem Soc Rev, 2013. **42**(16): p. 6801-6822.

199. Hansson, T., C. Oostenbrink, and W. van Gunsteren, *Molecular dynamics simulations.* Curr Opin Struct Biol, 2002. **12**(2): p. 190-196.

200. Steinhauser, M.O. and S. Hiermaier, *A review of computational methods in materials science: examples from shock-wave and polymer physics.* Int J Mol Sci, 2009. **10**(12): p. 5135-5216.

201. Steinbach, P.J. and B.R. Brooks, *New spherical-cutoff methods for long-range forces in macromolecular simulation.* J Comput Chem, 1994. **15**(7): p. 667-683.

202. Eastwood, J. and R. Hockney, *Shaping the force law in two-dimensional particle-mesh models.* J Comput Phys, 1974. **16**(4): p. 342-359.

203. Darden, T., D. York, and L. Pedersen, *Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems.* J Chem Phys, 1993. **98**(12): p. 10089-10092.

204. Cheatham, T.I., et al., *Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins.* J Am Chem Soc, 1995. **117**(14): p. 4193-4194.

205. Deserno, M. and C. Holm, *How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines.* J Chem Phys, 1998. **109**(18): p. 7678-7693.

206. Hockney, R.W. and J.W. Eastwood, *Computer simulation using particles.* 1988: CRC Press.

207. Pollock, E. and J. Glosli, *Comments on P 3 M, FMM, and the Ewald method for large periodic Coulombic systems.* Comput Phys Commun, 1996. **95**(2): p. 93-110.

208. Schulz, R., et al., *Functional rotation of the transporter AcrB: insights into drug extrusion from simulations.* PLoS Comput Biol, 2010. **6**(6): p. e1000806.

209. Schulz, R., et al., *Role of water during the extrusion of substrates by the efflux transporter AcrB.* J Phys Chem B, 2011. **115**(25): p. 8278-8287.

210. Schulz, R., et al., *Computational study of correlated domain motions in the AcrB efflux transporter.* BioMed Res Int, 2015. **2015**(2015): p. 12.

211. Vargiu, A.V., et al., *Effect of the F610A mutation on substrate extrusion in the AcrB transporter: explanation and rationale by molecular dynamics simulations.* J Am Chem Soc, 2011. **133**(28): p. 10704-10707.

212. Lomize, M.A., et al., *OPM database and PPM web server: resources for positioning of proteins in membranes.* Nucleic Acids Res, 2012. **40**(D1): p. D370-D376.

213. Jo, S., et al., *CHARMM-GUI: a web-based graphical user interface for CHARMM.* J Comput Chem, 2008. **29**(11): p. 1859-1865.

214. Maier, J.A., et al., *ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB.* J Chem Theory Comput, 2015. **11**(8): p. 3696-3713.

215. Pastor, R.W., B.R. Brooks, and A. Szabo, *An analysis of the accuracy of Langevin and molecular dynamics algorithms.* Mol Phys, 1988. **65**(6): p. 1409-1419.

216. Hopkins, C.W., et al., *Long-time-step molecular dynamics through hydrogen mass repartitioning.* J Chem Theory Comput, 2015. **11**(4): p. 1864-1874.

217. Roe, D.R. and T.E. Cheatham III, *PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data.* J Chem Theory Comput, 2013. **9**(7): p. 3084-3095.

218. Turner, P., *XMGRACE, Version 5.1.19.* 2005, Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR.

219. Chen, Y.-C., *Beware of docking!* Trends Pharmacol Sci, 2015. **36**(2): p. 78-95.

220. Trott, O. and A.J. Olson, *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.* J Comput Chem, 2010. **31**(2): p. 455-461.

221. Huang, S.Y. and X. Zou, *Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking.* Proteins: Struct, Funct, Bioinf, 2007. **66**(2): p. 399-421.

222. Amaro, R.E., R. Baron, and J.A. McCammon, *An improved relaxed complex scheme for receptor flexibility in computer-aided drug design.* J Comput Aided Mol Des, 2008. **22**(9): p. 693-705.

223. Vargiu, A.V. and H. Nikaido, *Multidrug binding properties of the AcrB efflux pump characterized by molecular dynamics simulations.* Proc Natl Acad Sci U S A, 2012. **109**(50): p. 20637-20642.

224. Collu, F., et al., *Recognition of imipenem and meropenem by the RND-transporter MexB studied by computer simulations.* J Am Chem Soc, 2012. **134**(46): p. 19146-19158.

225. Vargiu, A.V., et al., *Inhibition of E. coli AcrB multidrug efflux pump by MBX2319: molecular mechanism and comparison with other inhibitors.* Antimicrob Agents Chemother, 2014. **58**(10): p. 6224-6234.

226. Takatsuka, Y., C. Chen, and H. Nikaido, *Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of Escherichia coli.* Proc Natl Acad Sci U S A, 2010. **107**(15): p. 6559-6565.

227. Sousa, S.F., P.A. Fernandes, and M.J. Ramos, *Protein–ligand docking: current status and future challenges.* Proteins: Struct, Funct, Bioinf, 2006. **65**(1): p. 15-26.

228. Knapp, B., et al., *Is an intuitive convergence definition of molecular dynamics simulations solely based on the root mean square deviation possible?* J Comput Biol, 2011. **18**(8): p. 997-1005.

229. Karpen, M.E., D.J. Tobias, and C.L. Brooks III, *Statistical clustering techniques for the analysis of long molecular dynamics trajectories: analysis of 2.2-ns trajectories of YPGDV.* Biochemistry, 1993. **32**(2): p. 412-420.

230. Shao, J., et al., *Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms.* J Chem Theory Comput, 2007. **3**(6): p. 2312-2334.

231. Ruiz-Carmona, S., et al., *rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids.* PLoS Comput Biol, 2014. **10**(4): p. e1003571.

232. Morley, S.D. and M. Afshar, *Validation of an empirical RNA-ligand scoring function for fast flexible docking using RiboDock®.* J Comput Aided Mol Des, 2004. **18**(3): p. 189-208.

233. Afshar, M., et al., *Investigating the high affinity and low sequence specificity of calmodulin binding to its targets.* J Mol Biol, 1994. **244**(5): p. 554-571.

234. Ghose, A.K. and G.M. Crippen, *Use of physicochemical parameters in distance geometry and related three-dimensional quantitative structure-activity relationships: a demonstration using Escherichia coli dihydrofolate reductase inhibitors.* J Med Chem, 1985. **28**(3): p. 333-346.

235. Ghose, A.K. and G.M. Crippen, *Atomic physicochemical parameters for three-dimensional structure-directed quantitative structure-activity relationships I. Partition coefficients as a measure of hydrophobicity.* J Comput Chem, 1986. **7**(4): p. 565-577.

236. Gaillard, P., et al., *Molecular lipophilicity potential, a tool in 3D QSAR: method and applications.* J Comput Aided Mol Des, 1994. **8**(2): p. 83-96.

237. Oberhauser, N., A. Nurisso, and P.-A. Carrupt, *MLP Tools: a PyMOL plugin for using the molecular lipophilicity potential in computer-aided drug design.* J Comput Aided Mol Des, 2014. **28**(5): p. 587-596.

238. Harris, R.C., et al., *Opposites attract: shape and electrostatic complementarity in protein-DNA complexes.* Innovations in Biomolecular Modeling and Simulations. Vol. 2. 2012. 53-80.

239. Baker, N.A., et al., *Electrostatics of nanosystems: application to microtubules and the ribosome.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10037-10041.

240. Dolinsky, T.J., et al., *PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations.* Nucleic Acids Res, 2004. **32**(suppl 2): p. W665-W667.

241. Zakrzewska, K., R. Lavery, and B. Pullman, *The solvation contribution to the binding energy of DNA with non-intercalating antibiotics.* Nucleic Acids Res, 1984. **12**(16): p. 6559-6574.

242. Kulińska, K., et al., *Spatial distribution functions as a tool in the analysis of ribonucleic acids hydration—molecular dynamics studies.* Comput Chem, 2000. **24**(3): p. 451-457.

243. Abraham, M., et al., *GROMACS User Manual version 5.0.* 2014.

244. Van der Spoel, D., et al., *GROMACS User Manual version.* 2013.

245. Woo, H.-J. and B. Roux, *Calculation of absolute protein–ligand binding free energy from computer simulations.* Proc Natl Acad Sci U S A, 2005. **102**(19): p. 6825-6830.

246. Gilson, M.K. and H.-X. Zhou, *Calculation of protein-ligand binding affinities.* Annu Rev Biophys Biomol Struct, 2007. **36**(1): p. 21.

247. Steinbrecher, T. and A. Labahn, *Towards accurate free energy calculations in ligand protein-binding studies.* Curr Med Chem, 2010. **17**(8): p. 767-785.

248. Jiao, D., et al., *Calculation of protein–ligand binding free energy by using a polarizable potential.* Proc Natl Acad Sci U S A, 2008. **105**(17): p. 6290-6295.

249. Mitomo, D., et al., *Calculation of protein-ligand binding free energy using smooth reaction path generation (SRPG) method: a comparison of the explicit water model, GB/SA model and docking score function*, in *Genome Inform.* 2009. p. 85-97.

250. Rathore, R., et al., *Advances in binding free energies calculations: QM/MM-based free energy perturbation method for drug design.* Curr Pharm Des, 2013. **19**(26): p. 4674-4686.

251. Fogolari, F., A. Brigo, and H. Molinari, *Protocol for MM/PBSA molecular dynamics simulations of proteins.* Biophys J, 2003. **85**(1): p. 159-166.

252. Hou, T., et al., *Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations.* J Chem Inf Model, 2010. **51**(1): p. 69-82.

253. Srinivasan, J., et al., *Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices.* J Am Chem Soc, 1998. **120**(37): p. 9401-9409.

254. Kollman, P.A., et al., *Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models.* Acc Chem Res, 2000. **33**(12): p. 889-897.

255. Kongsted, J. and U. Ryde, *An improved method to predict the entropy term with the MM/PBSA approach.* J Comput Aided Mol Des, 2009. **23**(2): p. 63-71.

256. Massova, I. and P.A. Kollman, *Computational alanine scanning to probe protein-protein interactions: a novel approach to evaluate binding free energies.* J Am Chem Soc, 1999. **121**(36): p. 8133-8143.

257. Nishino, K. and A. Yamaguchi, *Analysis of a complete library of putative drug transporter genes in Escherichia coli.* J Bacteriol, 2001. **183**(20): p. 5803-5812.

258. Sulavik, M.C., et al., *Antibiotic Susceptibility Profiles ofEscherichia coli Strains Lacking Multidrug Efflux Pump Genes.* Antimicrob Agents Chemother, 2001. **45**(4): p. 1126-1136.

259. Opperman, T.J., et al., *Characterization of a novel pyranopyridine inhibitor of the AcrAB efflux pump of Escherichia coli.* Antimicrob Agents Chemother, 2014. **58**(2): p. 722-733.

260. Nishino, K., E. Nikaido, and A. Yamaguchi, *Regulation and physiological function of multidrug efflux pumps in Escherichia coli and Salmonella.* Biochim Biophys Acta Proteins Proteomics, 2009. **1794**(5): p. 834-843.

261. Li, X.-Z., P. Plésiat, and H. Nikaido, *The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria.* Clin Microbiol Rev, 2015. **28**(2): p. 337-418.

262. Schuster, S., M. Vavra, and W.V. Kern, *Evidence of a substrate discriminating entrance channel in the lower porter domain of the multidrug resistance efflux pump AcrB.* Antimicrob Agents Chemother, 2016: p. AAC. 00314-16.

263. Piddock, L.J., *Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria.* Clin Microbiol Rev, 2006. **19**(2): p. 382-402.

264. Ruggerone, P., et al., *RND efflux pumps: structural information translated into function and inhibition mechanisms.* Curr Top Med Chem, 2013. **13**(24): p. 3079-3100.

265. Consortium, U., *UniProt: a hub for protein information.* Nucleic Acids Res, 2014. **43**(D1): p. D204-D212.

266. Berman, H.M., et al., *The protein data bank.* Nucleic Acids Res, 2000. **28**(1): p. 235-242.

267. Shen, M.y. and A. Sali, *Statistical potential for assessment and prediction of protein structures.* Protein Sci, 2006. **15**(11): p. 2507-2524.

268. Ramachandran, G.N., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations.* J Mol Biol, 1963. **7**(1): p. 95-99.

269. Colovos, C. and T. Yeates, *ERRAT: an empirical atom-based method for validating protein structures.* Protein Sci, 1993. **2**: p. 1511-1519.

270. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(1): p. 12-21.

271. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics.* J Mol Graph, 1996. **14**(1): p. 33-38.

272. Schrödinger, L., *The PyMOL Molecular Graphics System. Version 1.5.*

273. Keil, M., T.E. Exner, and J. Brickmann, *Pattern recognition strategies for molecular surfaces: III. Binding site prediction with a neural network.* J Comput Chem, 2004. **25**(6): p. 779-789.

274. Abraham, M.J., et al., *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. **1**: p. 19-25.

275. Opperman, T.J. and S.T. Nguyen, *Recent advances toward a molecular mechanism of efflux pump inhibition.* Front Microbiol, 2015. **6**: p. 421.

276. Blair, J.M., et al., *AcrB drug-binding pocket substitution confers clinically relevant resistance and altered substrate specificity.* Proc Natl Acad Sci U S A, 2015. **112**(11): p. 3511-3516.

277. Nakashima, R., et al., *Structural basis for the inhibition of bacterial multidrug exporters.* Nature, 2013. **500**(7460): p. 102-106.

278. Fischer, N. and C. Kandt, *Porter domain opening and closing motions in the multi-drug efflux transporter AcrB.* Biochim Biophys Acta Biomembr, 2013. **1828**(2): p. 632-641.

279. Imai, T., et al., *Functionality mapping on internal surfaces of multidrug transporter AcrB based on molecular theory of solvation: implications for drug efflux pathway.* J Phys Chem B, 2011. **115**(25): p. 8288-8295.

280. Levy, Y., J.N. Onuchic, and P.G. Wolynes, *Fly-casting in protein-DNA binding: frustration between protein folding and electrostatics facilitates target recognition.* J Am Chem Soc, 2007. **129**(4): p. 738-739.

281. Selzer, T., S. Albeck, and G. Schreiber, *Rational design of faster associating and tighter binding protein complexes.* Nat Struct Mol Biol, 2000. **7**(7): p. 537-541.

282. Zgurskaya, H.I. and H. Nikaido, *Multidrug resistance mechanisms: drug efflux across two membranes.* Mol Microbiol, 2000. **37**(2): p. 219-225.

283. Fischbach, M.A. and C.T. Walsh, *Antibiotics for emerging pathogens.* Science, 2009. **325**(5944): p. 1089-1093.

284. Poole, K., *Pseudomonas aeruginosa: resistance to the max.* Front Microbiol, 2011. **2**(65): p. 90-102.

285. Nikaido, H., *Prevention of drug access to bacterial targets: permeability barriers and active efflux.* Science, 1994. **264**(5157): p. 382-388.

286. Hancock, R.E., *Resistance mechanisms in Pseudomonas aeruginosa and other nonfermentative gram-negative bacteria.* Clin Infect Dis, 1998. **27**(Supplement 1): p. S93-S99.

287. Nikaido, H., *Molecular basis of bacterial outer membrane permeability revisited.* Microbiol Mol Biol Rev, 2003. **67**(4): p. 593-656.

288. Nikaido, H. and R. Hancock, *Outer membrane permeability of Pseudomonas aeruginosa.* The Bacteria, A Treatise on Structure and Function, ed. J. Sokatch. Vol. 10. 2012: Elsevier. 145-93.

289. Poole, K., *Multidrug efflux pumps and antimicrobial resistance in Pseudomonas aeruginosa and related organisms.* J Mol Microbiol Biotechnol, 2001. **3**(2): p. 255-264.

290. Poole, K. and R. Srikumar, *Multidrug Efflux in Pseudomonas aeruginosa Components, Mechanisms and Clinical Significance.* Curr Top Med Chem, 2001. **1**(1): p. 59-71.

291. Poole, K., *Efflux-mediated resistance to fluoroquinolones in gram-negative bacteria.* Antimicrob Agents Chemother, 2000. **44**(9): p. 2233-2241.

292. Poole, K., *Pseudomonas aeruginosa efflux pumps*. Microbial Efflux Pumps: Current Research. 2013: Caister Academic Press. 175-206.

293. Poole, K., *Efflux-mediated multiresistance in Gram-negative bacteria*. Clin Microbiol Infect, 2004. **10**(1): p. 12-26.

294. Poole, K., *Efflux-mediated antimicrobial resistance*. J Antimicrob Chemother, 2005. **56**(1): p. 20-51.

295. Webber, M. and L. Piddock, *The importance of efflux pumps in bacterial antibiotic resistance*. J Antimicrob Chemother, 2003. **51**(1): p. 9-11.

296. Zechini, B. and I. Versace, *Inhibitors of multidrug resistant efflux systems in bacteria*. Recent Patents Anti-Infect Drug Disc, 2009. **4**(1): p. 37-50.

297. Fernández, L. and R.E. Hancock, *Adaptive and mutational resistance: role of porins and efflux pumps in drug resistance*. Clin Microbiol Rev, 2012. **25**(4): p. 661-681.

298. Delmar, J.A., C.-C. Su, and W.Y. Edward, *Bacterial multi-drug efflux transporters*. Annu Rev Biophys, 2014. **43**: p. 93.

299. Sun, J., Z. Deng, and A. Yan, *Bacterial multidrug efflux pumps: Mechanisms, physiology and pharmacological exploitations*. Biochem Biophys Res Commun, 2014. **453**(2): p. 254-267.

300. Gotoh, N., et al., *The outer membrane protein OprM of Pseudomonas aeruginosa is encoded by oprK of the mexA-mexB-oprK multidrug resistance operon*. Antimicrob Agents Chemother, 1995. **39**(11): p. 2567-2569.

301. Li, X.-Z., H. Nikaido, and K. Poole, *Role of mexA-mexB-oprM in antibiotic efflux in Pseudomonas aeruginosa*. Antimicrob Agents Chemother, 1995. **39**(9): p. 1948-1953.

302. Poole, K., et al., *Multiple antibiotic resistance in Pseudomonas aeruginosa: evidence for involvement of an efflux operon*. J Bacteriol, 1993. **175**(22): p. 7363-7372.

303. Mine, T., et al., *Expression in Escherichia coli of a new multidrug efflux pump, MexXY, from Pseudomonas aeruginosa*. Antimicrob Agents Chemother, 1999. **43**(2): p. 415-417.

304. Westbrock-Wadman, S., et al., *Characterization of a Pseudomonas aeruginosa efflux pump contributing to aminoglycoside impermeability*. Antimicrob Agents Chemother, 1999. **43**(12): p. 2975-2983.

305. Lee, A., et al., *Interplay between efflux pumps may provide either additive or multiplicative effects on drug resistance*. J Bacteriol, 2000. **182**(11): p. 3142-3150.

306. Llanes, C., et al., *Clinical strains of Pseudomonas aeruginosa overproducing MexAB-OprM and MexXY efflux pumps simultaneously*. Antimicrob Agents Chemother, 2004. **48**(5): p. 1797-1802.

307. Poole, K., *Efflux-mediated antimicrobial resistance*, in *Antibiotic Discovery and Development*. 2012, Springer. p. 349-395.

308. Krahn, T., et al., *Determinants of intrinsic aminoglycoside resistance in Pseudomonas aeruginosa*. Antimicrob Agents Chemother, 2012: p. AAC. 01446-12.

309. Srikumar, R., X.-Z. Li, and K. Poole, *Inner membrane efflux components are responsible for beta-lactam specificity of multidrug efflux pumps in Pseudomonas aeruginosa*. J Bacteriol, 1997. **179**(24): p. 7875-7881.

310. Murata, T., et al., *The substrate specificity of tripartite efflux systems of Pseudomonas aeruginosa is determined by the RND component*. Biochem Biophys Res Commun, 2002. **299**(2): p. 247-251.

311. Eda, S., H. Yoneyama, and T. Nakae, *Function of the MexB efflux-transporter divided into two halves*. Biochemistry, 2003. **42**(23): p. 7238-7244.

312. Köhler, T., et al., *Multidrug efflux in intrinsic resistance to trimethoprim and sulfamethoxazole in Pseudomonas aeruginosa*. Antimicrob Agents Chemother, 1996. **40**(10): p. 2288-2290.

313. Okamoto, K., N. Gotoh, and T. Nishino, *Pseudomonas aeruginosa reveals high intrinsic resistance to penem antibiotics: penem resistance mechanisms and their interplay.* Antimicrob Agents Chemother, 2001. **45**(7): p. 1964-1971.

314. Zhao, Q., et al., *Influence of the TonB energy-coupling protein on efflux-mediated multidrug resistance in Pseudomonas aeruginosa.* Antimicrob Agents Chemother, 1998. **42**(9): p. 2225-2231.

315. Ziha-Zarifi, I., et al., *In Vivo Emergence of Multidrug-Resistant Mutants of Pseudomonas aeruginosa Overexpressing the Active Efflux System MexA-MexB-OprM.* Antimicrob Agents Chemother, 1999. **43**(2): p. 287-291.

316. Mesaros, N., et al., *Pseudomonas aeruginosa: resistance and therapeutic options at the turn of the new millennium.* Clin Microbiol Infect, 2007. **13**(6): p. 560-578.

317. Hocquet, D., et al., *MexXY-OprM efflux pump is necessary for adaptive resistance of Pseudomonas aeruginosa to aminoglycosides.* Antimicrob Agents Chemother, 2003. **47**(4): p. 1371-1375.

318. Masuda, N., et al., *Contribution of the MexX-MexY-OprM efflux system to intrinsic resistance in Pseudomonas aeruginosa.* Antimicrob Agents Chemother, 2000. **44**(9): p. 2242-2246.

319. Kinana, A.D., A.V. Vargiu, and H. Nikaido, *Effect of site-directed mutations in multidrug efflux pump AcrB examined by quantitative efflux assays.* Biochem Biophys Res Commun, 2016. **480**(4): p. 552-557.

320. Bohnert, J.A., et al., *Site-directed mutagenesis reveals putative substrate binding residues in the Escherichia coli RND efflux pump AcrB.* J Bacteriol, 2008. **190**(24): p. 8225-8229.

321. Edward, W.Y., J.R. Aires, and H. Nikaido, *AcrB multidrug efflux pump of Escherichia coli: composite substrate-binding cavity of exceptional flexibility generates its extremely wide substrate specificity.* J Bacteriol, 2003. **185**(19): p. 5657-5664.

322. Nikaido, H. and Y. Takatsuka, *Mechanisms of RND multidrug efflux pumps.* Biochim Biophys Acta Proteins Proteomics, 2009. **1794**(5): p. 769-781.

323. Marsh, L., *Strong ligand-protein interactions derived from diffuse ligand interactions with loose binding sites.* BioMed Res Int, 2015. **2015**.

324. Biela, A., et al., *Ligand binding stepwise disrupts water network in thrombin: enthalpic and entropic changes reveal classical hydrophobic effect.* J Med Chem, 2012. **55**(13): p. 6094-6110.

325. Vargiu, A.V., et al., *Molecular mechanism of MBX2319 inhibition of Escherichia coli AcrB multidrug efflux pump and comparison with other inhibitors.* Antimicrob Agents Chemother, 2014. **58**(10): p. 6224-6234.

326. Ruggerone, P., et al., *Molecular dynamics computer simulations of multidrug RND efflux pumps.* Comput Struct Biotechnol J, 2013. **5**(6): p. 1-11.

327. Collu, F. and M. Cascella, *Multidrug resistance and efflux pumps: insights from molecular dynamics simulations.* Curr Top Med Chem, 2013. **13**(24): p. 3165-3183.

328. Malloci, G., et al., *A database of force-field parameters, dynamics, and properties of antimicrobial compounds.* Molecules, 2015. **20**(8): p. 13997-14021.

329. Alvarez-Garcia, D. and X. Barril, *Relationship between protein flexibility and binding: Lessons for structure-based drug design.* J Chem Theory Comput, 2014. **10**(6): p. 2608-2614.

330. Sennhauser, G., et al., *Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors.* PLoS Biol, 2007. **5**(1): p. e7.

331. Morris, G.M., et al., *AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.* J Comput Chem, 2009. **30**(16): p. 2785-2791.

332. Hetényi, C. and D.v.d. Spoel, *Toward prediction of functional protein pockets using blind docking and pocket search algorithms.* Protein Sci, 2011. **20**(5): p. 880-893.

333. Hetényi, C. and D. van der Spoel, *Blind docking of drug-sized compounds to proteins with up to a thousand residues.* FEBS Lett, 2006. **580**(5): p. 1447-1450.

334. Genheden, S. and U. Ryde, *The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities.* Expert Opin Drug Discov, 2015. **10**(5): p. 449-461.

335. Laskowski, R.A. and M.B. Swindells, *LigPlot+: multiple ligand–protein interaction diagrams for drug discovery.* J Chem Inf Model, 2011. **51**(10): p. 2778-2786.

336. Salentin, S., et al., *PLIP: fully automated protein–ligand interaction profiler.* Nucleic Acids Res, 2015. **43**(W1): p. W443-W447.

337. Neyfakh, A.A., *Mystery of multidrug transporters: the answer can be simple.* Mol Microbiol, 2002. **44**(5): p. 1123-1130.

338. Dominguez, C., R. Boelens, and A.M. Bonvin, *HADDOCK: a protein-protein docking approach based on biochemical or biophysical information.* J Am Chem Soc, 2003. **125**(7): p. 1731-1737.

339. van Zundert, G., et al., *The HADDOCK2. 2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes.* J Mol Biol, 2016. **428**(4): p. 720-725.

340. Dolghih, E., et al., *Predicting binding to p-glycoprotein by flexible receptor docking.* PLoS Comput Biol, 2011. **7**(6): p. e1002083.

341. Liu, M., et al., *The flexibility of P-glycoprotein for its poly-specific drug binding from molecular dynamics simulations.* J Biomol Struct Dyn, 2013. **31**(6): p. 612-629.

342. Van Dijk, A.D. and A.M. Bonvin, *Solvated docking: introducing water into the modelling of biomolecular complexes.* Bioinformatics, 2006. **22**(19): p. 2340-2347.

343. Fischer, N. and C. Kandt, *Three ways in, one way out: Water dynamics in the trans-membrane domains of the inner membrane translocase AcrB.* Proteins: Struct, Funct, Bioinf, 2011. **79**(10): p. 2871-2885.

344. Vangone, A., et al., *Sense and Simplicity in HADDOCK Scoring: Lessons from CASP-CAPRI Round1.* Proteins: Struct, Funct, Bioinf, 2016.

345. Janin, J., *Protein–protein docking tested in blind predictions: the CAPRI experiment.* Mol BioSyst, 2010. **6**(12): p. 2351-2362.

346. Lensink, M.F., et al., *Prediction of homoprotein and heteroprotein complexes by protein docking and template-based modeling: A CASP-CAPRI experiment.* Proteins: Struct, Funct, Bioinf, 2016.

347. Peterson, L.X., et al., *Human and server docking prediction for CAPRI round 30-35 using LZerD with combined scoring functions.* Proteins: Struct, Funct, Bioinf, 2016.

348. Jorgensen, W.L. and J. Tirado-Rives, *The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin.* J Am Chem Soc, 1988. **110**(6): p. 1657-1666.

349. Fernandez-Recio, J., M. Totrov, and R. Abagyan, *Identification of protein–protein interaction sites from docking energy landscapes.* J Mol Biol, 2004. **335**(3): p. 843-865.

350. Van Aalten, D.M., et al., *PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules.* J Comput Aided Mol Des, 1996. **10**(3): p. 255-262.

351. SchuÈttelkopf, A.W. and D.M. Van Aalten, *PRODRG: a tool for high-throughput crystallography of protein–ligand complexes.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(8): p. 1355-1363.

352. Meyer, E.A., R.K. Castellano, and F. Diederich, *Interactions with aromatic rings in chemical and biological recognition.* Angew Chem, Int Ed, 2003. **42**(11): p. 1210-1250.

353. Hunter, C.A., et al., *Aromatic interactions.* J Chem Soc, Perkin Trans 2, 2001(5): p. 651-669.

354. Edward, W.Y., et al., *Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump.* Science, 2003. **300**(5621): p. 976-980.

355.   Wang, J., et al., *Development and testing of a general amber force field.* J Comput Chem, 2004. **25**(9): p. 1157-1174.
356.   da Silva, A.W.S. and W.F. Vranken, *ACPYPE-Antechamber python parser interface.* BMC Res Notes, 2012. **5**(1): p. 1.
357.   De Vries, S.J., M. van Dijk, and A.M. Bonvin, *The HADDOCK web server for data-driven biomolecular docking.* Nat Protoc, 2010. **5**(5): p. 883-897.